Active-site variants of *Streptomyces griseus* protease B with peptide-ligation activity

R Jeffery Elliott[1], Andrew J Bennet[2], Curtis A Braun[1]*, Alasdair M MacLeod[1] and Thor J Borgford[1]

Background: Peptide-ligating technologies facilitate a range of manipulations for the study of protein structure and function that are not possible using conventional genetic or mutagenic methods. To different extents, the currently available enzymatic and nonenzymatic methodologies are synthetically demanding, sequence-dependent and/or sensitive to denaturants. No single coupling method is universally applicable. Accordingly, new strategies for peptide ligation are sought.

Results: Site-specific variants (Ser195→Gly, S195G, and Ser195→Ala, S195A) of *Streptomyces griseus* protease B (SGPB) were generated that efficiently catalyze peptide ligation (i.e., aminolysis of ester-, thioester- and *para*-nitroanilide-activated peptides). The variants also showed reduced hydrolytic activity relative to the wild-type enzyme. The ratio of aminolysis to hydrolysis was greater for the S195A variant, which was also capable of catalyzing ligation in concentrations of urea as high as 2 M.

Conclusions: Mutagenic substitution of the active-site serine residue of SGPB by either glycine or alanine has created a unique class of peptide-ligating catalysts that are useful for coupling relatively stable ester- and *para*-nitroanilide-activated substrates. Ligation proceeds through an acyl–enzyme intermediate involving His57. Serine to alanine mutations may provide a general strategy for converting proteases with chymotrypsin-like protein folds into peptide-coupling enzymes.

Addresses: [1]Institute of Molecular Biology and Biochemistry and [2]Department of Chemistry, Simon Fraser University, Burnaby, BC, Canada, V5A 1S6.

*Present address: De Novo Enzyme Corporation, 8081 Lougheed Highway, Burnaby, BC, Canada, V5A 1W9.

Correspondence: Thor J Borgford
E-mail: borgford@sfu.ca

Key words: non-natural amino acids, peptide ligation, protein engineering, semi-synthesis, *Streptomyces griseus* protease B

Received: **21 December 1999**
Accepted: **13 January 2000**

Published: **15 February 2000**

Chemistry & Biology 2000, **7**:163–171

1074-5521/00/$ – see front matter

Introduction

Gene synthesis and oligonucleotide-directed mutagenesis techniques make possible virtually any variation to the primary sequence of proteins [1,2]. These site-specific changes in primary sequence have proved, in turn, to be powerful methods for dissecting relationships between protein structure and function. Notably, these methods are also a means for making useful modifications to protein activity [3,4]. Although powerful and useful, conventional molecular biology techniques are restricted to the use of the 20 amino acids dictated by the genetic code. A far wider range of manipulations to protein structure and activity is conceivable if non-natural amino acids are used in protein engineering [5–8].

Solid-phase synthesis permits the routine incorporation of non-natural amino acids into short peptides [9]. Such methods are currently too costly and synthetically challenging to permit the routine synthesis of entire proteins containing novel amino acids. Using a ligation strategy, peptides produced chemically, recombinantly or by partial proteolysis can be spliced together to create full-length proteins. Hence, non-natural amino acids may be site-specifically incorporated into semi-synthetic proteins [10].

Existing methods of peptide ligation have been used to generate proteins with a wide variety of novel features, including unnatural catalytic residues, nonpeptidic backbones, and unique conformational probes [11–15]. Several enzymatic and nonenzymatic peptide-coupling strategies are currently in use; virtually all are performed with unprotected peptides and in aqueous phase. The formation of a peptide bond from free amines and carboxylic acids is, however, thermodynamically unfavourable in water. Substrate peptides are therefore chemically pre-activated and reaction conditions designed to allow the products of ligation to be 'kinetically trapped' [16].

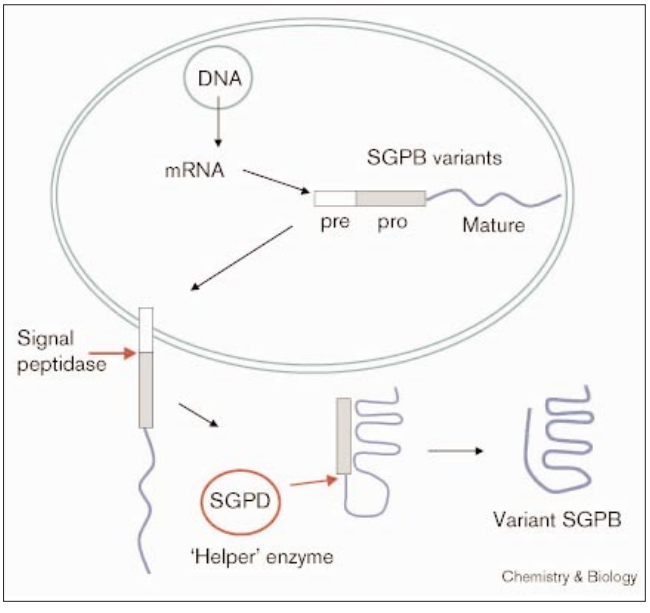
Kent and co-workers [17] have described an elegant nonenzymatic method for ligating peptides in aqueous phase. Donor peptides, activated with carboxy-terminal thioesters, are reacted with acceptor peptides that have unblocked amino-terminal cysteine residues. Coupling proceeds through a transient thioester intermediate (between the cysteine thiol and the carbonyl carbon of the activated peptide), which undergoes a spontaneous intramolecular S→N acyl transfer to form a native peptide bond. The development of methods to generate recombinant proteins activated with carboxy-terminal thioesters

has made this technology more accessible [18–23], and a variety of proteins containing non-natural amino acids have been generated [15,22,24]. Major disadvantages of the technology are the relative lability of thioesters and the sequence restrictions (X-Cys, X-Gly, Gly-X) at the junction site [25].

The potential advantage of enzyme-catalyzed ligation technology is that reactions can be performed using a variety of activated leaving groups on the donor peptide (i.e., not solely thioesters). Proteases (or protease variants) can ensure the regiospecific coupling of amino- and carboxy-terminal ends of reacting peptides, and consequently are less dependent on a particular leaving-group chemistry. Furthermore, many proteases have been characterized, covering a wide range of substrate specificities, thereby presenting the possibility of coupling virtually any sequence of amino acids. Fundamental objectives in developing protease-based methods for peptide ligation are to find mutant enzymes that can couple a broad range of sequences using activated peptides that resist uncatalyzed hydrolysis.

