Pergamon

Tetrahedron 56 (2000) 9477–9485

TETRAHEDRON

Generation of Active Trypsin by Chemical Cleavage

Teaster Baird Jr.,[a,†] Bixun Wang,[b,†] Michiel Lodder,[b] Sidney M. Hecht[b] and Charles S. Craik[a,*]

[a]*Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, CA 94143, USA*
[b]*Departments of Chemistry and Biology, University of Virginia, Charlottesville, VA 22901, USA*

Received 15 May 2000; accepted 6 July 2000

Abstract—A novel strategy has been developed to generate trypsin from trypsinogen without proteolytic processing. The unnatural amino acid allylglycine was incorporated into a trypsinogen variant to permit cleavage specifically at Ile-16, the prototypical activation site, and generate the active protease, trypsin. A suppressor tRNA, activated with allylglycine, was used to synthesize the allylglycine trypsinogen variant by in vitro translation of an mRNA transcript that placed the UAG stop codon immediately prior to the authentic trypsin coding region. Iodine treatment of the variant trypsinogen resulted in trypsin that was fully active and kinetically indistinguishable from wild type recombinant trypsin. Using single substrates, the iodine treated variant exhibited a catalytic profile essentially identical to that of wild type recombinant rat trypsin (k_{cat}/K_M=3.7±0.5 vs 3.7±0.8 $\mu M^{-1}s^{-1}$, respectively). Furthermore, the specificity profiles generated from substrate combinatorial libraries are indistinguishable. These results illustrate that (1) iodine activated allylglycine trypsinogen is virtually identical to enteropeptidase activated recombinant trypsinogen and (2) this novel technique is a feasible alternative to achieve site specific cleavage of serine protease zymogens.

Introduction

Trypsin (Tn), the canonical serine protease, is an endopeptidase that catalyzes the hydrolysis of peptide bonds on the carboxyl side of arginine and lysine residues. It belongs to the pancreatic serine protease family whose members are related by similar structure and function.[22] Trypsin is expressed within the acinar cells of the pancreas as trypsinogen (Tg), the inactive precursor or zymogen form of the enzyme. The NH_2-terminal region of the zymogen consists of two distinct amino acid sequences that act as signal[5] and activation peptides. The most N-terminal sequence specifies vectoral transport of trypsinogen into the cisternae of the rough endoplasmic reticulum, the first step in sequestering the protein for secretion.[27] After the signal peptide is removed, the activation peptide sequence (also referred to as the propeptide) keeps the protein in an inactive conformation and prevents uncontrolled peptide digestion by the enzyme.

Trypsinogen is secreted into the small intestine where it is activated by enteropeptidase,[18] a highly selective serine protease that releases the activation peptide from the N-terminus of trypsinogen via proteolytic cleavage. The activation peptide of trypsinogen is typically an octa- or hexapeptide sequence that contains a cluster of anionic amino acids, e.g. Val-Asp_4-Lys.[17] Enteropeptidase recognizes this sequence and cleaves the peptide bond carboxyl to the lysine residue. The conversion of trypsinogen to trypsin is the first step of the digestive cascade–activation sequence and consequently relies on the high specificity of enteropeptidase. Once activated, trypsin activates other pancreatic zymogens, including trypsinogen.

A free amino terminus at position 16 (chymotrypsinogen numbering is used throughout) is absolutely required to obtain a fully active enzyme. Chemical modification experiments in which the N-terminus of trypsin was acylated illustrated that such treatment completely abolished any

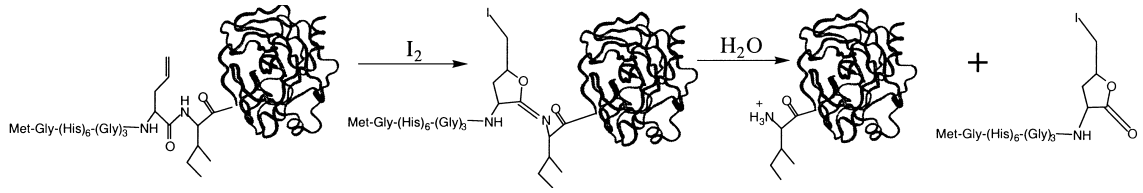

Figure 1. Scheme illustrating how treatment of an allylglycine-containing trypsinogen with iodine can lead to site-specific cleavage of the protein backbone and conversion to catalytically competent trypsin.