Through multiple mutations, Wells and co-workers [14,26,27] converted the protease subtilisin BPN′ into the peptide-ligating catalyst referred to as 'subtiligase'. Subtiligase contains a catalytic serine to cysteine mutation (S221C), which greatly reduces the amidase activity of the enzyme, rendering it incapable of hydrolyzing the products of peptide-coupling reactions. Nevertheless, the enzyme can be efficiently acylated by glycolate-ester-activated peptides. Attack on the resulting acyl–enzyme intermediate is more facile for amine nucleophiles than for water, and therefore aminolysis (ligation) is favored over hydrolysis. Subtiligase variants with increased activities and oxidative stabilities have also been selected using phage display [28]. To date, subtiligase and its derivatives are the only reported genetic enzyme variants active in peptide coupling.

Figure 1



Production of SGPB variants from *B. subtilis* strain DB104. The enzymes are produced as pre-pro-mature proteins. The pre- region directs the protein to the cell surface for secretion, where it is cleaved from the pro-mature region by signal peptidase. A 'helper' enzyme, SGPD, which has the same specificity as SGPB, cleaves the mature protein from the pro- region releasing it into the culture supernatant.

In the current study, mutants of a chymotrypsin-like serine protease (*Streptomyces griseus* protease B, SGPB) were investigated for their ability to couple peptides. Enzymes of the chymotrypsin family of serine proteases have a protein fold that is distinct from the subtilisin family [29]. A diversity of chymotrypsin-like enzymes have been described, including those with primary specificity for basic and acidic amino acids [29–31].

Results and discussion

Expression of SGPB variants

SGPB is a relatively small (18.6 kDa) monomeric, secreted bacterial enzyme of the chymotrypsin clan of serine proteases, containing a catalytic triad composed of His57, Asp102 and Ser195 [29,32,33]. SGPB has primary specificity for peptide bonds following large hydrophobic amino acids [34,35]. The protein contains two disulfide bonds and is extremely stable, with activity in high concentrations of denaturants (6 M urea and 8 M guanidine HCl (GnHCl)) and at high temperatures [36,37]. These are potentially valuable properties in a peptide ligase, as the semi-synthesis of proteins from relatively long peptides is likely to involve denaturing conditions.

Two categories of mutation were investigated. Firstly, the active-site serine residue of SGPB was mutated to cysteine, in the expectation that an enzyme with subtiligase-like activity would be generated. Secondly, a category of mutants was made in which the active-site serine residue was replaced with either glycine or alanine. We postulated that variant enzymes in the second category would act as a platform for simultaneous binding of donor and acceptor peptides, allowing the direct nucleophilic attack of the acceptor amine on the carbonyl group of the activated donor. Variants would contribute specificity through substrate binding energy, and the negatively charged transition state would be stabilized by the oxyanion hole of SGPB [38].

In total, four genetic variants of SGPB were constructed: Ser195 to cysteine (S195C), alanine (S195A), or glycine (S195G), and a double mutant Ser195 to glycine/His57 to asparagine (S195G/H57N). Variant genes were cloned in

the *Escherichia coli*/*Bacillus subtilis* shuttle vector pEB-11 and expressed in *B. subtilis* DB104 [31].

Wild-type SGPB is normally translated as a pre-pro-mature enzyme. Subsequently, the zymogen is processed at the pre-pro junction by a host-encoded signal peptidase and processed at the pro-mature junction autocatalytically [30,39]. As the mutants are poor amidases, a 'helper' enzyme was required to perform the pro-mature processing (Figure 1). *Streptomyces griseus* protease D (SGPD) has similar specificity to SGPB [30] and was therefore used to perform the pro-mature processing event. Significantly, SGPD differs from SGPB in terms of its size (SGPD ≅ 36 kDa versus SGPB ≅ 18 kDa) and isoelectric point (SGPB has a basic pI whereas SGPD has an acidic pI) [30]. Mutants can, therefore, be readily separated from the helper enzyme by ion-exchange chromatography.

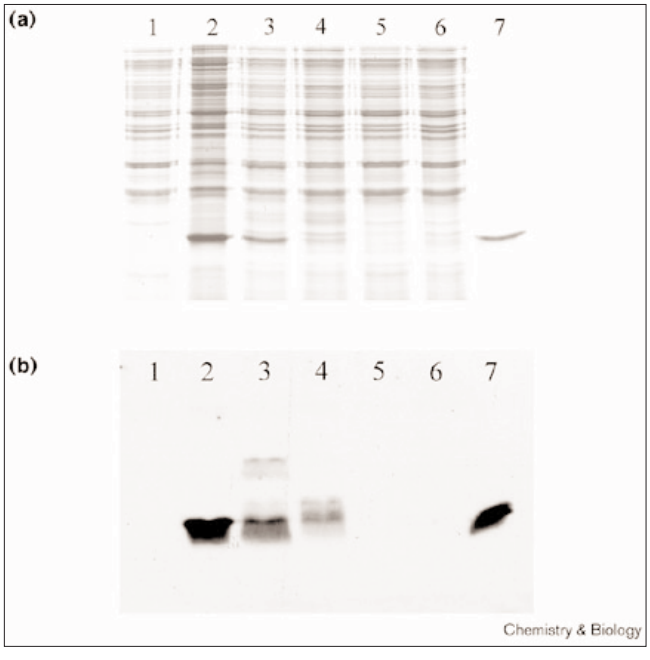
S195A and S195G proteins were expressed and secreted at ~1 mg/l by recombinant *Bacillus*; however, several attempts to express protein from S195C and S195G/H57N recombinants proved unsuccessful (Figure 2). Further studies of these two variants were, therefore, not possible. The S195A variant was produced in greater quantities than the S195G variant, but at lower levels than the wild-type enzyme. It is possible that the levels of variant expression, and the absence of S195C and S195G/H57N proteins, are related to nonproductive folding events. A relationship between the folding and expression of SGPB was observed previously [40]. The S195G and S195A variants were purified to homogeneity, and the serine to glycine and serine to alanine mutations confirmed by mass spectrometry.

Ligation and hydrolysis activity of SGPB variants

Purified SGPB variants were tested for their ability to catalyze reactions between activated 'donor' peptides (i.e., peptides having carboxy-terminal ester, thioester or *para*-nitroanilide (-*p*NA) groups) and nucleophilic 'acceptor' peptides (i.e., peptides having a free amino-terminal amine).

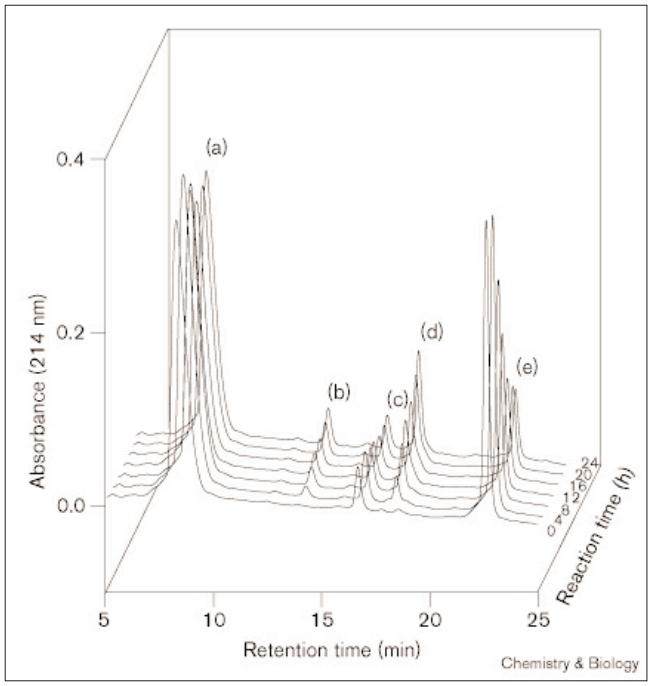
A time course for the S195A-catalyzed ligation of two peptides is shown in Figure 3. The S195G variant was also active in peptide ligation; however, it showed a lower aminolysis to hydrolysis ratio compared with the S195A mutant (Figure 4). Consequently, the S195G variant was not further characterized in ligation reactions. Ligation reactions show saturation kinetics (Figure 4) and do not

Figure 2



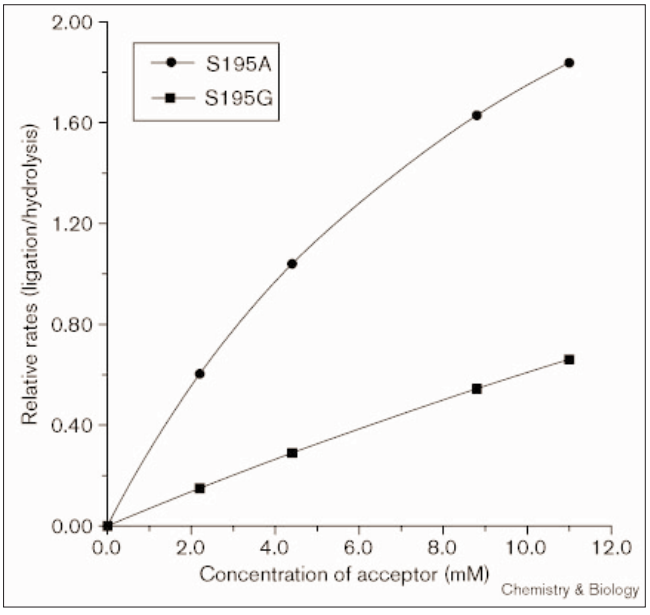
Expression of SGPB variants. **(a)** Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and **(b)** western blot of small-scale expression tests of the SGPB variants. Lanes 1, pEB-11 (vector only); lanes 2, pEB-B8 (vector encoding wild-type SGPB); lanes 3, pEB-HDa (vector encoding S195A); lanes 4, pEB-HDg (vector encoding S195G); lanes 5, pEB-HDc (vector encoding S195C); lanes 6, pEB-nDg (vector encoding S195G/H57N); and lanes 7, purified SGPB.

Figure 3



Time course of the S195A-catalyzed ligation reaction, analyzed by reversed phase high performance liquid chromatography (HPLC). Peaks are assigned as follows: **(a)** 8 min (F-A-A-S-R-NH_2, nucleophile); **(b)** 14 min (*N*-succinyl-A-A-A-P-F-OH, hydrolysis product); **(c)** 16.5 min (1,4-butanediol); **(d)** 19 min (*N*-succinyl-A-A-P-F-F-A-A-S-R-NH2, ligated product); and **(e)** 23 min (*N*-succcinyl-A-A-P-F-SBz, donor peptide).

Figure 4



Ligation to hydrolysis ratios for the S195A- and S195G-catalyzed ligation reactions as a function of nucleophile (NH_2-F-A-A-S-R-NH_2) concentration. The standard ligation reactions were performed for 16 h, with varying amounts of the nucleophilic peptide. Data were fit to a model of saturation kinetics.

occur in the absence of enzyme. Notably, significant uncatalyzed hydrolysis of thioester-activated donor peptides took place during the ligation reaction, whereas uncatalyzed hydrolysis of *p*NA-activated donors was not observed — underscoring the advantage of working with more stable donor peptides.

Kinetic studies indicate that both ligase variants have reduced amidase activities compared with the wild-type enzyme (Table 1), an anticipated consequence of the replacement of Ser195. Interestingly, the serine to alanine mutation in SGPB (S195A) reduced the hydrolytic activity of the enzyme (k_{cat}/K_m) by a factor of 10^4, whereas the analogous substitution in subtilisin (S221A) resulted in a 10^7-fold reduction [41].

A minimal kinetic scheme for reactions catalyzed by the S195A variant is presented in Figure 5 and justified in the discussion below. Significantly, the reaction kinetics are consistent with a catalytic mechanism involving an enzyme-bound intermediate.

Mechanism of ligation: evidence for an acyl-imidazole intermediate

The hydrolysis reactions catalyzed by S195A (k_{cat}/K_m) differ according to the leaving group of the donor peptides, with thioesters (thiomethyl and thiobenzyl) reacting more rapidly than esters, which react more rapidly than

Table 1

Kinetic constants for the hydrolysis of *N*-succinyl-A-A-P-F-*p*NA by wild-type SGPB, S195G and S195A.

	k_{cat} (s^{-1})	K_m (M)	k_{cat}/K_m ($s^{-1}\,M^{-1}$)
SGPB*	31.4	6.2×10^{-5}	5.0×10^{5}
S195G†	1.1×10^{-3}	3.0×10^{-4}	3.7×10^{0}
S195A‡	7.9×10^{-3}	5.0×10^{-4}	1.6×10^{1}

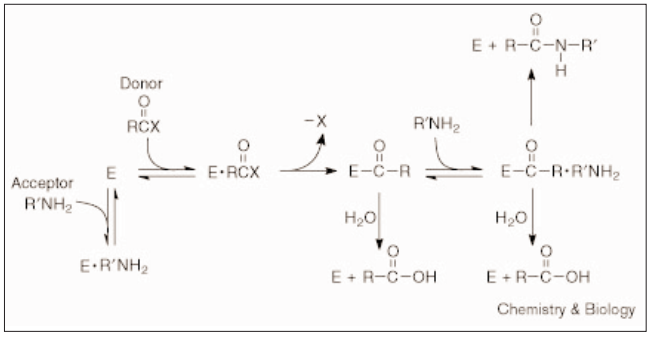
*From [47]. †Errors were estimated at 30%. ‡Errors were estimated at 15%.

para-nitroanilides. The product ratio (i.e., the ratio of aminolysis to hydrolysis at a given acceptor concentration) remains the same regardless of the donor peptide. These two observations, as explained in detail by Fersht [42], provide very strong evidence that the coupling reactions all proceed via a common enzyme-bound intermediate [43].