Keywords: protease; hydrolysis; unnatural amino acids; zymogen activation.
* Corresponding authors. Tel.: +1-415-476-8146; fax: +1-415-502-8298; e-mail: craik@cgl.ucsf.edu
† Authors contributed equally to this work.

0040–4020/00/$ - see front matter
PII: S0040-4020(00)00832-2

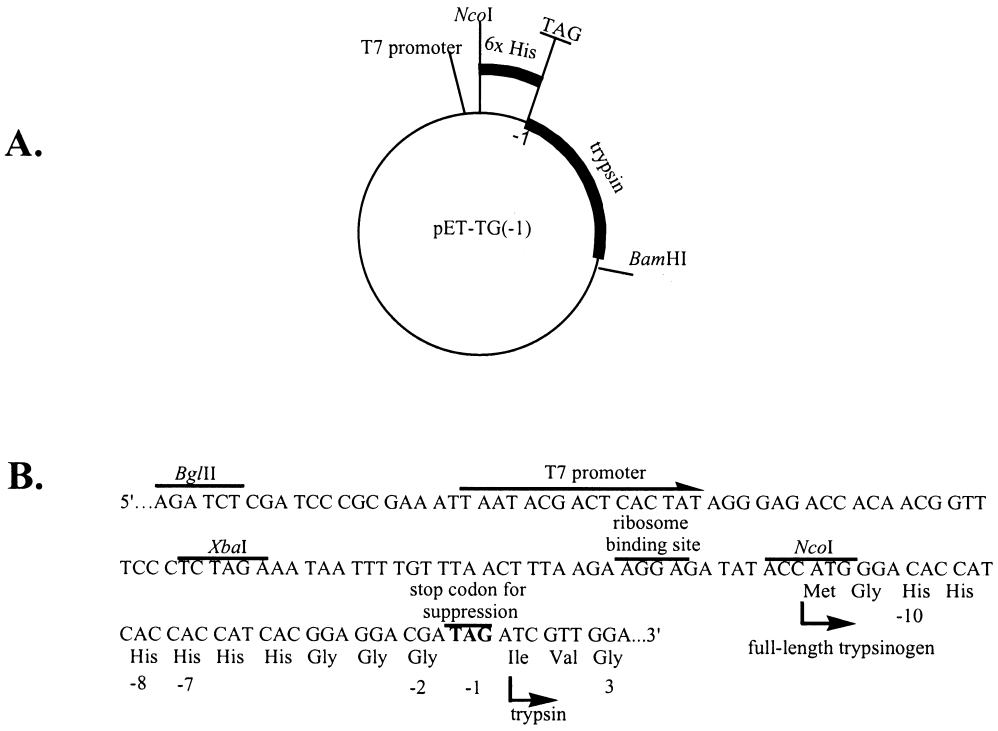

Figure 2. (A) Plasmid construct used for the in vitro elaboration of allylglycine-containing trypsinogen (^{A}Gly-Tg). (B) Nucleotide sequence and deduced amino acid sequence at the 5′-end of the trypsinogen gene variant. The activation peptide of wild-type trypsinogen was replaced by the peptide sequence Met-Gly-$(His)_6$-Gly-Gly-Gly-^{A}Gly. The inclusion of hexahistidine moiety was intended to facilitate purification of the derived ^{A}Gly-Tg by affinity chromatography on Ni–NTA agarose. Amino acids are numbered starting from the first Ile of wild-type trypsin. The nonsense codon TAG was introduced at position −1, immediately prior to the authentic trypsin coding region. The authentic stop codon for the gene was changed from TAG to TGA.