Formation of an enzyme intermediate can be monitored by the release of *para*-nitroaniline from a *p*NA-activated donor peptide. Moreover, for reactions that involve an enzyme-bound intermediate, the rate of *p*NA release should be independent of the concentration of acceptor peptide. The alternative possibility that an intermediate does not exist requires that the coupling rate, as well as the rate of *p*NA release, increase as acceptor concentration increases.

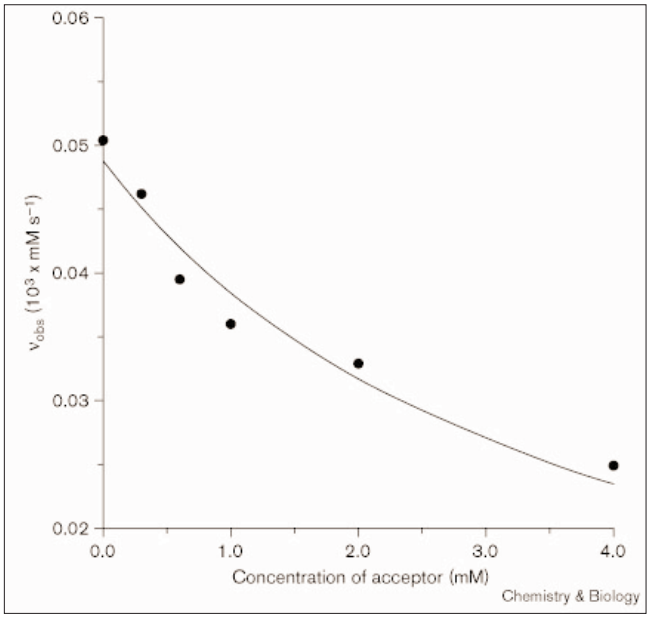
The ratio of coupling to hydrolysis was observed to increase with increasing acceptor concentration (see Figure 4), yet a simultaneous decrease in the rate of *p*NA release was also observed (Figure 6). As *p*-NA release and ligated product formation clearly have different dependencies on nucleophile concentration, a minimum of two steps are involved in ligation. Consequently, the reaction mechanism involves an enzyme-bound intermediate (Figure 5). The decrease in the rate of *p*NA release at high acceptor concentrations was fit to a competitive inhibition model (Figure 6), thereby suggesting competition between donor and acceptor peptides for the donor (P_1)-binding site (Figure 5).

Figure 5



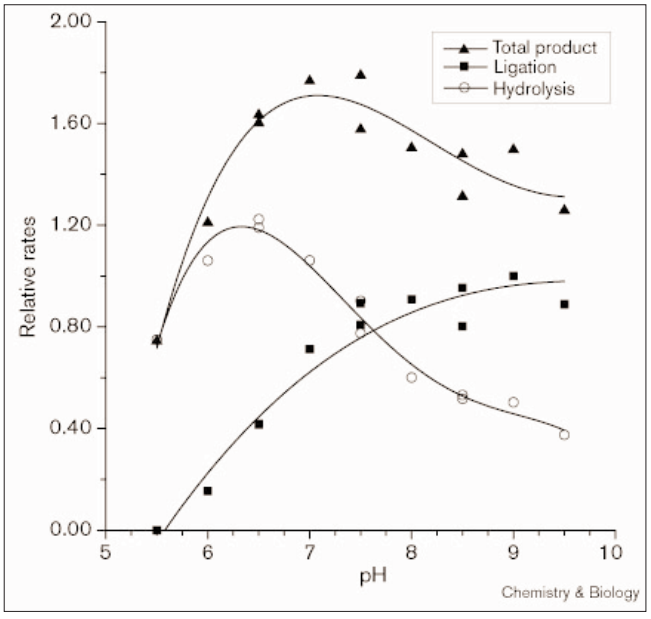
Mechanistic scheme for variant-catalyzed reactions. The leaving group is indicated by X and is one of -*p*NA, -Sbz, -SMe or -OMe.

Figure 6



Rate of S195A-catalyzed -*p*NA release as a function of nucleophilic acceptor peptide concentration (NH_2-FAASR-NH_2). Hydrolysis was performed with 2.0 mM *N*-succinyl-A-A-P-F-*p*NA, in 200 mM tricine (pH 8.0). Data were fit to a competitive inhibition model.

Figure 7



pH profile of the S195A-catalyzed ligation reaction. The standard ligation reaction was performed for 16 h, with varying buffer pH. Buffers used: MES (2-[N-morpholino]ethanesulfonic acid) pH 5.5–6.5; MOPS (3-[N-morpholino]propanesulfonic acid) pH 6.5–7.5; tricine pH 7.5–8.5; CHES (2-[N-cyclohexlamino]sulfonic acid) pH 8.0–10. The maximum ligation rate was set to 1. Trend lines are for the visual aid of the reader.

The pH profile of the S195A variant is consistent with the enzyme-bound intermediate being an acylated imidazole (Figure 7). At pH values below 6.5, the activity of the enzyme, as measured by total product formation, falls off dramatically. Histidine residues typically titrate in the vicinity of pH 6.5 and protonation of the histidine's imidazole ring eliminates the nucleophilicity of this group. Furthermore, there are well-established precedents for histidine's role as a nucleophile [44,45]. For example, Quinn *et al.* [44] proposed that *para*-nitrophenyl acetate acylates the active-site histidine of α-lytic protease (a homolog of SGPB) to generate an acyl-imidazole, and the intermediate subsequently undergoes an N→O acyl transfer to give the acylated serine.

Consistent with the formation of an acyl-imidazole intermediate, both the hydrolysis and aminolysis activities of the S195A variant were eliminated by the addition of the irreversible serine protease inhibitor phenylmethylsulfonyl fluoride (PMSF) [46]. The enzyme was not irreversibly inactivated by PMSF as is the case for wild-type SGPB, however. The S195A variant regained ~10% of its activity over the course of a 24-hour incubation. Similar results were obtained with the S195G variant. It is postulated, therefore, that PMSF forms a covalent adduct with the active-site histidine (His57), preventing reaction with the donor peptide. Subsequently, the adduct slowly hydrolyzes, regenerating an active ligase. Preliminary attempts were made to isolate and directly characterize the PMSF adduct using mass spectroscopy. Unfortunately, because the intermediate is unstable, the attempts proved unsuccessful.

Taken together, the results from the inhibitor, pH, product ratio and kinetic studies indicate that the ligation reaction proceeds through an acyl-enzyme intermediate resulting from the nucleophilic attack on the activated peptide by the imidazole of His57. The acyl-imidazole intermediate is subsequently deacylated by either water (hydrolysis) or the amine of the nucleophilic acceptor peptide (ligation).