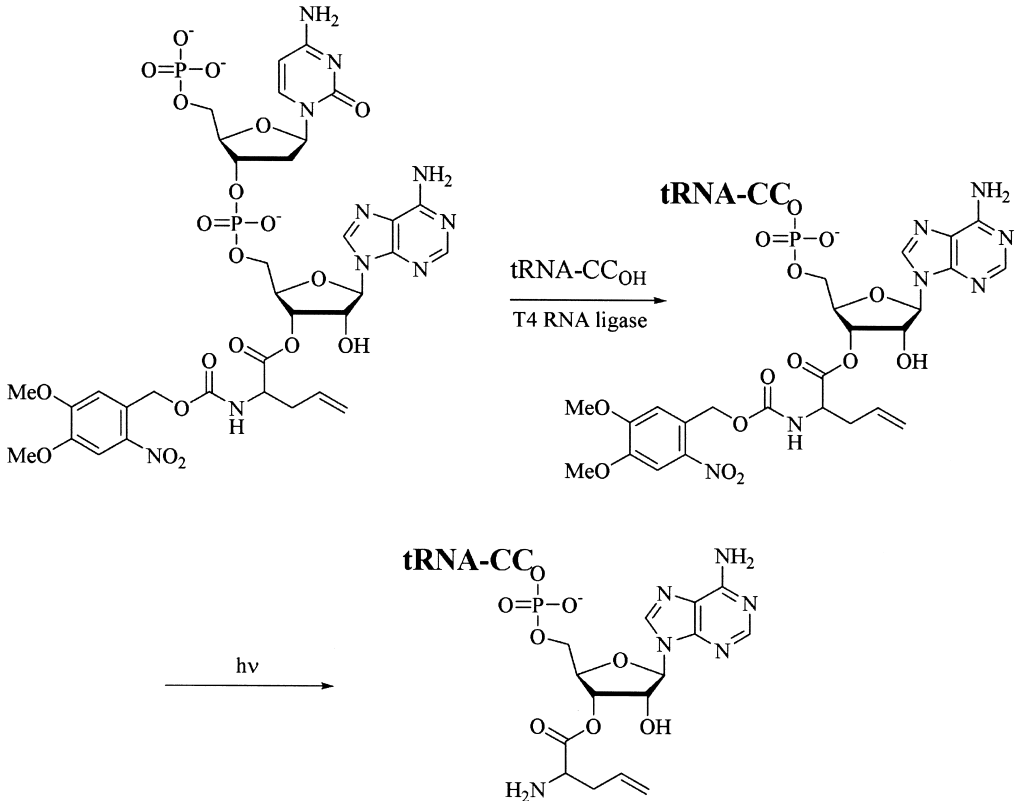

Figure 3. 'Chemical aminoacylation' strategy employed for the preparation of the mis-acylated suppressor $tRNA_{CUA}$ used for the elaboration of allylglycine-containing trypsinogen.