Ligation efficiency

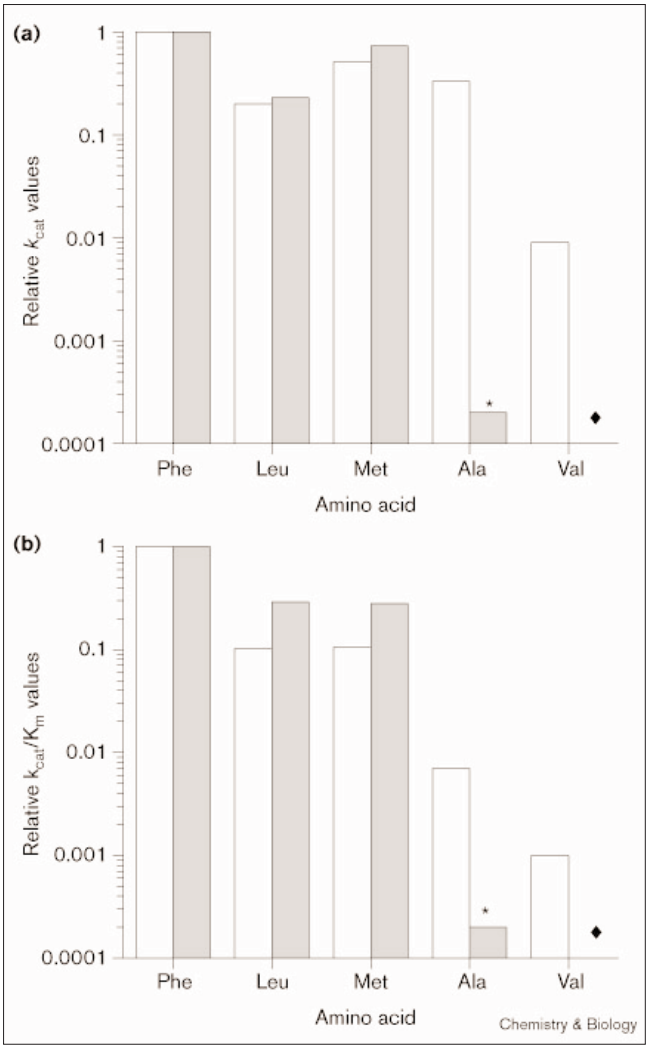
The efficiency of ligation to hydrolysis increases as the pH of the solution increases (Figure 7). In addition, more aminolysis is observed at higher acceptor concentrations (Figure 4). These observations are consistent with the aminolysis reaction requiring a neutral (deprotonated) amine nucleophile, and ligation efficiency may therefore be improved by working at higher pH and with greater concentrations of acceptor peptide.

Substrate specificity

To investigate the substrate specificity of the S195A variant, the enzymatic hydrolysis of several substrate peptides (*N*-succinyl-Ala–Ala–Pro–Xaa-*p*NA, where Xaa is the

P1 amino acid) was examined (Figure 8). The kinetic constants k_{cat} and k_{cat}/K_m are observed to follow the same trend for the S195A variant as for the wild-type protease, indicating the retention of primary specificity. Although the requirement is less stringent, SGPB also shows P1′ specificity toward large hydrophobic amino acid residues. This observation is reflected in the ligation activity of the S195A variant, which is capable of ligating acceptor peptides containing an amino-terminal phenylalanine, leucine or alanine residue (listed in order of preference, data not shown).

Figure 8



Comparison of wild-type and S195A variant substrate specificity. Shown are relative **(a)** k_{cat} and **(b)** k_{cat}/K_m values for the SGPB- and S195A-variant-catalyzed hydrolysis of *N*-succinyl-AAPX-*p*NA. For both wild-type (white bars) and variant (gray bars) enzymes, k_{cat} and k_{cat}/K_m for hydrolysis of *N*-succinyl-AAPF-*p*NA were set to 1. For the hydrolysis of the alanine (*) substrate by S195A, a rate was observed; however, it was not possible to determine accurate kinetic constants. The valine (♦) substrate was not hydrolyzed by the S195A variant.

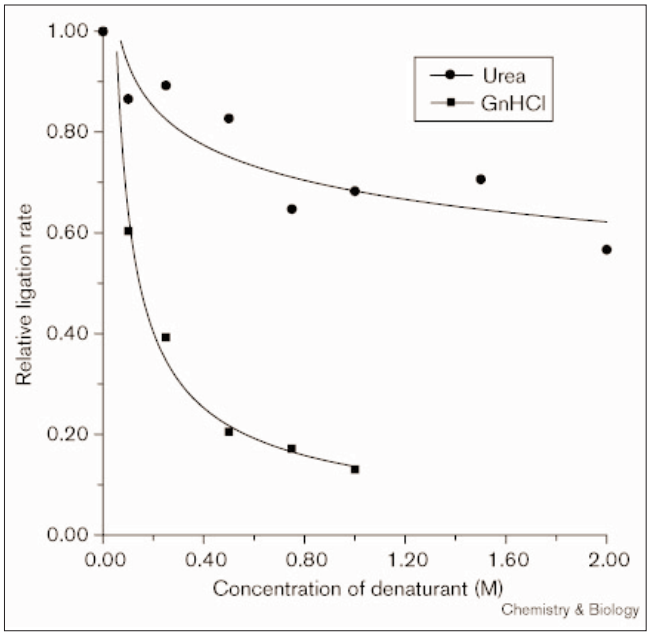
Stability of the S195A ligase variant

The ligation activity of the S195A variant was examined in various concentrations of urea and GnHCl (Figure 9). Ligation was observed in concentrations of urea as high as 2 M, but the enzyme was inactivated by relatively low concentrations of GnHCl.

The S195A-catalyzed ligation was also investigated over a range of temperatures (Figure 10). Ligation reactions proceeded at temperatures as high as 55°C, although a dramatic decrease in rate occurs above ~42°C, which probably indicates enzyme denaturation. The thermostability of the enzyme was evaluated in the absence of substrates. Purified protein was heated at various temperatures, and then tested for ligase activity. The enzyme was completely inactivated in less than 1 minute at 45°C, and less than 20 minutes at 37°C. Wild-type SGPB has a $t_{1/2}$ of 8.3 minutes at 55°C under similar experimental conditions.