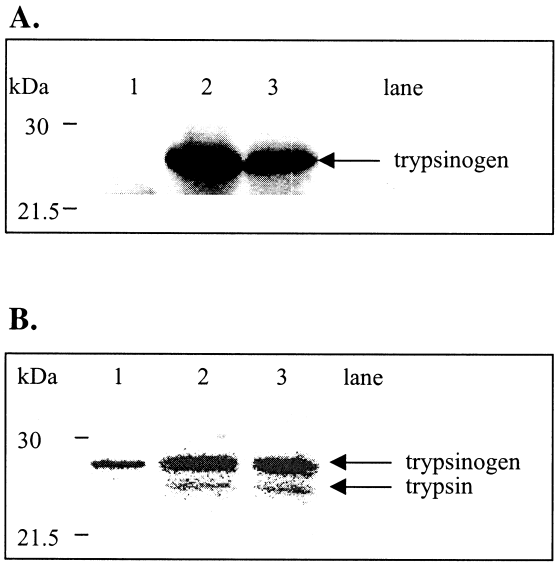

Figure 4. Phosphorimager analysis of a 20% SDS–polyacrylamide gel illustrating the in vitro synthesis of allylglycine-containing trypsinogen and iodine induced cleavage. (A) Protein synthesis was carried out in the presence of ^{35}S-methionine using a rabbit reticulocyte lysate, mRNA containing UAG at codon position position −1 and a deacylated suppressor $tRNA_{CUA}$ (lane 1) or a suppressor tRNA activated with either valine (lane 2) or allylglycine (lane 3). The reaction mixtures were analyzed following incubation at 30°C for 1 h. (B) The purified protein was treated with iodine in 100:15 water–tetrahydrofuran at 25°C for 30 min, then analyzed on a 20% SDS–polyacrylamide gel: lane 1, ^{A}Gly-Tg without I_2 treatment; lane 2, ^{A}Gly-Tg+1 mM I_2; lane 3, ^{A}Gly-Tg+1.5 mM I_2.

observable activity.[24] The isoleucine side chain is also important for activity of the mature trypsin since removing residue 16 greatly decreases the activity of trypsin.[24] Trypsinogen contains the entire amino acid sequence that is ultimately present in the mature enzyme, but it exhibits little trypsin activity. This inactivity is largely due to the distorted conformations of the S1 binding pocket and oxyanion hole in Tg relative to active Tn. After the activation peptide is lost, Asp-194 rotates about the backbone and forms a salt bridge with the new N-terminal amine of Ile-16.[6,15] This subtle structural rearrangement concurrently stabilizes the oxyanion hole and the S1 binding pocket to yield the active enzyme.[12]

Proteolytic processing by a suitable protease is currently the only method for activating a zymogen either in vivo or in vitro. However, we have described recently a novel strategy that may be used in vitro to specifically cleave engineered versions of dihydrofolate reductase and trypsinogen with a simple chemical reagent to produce truncated versions of these proteins.[28] We now describe in more detail this scheme for activating rat anionic trypsinogen without proteolytic processing. This technique allows site-specific cleavage of the protein backbone at a single, predetermined site and involves the elaboration of protein analogues that contain the non-native amino acid allylglycine (^{A}Gly). We used nonsense suppression[10] to incorporate allylglycine into rat trypsinogen at the prototypical site of proteolytic cleavage. Readthrough of a UAG codon[21,2] in the mRNAs of interest afforded full-length proteins that contained allylglycine at the desired position. Subsequent iodine treatment of the allylglycine-containing proteins resulted in cleavage at the engineered site through a presumed iodolactone intermediate and resulted in the activation of trypsin by chemical removal of the activation peptide (Fig. 1). This strategy thus provides an alternative and simple means to activate a zymogen by mild chemical treatment.

Results and Discussion

Template and transcript preparation

Plasmid pET-Tg (−1) was created for the in vitro synthesis of allylglycine-containing trypsinogen (^{A}Gly-Tg) (Fig. 2). The construct contained a modified rat anionic trypsinogen gene II under the control of a T7 promoter and efficient translation initiation signals. The activation peptide sequence was changed from Phe-Pro-Val-$(Asp)_4$-Lys to