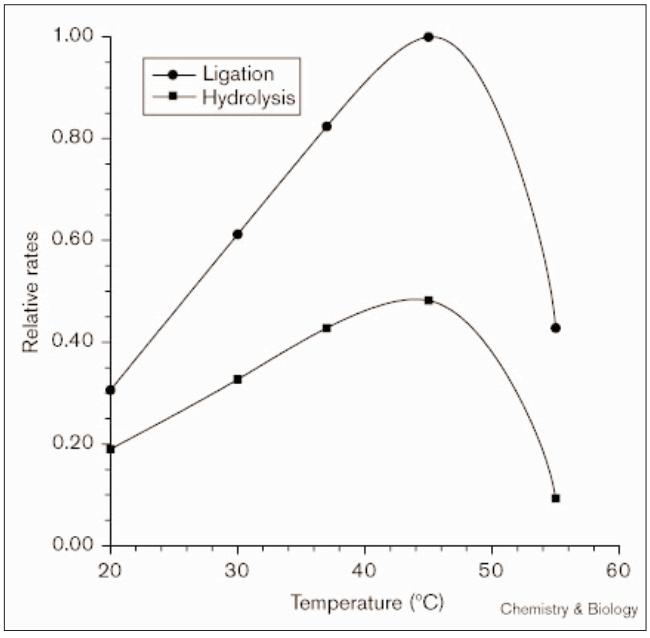
These results indicate that the S195A variant is less stable than the wild-type enzyme. Several variants of the wild-type enzyme with increased thermal stability have been generated ($t_{1/2}$ of 34–84 minutes at 55°C) [47], and we are currently in the process of making the serine to alanine substitutions in these, in an attempt to create a more stable ligase.

Figure 9



Relative ligation rates by S195A in urea and guanidine hydrochloride. Standard ligation reactions were performed for 16 h, with varying amounts of urea or guanidine hydrochloride. Ligation reactions were analyzed by HPLC, and the relative ligation rate calculated. The maximum ligation rate was set to 1. Trend lines are for the visual aid of the reader.

Figure 10



Temperature profile of the S195A-catalyzed ligation reaction. Standard ligation reactions were performed for 16 h at various reaction temperatures. The maximum ligation rate was set to 1. Trend lines are for the visual aid of the reader.

Significance

Recent developments in peptide ligation, including the enzymatic coupling of protein fragments, have allowed tremendous advances in the study of protein structure and function. It has become relatively practical to create proteins with non-natural amino acids incorporated site-specifically, which can serve as probes of catalytic activity, binding and structural alterations.

An enzyme of the chymotrypsin superfamily (*Streptomyces griseus* protease B, SGPB) was made into an effective coupling catalyst by single amino-acid substitutions at Ser195. Replacement of the catalytic serine residue with small hydrophobic amino acids resulted in a novel class of catalysts that couple peptides via an acyl-imidazole intermediate. The mutant enzymes were able to efficiently ligate peptides activated with thioester, ester and *para*-nitroanilide activating groups, and showed poor amidase activity.

The potential to apply this mutagenic strategy to the proteases contained within the chymotrypsin superfamily (including variants of SGPB itself) may enable a more general method for creating peptide ligases, covering a wide range of specificities and physical characteristics. An array of coupling enzymes would undoubtedly expand the application of semi-synthetic strategies in protein engineering. Furthermore, the ability of such mutants to use readily accessible and hydrolytically stable anilide and methyl esters as acyl donors permits the use of high pH (8–10), as required for deprotonation of amine nucleophiles. The ability to work at high pH provides a versatility not achieved by peptide-ligation strategies involving the use of less hydrolytically stable thioesters.

Materials and methods

Bacterial strains

DNA manipulation was performed as previously described [43] in *E. coli* strain DH5α. Protein expression was performed in the protease-deficient strain *B. subtilis* DB104.

Plasmids

The pNC-B plasmid is a pUC-19 derivative containing the full-length SGPB gene (*sprB*) and was used as a template for mutagenic PCR. The pEB-11 plasmid is an *E. coli/B. subtilis* shuttle vector, and contains components of pUC-18 for maintenance in *E. coli*, and components of pUB110 for maintenance in *B. subtilis* [31]. In addition, it contains the promoter and pre- regions of the subtilisin BPN gene, allowing expression and secretion of downstream gene products in *B. subtilis*. The pEB-B8 plasmid contains the pro-mature region of the *sprB* gene (wild-type SGPB) cloned in frame behind the promoter and pre- regions of subtilisin in pEB-11 [48].

Primers

Primers (Table 2) were synthesized on an Applied Biosystems 392 DNA/RNA synthesizer. Mismatch codons are shown in bold (shown with the mutant amino acid encoded above) and restriction sites are in italics (labeled with specific restriction site).

Construction of the expression vectors pEB-HDg/-HDa/-HDc

All three vectors were constructed in the same manner. pEB-HDx (x = a, g or c) contains the promoter region of subtilisin and a gene construct encoding the pre- region of subtilisin, followed by the propeptide and mature region of SGPB (with a S195X mutation).

BF1 and S195X (X = Ala, Gly or Cys) were used as primers in a PCR with pNC-B as template, and the 658 base pair amplified product isolated. The fragment was digested with *Kpn*I and ligated into pUC-19 (digested with *Kpn*I and *Hinc*II and treated with CIP (calf intestinal

Table 2

PCR primers.

BF1:	5′ *GGTACC*TGCGCAACCCCCCGGACG 3′ *Kpn*I
S195G:	5′ GAG*CGGACCGCC***GCC**GTCGCC 3′ *Rsr*II (Gly)
S195A:	5′ GAG*CGGACCGCC***GGC**GTCGCC 3′ *Rsr*II (Ala)
S195C:	5′ GAG*CGGACCGCC***GCA**GTCGCC 3′ *Rsr*II (Cys)
H57N:	5′ ACCGCCGGC**AAC**TGCACGGACGGC 3′ (Asn)
pEBloPvuII-:	5′ GGCCTCTTCGCTATTAC 3′

alkaline phosphatase). Vectors with inserts of the correct size were identified by restriction analysis and were partially sequenced to verify accurate amplification, mutagenesis and ligation.

Plasmids containing the mutagenized inserts were digested with *Dra*III and *Rsr*II, and the 310 base pair fragment gel-purified and ligated into pEB-B8 (digested with *Dra*III and *Rsr*II and treated with CIP). A vector containing the correct insert was isolated on the basis of restriction digest analysis and sequenced.

pEB-nDg

pEB-nDg contains the subtilisin promoter and contains the gene construct encoding the subtilisin pre- region, followed by the propeptide and mature region of SGPB (with H57N and S195G mutations).

H57N and pEBloPvuII- (primes downstream of gene construct) were used in a PCR with pEB-HDg as template. The ~600 base pair amplified product was gel-purified, digested with *Pst*I and ligated into pUC-19 (digested with *Pst*I and *Sma*I and treated with CIP). Vectors containing inserts of the correct size were identified by restriction analysis and labeled pJE-Bng.

pJE-Bng was digested with *Ngo*MI and *Pst*I, and the 465 base pair fragment was gel-purified, and ligated into pEB-HDg (digested with *Ngo*MI and *Pst*I and treated with CIP). A vector containing the correct insert was identified by restriction analysis, sequenced to ensure mutation, and designated pEB-nDg.

Production of the SGPB variants

B. subtilis DB104 cultures harboring the respective expression vectors were grown at 30°C in PEM (32 g tryptone, 2 g yeast extract, 5 g NaCl per l) + kanamycin (50 mg/l). For large-scale expression of the S195A and S195G variants, the cultures were scaled up to 15 l and grown at 30°C with 6 l/min airflow at 200 rpm in a Chemap Fermenter. SGPD was added to a final concentration of 0.35 U/l [30] when the A_{600} had reached 0.6, after which the culture was grown for an additional 48 h and harvested for purification.

Protein gels and western blots

SDS–PAGE and western blots were performed as previously described [40,49].

Purification of S195A and S195G variants

Bacteria were removed from the large-scale (15 l) fermenter culture by filtration with a Millipore Pellicon apparatus equipped with two 0.45 µm cutoff membrane cassettes. The filtrate (culture supernatant) was exchanged into 25 mM sodium acetate (pH 4.6) and concentrated to approximately 1 l with two 10,000 molecular weight cutoff membrane cassettes. The retentate (concentrated supernatant) was centrifuged at $10{,}000 \times g$ to remove any material. Purification of the SGPB variants from the retentates was performed by a combination of ammonium sulfate precipitation (S195G variant only) and ion-exchange chromatography on Source-15Q and Source-15S columns (Pharmacia). Purified proteins were analyzed by matrix assisted laser desorption/ionization–time-of-flight (MALDI–TOF) mass spectrometry (PerSeptive Biosystems Voyager-DE), and amino-terminal amino acid sequencing (Applied Biosystems 473 peptide sequencer). Protein concentration was determined by the method of Scopes [50].

Peptides

The acceptor pentapeptide, NH_2-F-A-A-S-R-NH_2 (using single-letter amino-acid code), was synthesized on an Applied Biosystems 432A peptide synthesizer or purchased from SynPep Corporation (Dublin, CA).

Synthesis of N*-succinyl-A-A-P-F-OMe*

N-succinyl-A-A-P-F-OH was purchased from SynPep. 50 mg of peptide was dissolved in anhydrous ether and allowed to react with an excess of diazomethane for 1 h [51]. The peptide methyl ester was lyophilized, resuspended in 10% methanol and purified by reverse phase (C-18) HPLC.

Kinetic studies

Amidase activity was determined using the following assay mixture: 100–2000 µM substrate (*N*-succinyl-Ala–Ala–Pro–Xaa-*p*-nitroanilide, where Xaa is phenylalanine, methionine, leucine, alanine or valine) in 200 mM Tricine buffer (pH 8.0), in a total volume of 1 ml, 25°C. The release of *p*-NA was detected spectrophotometrically at 412 nm.

Ligation reactions

Typically, ligation reactions were composed of 1.2 mM donor peptide (*N*-succinyl-A-A-P-F-SBz (Sigma), *N*-succinyl-A-A-P-F-*p*NA (BACHEM, Torrance, CA) or *N*-succinyl-A-A-P-F-SMe (Gryphon Sciences, San Francisco) or *N*-succinyl-A-A-P-F-OMe), 11 mM F-A-A-S-R-NH_2 (SynPep), 4% 1,4-butanediol (for methyl ester and thioesters (Sigma)), 50 mM Tricine (*N*-tris[hydroxymethyl]methylglycine) (pH 8.0) and ~1 µM enzyme, and were performed at 30°C. Deviations from these conditions are noted in the figure legends. Ligation reactions were analyzed by reverse-phase HPLC (RP-HPLC) on a Waters LC Module 1 equipped with an analytical Nova-pak C-18 column (8×10). Solvent A was water containing 0.05% TFA; solvent B was acetonitrile containing 0.045% TFA. Gradients were 5% to 65% B over 30 min. Peaks from HPLC analysis were identified by MALDI–TOF mass spectrometry.

Acknowledgements

We thank K. Joe and A. Schultz for help in protein purification. R.J.E. is the recipient of a Medical Research Council of Canada Studentship. This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

References

1. Smith, M. (1985). *In vitro* mutagenesis. *Annu. Rev. Genet.* **19**, 423-462.
2. Hutchison, C.A.D., Phillips, S., Edgell, M.H., Gillam, S., Jahnke, P. & Smith, M. (1978). Mutagenesis at a specific position in a DNA sequence. *J. Biol. Chem.* **253**, 6551-6560.
3. Fersht, A. & Winter, G. (1992). Protein engineering. *Trends Biochem. Sci.* **17**, 292-295.
4. Knowles, J.R. (1987). Tinkering with enzymes: what are we learning? *Science* **236**, 1252-1258.
5. Offord, R. (1991). Going beyond the code. *Protein Eng.* **4**, 709-710.
6. Noren, C.J., Anthony-Cahill, S.J., Griffith, M.C. & Schultz, P.G. (1989). A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* **244**, 182-188.
7. Doyle, D.F. & Corey, D.R. (1998). Engineering proteins: keeping up with biology. *Chem. Biol.* **5**, R157-R160.
8. Wilken, J. & Kent, S.B. (1998). Chemical protein synthesis. *Curr. Opin. Biotechnol.* **9**, 412-426.
9. Marglin, A. & Merrifield, R.B. (1970). Chemical synthesis of peptides and proteins. *Annu. Rev. Biochem.* **39**, 841-866.
10. Offord, R. (1980). *Semisynthetic Proteins*. J. Wiley, New York.
11. Cotton, G.J. & Muir, T.W. (1999). Peptide ligation and its application to protein engineering. *Chem. Biol.* **6**, R247-R256.
12. Lu, W., Qasim, M.A., Laskowski, M., Jr. & Kent, S.B. (1997). Probing intermolecular main chain hydrogen bonding in serine proteinase-protein inhibitor complexes: chemical synthesis of backbone-engineered turkey ovomucoid third domain. *Biochemistry* **36**, 673-679.
13. Lu, W., Randal, M., Kossiakoff, A. & Kent, S.B. (1999). Probing intermolecular backbone H-bonding in serine proteinase–protein inhibitor complexes. *Chem. Biol.* **6**, 419-427.
14. Jackson, D.Y., Burnier, J., Quan, C., Stanley, M., Tom, J. & Wells, J.A. (1994). A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues. *Science* **266**, 243-247.
15. Muir, T.W., Sondhi, D. & Cole, P.A. (1998). Expressed protein ligation: a general method for protein engineering. *Proc. Natl Acad. Sci. USA* **95**, 6705-6710.
16. Kullman, W. (1987). *Enzymatic Peptide Synthesis*. CRC Press, Boca Raton, Florida, USA.
17. Dawson, P.E., Muir, T.W., Clark-Lewis, I. & Kent, S.B. (1994). Synthesis of proteins by native chemical ligation. *Science* **266**, 776-779.
18. Mathys, S., *et al.,* & Xu, M.Q. (1999). Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. *Gene* **231**, 1-13.