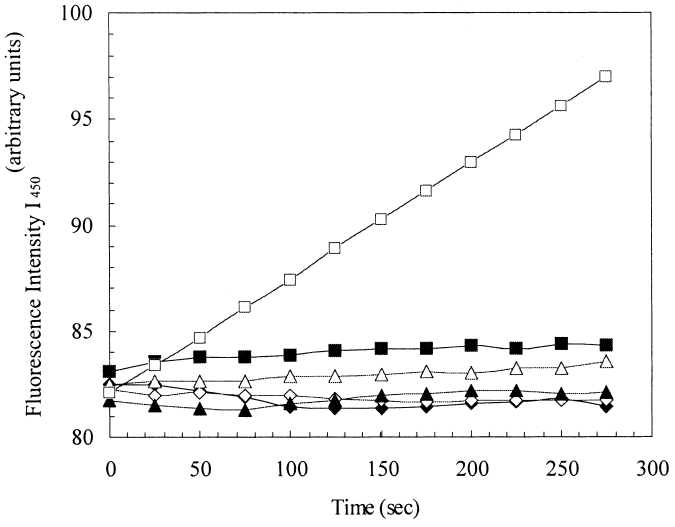

Figure 5. *Activity profile of trypsinogen containing allylglycine- (−1) (^{A}Gly-Tg) or containing valine- (−1) (Val-Tg) before and after iodine treatment.* ^{A}Gly-Tg and Val-Tg were synthesized in vitro, purified, dialyzed against 1 mM HCl, and subjected to 1 mM I_2 treatment in 10:1 THF–H_2O. Trypsin hydrolysis activity of 0.2 mM N_α-benzoyl-(L)-arginine 7-amido-4 methylcoumarin was compared among ^{A}Gly-Tg before (■) and after (□) iodine treatment; Val-Tg before (▲) and after (△) iodine treatment; and wild type bovine trypsinogen before (◆) and after (◇) iodine treatment.

Table 1. Comparison of kinetic parameters obtained from trypsin generated by iodine treatment of trypsinogen containing allylglycine (^{A}Gly-Tg) and wild-type trypsin (Tn[a]); wild type rate trypsin was added to rabbit reticulocyte lysate and subjected to the purification and iodine activation schemes used to activate ^{A}Gly-Tg

Sample	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($\mu M^{-1}\ s^{-1}$)
Tn[a]	44.1±1.6	12±2	3.7±0.8
Tn+L[b]	42.3±1.3	10.5±1.4	4.0±0.4
Tn+L+THF	42.5±2.4	14.1±2.8	3.0±0.5
Tn+L+THF+I_2	41.2±2.7	11.6±2.9	3.5±0.7
^{A}Gly-Tg+THF+I_2	42.1±1.7	11.4±1.6	3.7±0.5

[a] Synthesized in vivo as rat trypsinogen and subsequently activated with enterokinase (1:20, w/w). Active Tn purified by sequential hydrophobic and affinity column chromatography methods.
[b] Wild-type Tn was combined with rabbit reticulocyte lysate (L) and subjected to the same treatment used to prepare ^{A}Gly-Tg.

Met-Gly-$(His)_6$-Gly-Gly-Gly-^{A}Gly by substituting the sequence TTT CCC GTC GAT GAT GAT GAC AAG in the trypsinogen gene with ATG GGA CAC CAT CAC CAC CAT CAC GGA GGA GGA TAG. The hexahistidine sequence was included to facilitate purification of the in vitro synthesized ^{A}Gly-Tg by affinity chromatography on Ni–NTA agarose.[13] The three glycine residues following the hexahistidine sequence assisted in making the histidines accessible to the affinity resin. The nonsense codon TAG was introduced at position −1, immediately before the authentic trypsin coding region. Additionally, the authentic stop codon for the gene was changed from TAG to TGA to preclude readthrough by allylglycyl–$tRNA_{CUA}$. The final DNA construct, pET-Tg (−1), was obtained by subcloning the modified rat trypsinogen gene from the intermediate vector pGem7-Tg (−1) into the expression vector pET-14b. The mRNA was produced by run-off in vitro transcription of *Bam*HI-linearized pET-Tg (−1) plasmid DNA. The synthesized mRNAs were analyzed by polyacrylamide gel electrophoresis and shown to have the expected lengths of approximately 700 nucleotides (data not shown).

tRNA synthesis

The suppressor tRNA used in this study was a modified yeast $tRNA^{Phe}$ transcript[23,25] that contains a CUA anticodon. A truncated form of the tRNA lacking the 3′-terminal pCpA was prepared by in vitro T7 transcription of linearized pYRNA8 plasmid DNA[14] bearing a gene for the yeast