19. Chong, S., *et al.,* & Xu, M.Q. (1997). Single-column purification of free recombinant proteins using a self- cleavable affinity tag derived from a protein splicing element. *Gene* **192**, 271-281.
20. Chong, S., Shao, Y., Paulus, H., Benner, J., Perler, F.B. & Xu, M.Q. (1996). Protein splicing involving the *Saccharomyces cerevisiae* VMA intein. The steps in the splicing pathway, side reactions leading to protein cleavage, and establishment of an *in vitro* splicing system. *J. Biol. Chem.* **271**, 22159-22168.
21. Xu, M.Q. & Perler, F.B. (1996). The mechanism of protein splicing and its modulation by mutation. *EMBO J.* **15**, 5146-5153.
22. Xu, R., Ayers, B., Cowburn, D. & Muir, T.W. (1999). Chemical ligation of folded recombinant proteins: segmental isotopic labeling of domains for NMR studies. *Proc. Natl Acad. Sci. USA* **96**, 388-393.
23. Welker, E. & Scheraga, H.A. (1999). Use of benzyl mercaptan for direct preparation of long polypeptide benzylthio esters as substrates of subtiligase. *Biochem. Biophys. Res. Commun.* **254**, 147-151.
24. Muir, T.W., Dawson, P.E. & Kent, S.B. (1997). Protein synthesis by chemical ligation of unprotected peptides in aqueous solution. *Methods Enzymol.* **289**, 266-298.
25. Canne, L.E., Bark, S.J. & Kent, S.B.H. (1996). Extending the applicability of native chemical ligation. *J. Am. Chem. Soc.* **118**, 5891-5896.
26. Braisted, A.C., Judice, J.K. & Wells, J.A. (1997). Synthesis of proteins by subtiligase. *Methods Enzymol.* **289**, 298-313.
27. Abrahmsen, L., Tom, J., Burnier, J., Butcher, K.A., Kossiakoff, A. & Wells, J.A. (1991). Engineering subtilisin and its substrates for efficient ligation of peptide bonds in aqueous solution. *Biochemistry* **30**, 4151-4159.
28. Atwell, S. & Wells, J.A. (1999). Selection for improved subtiligases by phage display. *Proc. Natl Acad Sci. USA* **96**, 9497-9502.
29. Rawlings, N.D. & Barrett, A.J. (1994). Families of serine peptidases. *Methods Enzymol.* **244**, 19-61.
30. Sidhu, S.S., Kalmar, G.B., Willis, L.G. & Borgford, T.J. (1995). Protease evolution in *Streptomyces griseus*. Discovery of a novel dimeric enzymes. *J. Biol. Chem.* **270**, 7594-7600.
31. Sidhu, S.S., Kalmar, G.B. & Borgford, T.J. (1993). Characterization of the gene encoding the glutamic-acid-specific protease of *Streptomyces griseus*. *Biochem. Cell Biol.* **71**, 454-461.
32. Jurasek, J., Johnson, P., Olafson, R.W. & Smillie, L.B. (1971). An improved fractionation system for pronase on CM-sephadex. *Can. J. Biochem.* **49**, 1195-1201.
33. Jurasek, L., Carpenter, M.R., Smillie, L.B., Gertler, A., Levy, S. & Ericsson, L.H. (1974). Amino acid sequence of *Streptomyces griseus* protease B, a major component of pronase. *Biochem. Biophys. Res. Commun.* **61**, 1095-1100.
34. Bauer, C.A. (1978). Active centers of *Streptomyces griseus* protease 1, *Streptomyces griseus* protease 3, and alpha-chymotrypsin: enzyme-substrate interactions. *Biochemistry* **17**, 375-380.
35. Bauer, C.A. (1980). Active centers of alpha-chymotrypsin and of *Streptomyces griseus* proteases 1 and 3. S2–P2 enzyme–substrate interactions. *Eur. J. Biochem.* **105**, 565-570.
36. Johnson, E.R., Lamkin, R.M., Jordan, K.J., Gordon, J.A. & Lindley, B.R. (1987). Determination of the cleavage specificity of *Streptomyces griseus* protease B in the presence of guanidinium chloride. *Int. J. Pept. Protein Res.* **30**, 170-176.
37. Siegel, S., Brady, A.H. & Awad, W.M. (1972). The proteolytic enzymes of the K-1 strain of *Streptomyces griseus* obtained from a commercial preparation (Pronase). II. The activity of a serine enzyme in 6 M guanidinium chloride. *J. Biol. Chem.* **247**, 4155-4159.
38. Polgár, L. (1989). *Mechanisms of Protease Action*. CRC Press, Boca Raton, Florida.
39. Henderson, G., Krygsman, P., Liu, C.J., Davey, C.C. & Malek, L.T. (1987). Characterization and structure of genes for proteases A and B from *Streptomyces griseus*. *J. Bacteriol.* **169**, 3778-3784.
40. Baardsnes, J., *et al.,* & Borgford, T.J. (1998). *Streptomyces griseus* protease B: secretion correlates with the length of the propeptide. *J. Bacteriol.* **180**, 3241-3244.
41. Carter, P. & Wells, J.A. (1988). Dissecting the catalytic triad of a serine protease. *Nature* **332**, 564-568.
42. Fersht, A. (1985). *Enzyme Structure and Mechanism*. (2nd edn), W.H. Freeman, New York.
43. Jencks, W.P. (1969). *Catalysis in Chemistry and Enzymology*. McGraw-Hill, New York.
44. Quinn, D.M., Elrod, J.P., Ardis, R., Friesen, P. & Schowen, R.L. (1980). Protonic reorganization in catalysis by serine protease: acylation by small substrates. *J. Am. Chem. Soc.* **102**, 5358-5365.
45. Ryan, D.S. & Feeney, R.E. (1975). The interaction of inhibitors of proteolytic enzymes with 3-methylhistidine-57-chymotrypsin. *J. Biol. Chem.* **250**, 843-847.
46. Gold, A.M. (1967). Sulfonylation with sulfonyl halides. *Methods Enzymol.* **11**, 706-711.
47. Sidhu, S.S. & Borgford, T.J. (1996). Selection of *Streptomyces griseus* protease B mutants with desired alterations in primary specificity using a library screening strategy. *J. Mol. Biol.* **257**, 233-245.
48. Sidhu, S.S., Kalmar, G.B., Willis, L.G. & Borgford, T.J. (1994). *Streptomyces griseus* protease C. A novel enzyme of the chymotrypsin superfamily. *J. Biol. Chem.* **269**, 20167-20171.
49. Ausubel, F.M. (1987). *Current Protocols in Molecular Biology*. J. Wiley, Brooklyn, New York.
50. Scopes, R.K. (1994). *Protein Purification: Principles and Practice*. (3rd edn), Springer-Verlag, New York.
51. Furniss, B.S., Hannaford, A.J., Smith, P.W.G. & Tatchell, A.R. (1989). *Vogel's Textbook of Practical Organic Chemistry*. (5th edn), Longman Scientific and Technical, London.