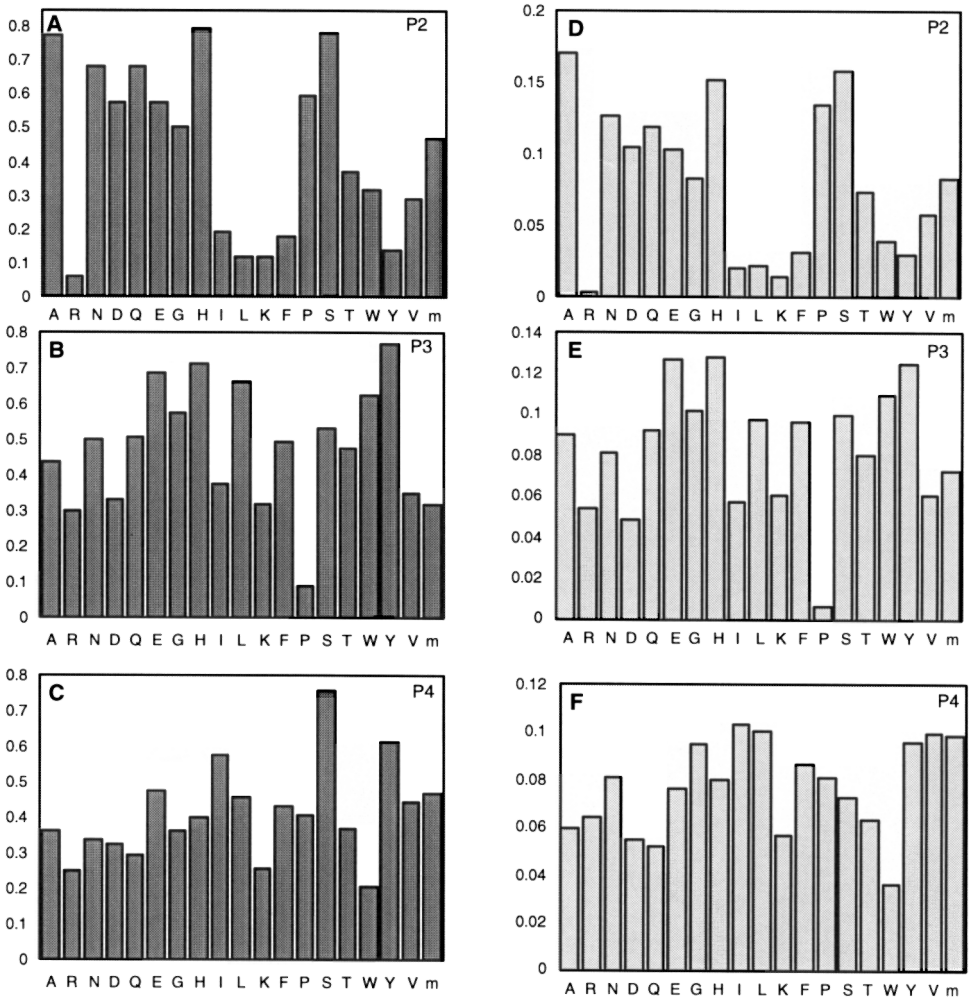

Figure 6. *Substrate specificity profiles of WT Tn and iodine activated ^{A}Gly-Tg.* The activity of wild type Tn or ^{A}Gly-Tn was assayed against a semi-randomized tetrapeptide substrate library to provide a *qualitative* assessment of the specificity profiles. P2, P3, and P4 indicated the substrate positions in which the amino acid indicated on the *X*-axis was held constant (m=norleucine). The units for the *Y*-axis are relative fluorescence units. Panels A, B, and C are for wild type Tn. Panels D, E, and F are for iodine activated ^{A}Gly-Tg.

$tRNA^{Phe}_{CUA}$. The mis-acylated $tRNA_{CUA}$ used for readthrough of the TAG codon was prepared by T4 RNA ligase-mediated coupling of *N*-(6-nitroveratryloxycarbonyl)-(*R*,*S*)-allylglycyl pCpA with the truncated tRNA transcript (Fig. 3). Valyl-$tRNA_{CUA}$ was prepared similarly. The mis-acylated suppressor $tRNA_{CUA}$ synthesized in this fashion included both *R*- and *S*-isomers of allylglycine. However, this is of little consequence since ribosome-mediated protein biosynthesis results in essentially exclusive incorporation of the *S*-amino acid into proteins.[10]

Synthesis and purification of ^{A}Gly-Tg

The rat trypsinogen analogues were produced using rat trypsinogen mRNA that was produced from runoff transcription of the linearized pET-Tg (−1) vector in an in vitro protein biosynthesizing system (rabbit reticulocyte lysate). The synthesis of full-length trypsinogen was dependent on the presence of an aminoacylated $tRNA_{CUA}$ (Fig. 4A). In the presence of unacylated suppressor tRNA, no ^{35}S-labeled protein was synthesized (Fig. 4A, lane 1). In the presence of either valyl (lane 2) or allylglycyl (lane 3) aminoacylated $tRNA_{CUA}$s, full-length proteins having the expected molecular weight (25 kDa) were produced.

The in vitro synthesized trypsinogen analogues were purified using anion exchange and affinity chromatography methods. The crude translation mixtures were first loaded onto DEAE Sepharose CL-6B to remove hemoglobin and cationic proteins that are present in the rabbit reticulocyte lysate. The trypsinogen analogues bound to the DEAE resin at pH 7.0 as expected from their anionic nature. After elution, the fractions that contained Tg (see Experimental section) were loaded onto a Ni–NTA agarose column[13] to retain specifically those proteins that have a conformationally accessible hexahistidine moiety.

The in vitro synthesized trypsinogen analogues exhibited minimal trypsin activity, comparable to wild-type recombinant Tg. Relative to wild-type trypsin, the ^{A}Gly-Tg was 10^6–10^8 fold less active. This result was consistent with the expectation that the in vitro synthesized ^{A}Gly-Tg would be an inactive enzyme since the engineered protein would be in a zymogen conformation.

To demonstrate that ^{A}Gly-incorporation produced a stable zymogen and precluded autoactivation of the protein, the trypsinogen analogues were assayed for their abilities to bind to an LTI (lime bean trypsin inhibitor)-agarose resin. Upon binding LTI, trypsinogen adopts a conformation resembling that of trypsin.[3,12] Furthermore, human trypsinogen will bind LTI–agarose resin only at low ionic strength, while trypsin is quantitatively bound even in the presence of 0.5 M NaCl.[4] Therefore, one would conclude that if the variant trypsinogens bound LTI agarose at low ionic strength, but eluted at high ionic strength, the polypeptide adopted the characteristic trypsinogen conformation. To test this characteristic, the affinities of ^{A}Gly-Tg and Val-Tg for LTI–agarose resin were tested quantitatively at low and high ionic strength. Using the [^{35}S]methionine in the Tg analogues as a quantifiable marker, we observed that at low ionic strength about 80% of the radioactivity was retained on the resin while at high ionic strength (0.5 M NaCl), 95% of the radiolabel passed through the column.

Treatment of the trypsinogen analogue ^{A}Gly-Tg with I_2 afforded a cleavage product having the expected size while the same treatment of Val-Tg gave no cleavage product (data not shown) verifying that the allylglycine at position 15 is required to produce a mature enzyme with iodine treatment. Phosphorimager analysis demonstrated that 9% cleavage was achieved under these conditions. The wild-type trypsinogen and trypsinogen analogues containing either allylglycine or valine at the normal cleavage site were subjected to a fluorescence-based trypsin activity assay, prior to and after iodine treatment. As shown in Fig. 5, only ^{A}Gly-Tg could be activated by iodine treatment.

Characterization of I_2 treated ^{A}Gly-Tg

Effect of iodine on trypsin activity. Since molecular iodine is a powerful oxidant, it could have detrimental effects on environment-sensitive macromolecules such as proteins. In fact, if wild-type recombinant trypsin is incubated with 1 mM iodine under the conditions used in the activation of ^{A}Gly-Tg, all observable activity is lost, presumably through destruction of the protein by iodine (data not shown). However, in the presence of a macromolecular protecting agent such as BSA (5 mg/ml), the effect of iodine is neutralized and the enzyme retains full activity (data not shown). Since initial qualitative experiments illustrated that iodine treated ^{A}Gly-Tg produced active enzyme, some protecting agent must have been present. It seemed plausible that the material remaining from the in vitro translation after the final purification step may act as the protecting agent, minimizing any adverse effects the iodine may have had on the in vitro translated zymogen. To test this hypothesis, wild type rat Tn was added to the rabbit reticulocyte lysate mixture used in the in vitro translation reaction and subjected to the purification and activation protocols used for ^{A}Gly-Tg. The Tn concentration was determined by active site titration before and after iodine treatment. The activity was measured after iodine treatment and compared to wild-type trypsin that was not treated with iodine. As Table 1 illustrates, the components of the activation solution (tetrahydrofuran and iodine) had no observable effect on the activity of trypsin in the presence of the in vitro translation components. These results suggest that the iodine treatment has inconsequential effects on the ^{A}Gly-Tg other than to initiate cleavage and removal of the propeptide in activation.

After removal of the variant propeptide by iodine activation, the resultant polypeptide should be identical to wild type mature rat trypsin. Consequently, one would expect the mature trypsin produced from iodine treatment of ^{A}Gly-Tg to be indistinguishable from wild type mature rat trypsin in all respects. To test this hypothesis, we measured the specific activity of iodine treated ^{A}Gly-Tg, determined its substrate specificity profile and compared the results with those of mature wild type rat trypsin.

Quantification of ^{A}Gly-Tn. The yield of active enzyme from iodine treatment of ^{A}Gly-Tg is too low to quantify

accurately with active site titration. Therefore, to determine the concentration of active enzyme produced by iodine treatment of ^{A}Gly-Tg, the specific activity of the active-site titrated wild type recombinant rat trypsinogen was measured using the fluorogenic substrate Z-GPR-AMC (*N*-α-benzyloxycarbonyl-L-glycylprolylarginine 7-amido-4-methylcoumarin)[11] and compared to the activity of iodine treated ^{A}Gly-Tg. Once determined, this concentration was used to measure the initial rates of hydrolysis against varying concentrations of Z-GPR-AMC. One possible caveat of this quantification method is the inherent assumption that each molecule of enzyme is either completely active or completely inactive. However, since iodine treatment of the wild type rat Tn in rabbit reticulocyte lysate did not affect the specific activity of the enzyme significantly (see Table 1), the assumption that Tn liberated from ^{A}Gly-Tg is ~100% active is likely to be valid. Additionally, no observable activity in the sample was apparent prior to treatment with iodine (Fig. 5). As expected, the activity of the ^{A}Gly-Tn was nearly identical to that of wild type rat trypsin (see Table 1). Nonetheless, the values of k_{cat} and K_M are nearly the same for both samples (±5%) but their physiologically relevant k_{cat}/K_M values are identical. These results suggest that the iodine treatment does not have any observable unfavorable effects on the kinetic properties of the mature enzyme.

Substrate specificity profile. Although the primary specificity pocket (the S1 site) is highly selective for Arg or Lys residues, the extended specificity of trypsin is more relaxed. Even so, the profile of this extended specificity can be a highly sensitive probe of enzyme structure. Positional scanning substrate combinatorial libraries are a powerful tool for exploring the extended substrate specificity of proteases.[8,1] Recently a method was developed that allows for the configuration of various substrate libraries to rapidly identify the primary and extended specificity of proteases.[7] One of these combinatorial libraries was used to compare the extended substrate specificities of wild type recombinant trypsin and ^{A}Gly-Tn. In the experimental design, each library was composed of 361 tetrapeptides (P4-P3-P2-P1) that contained an Arg in the P1 position. Positions P2, P3, and P4 were occupied by one of the 19 standard amino acids (cysteine was excluded) or norleucine (used to approximate methionine) while the remaining positions were randomized. For example, in testing for a preference at P2, the library would consist of peptides with the configuration X-X-P2-Arg where X is any amino acid and P2 is a specific amino acid. As Fig. 6 shows, wild type trypsin exhibits no preference at P2, but positively charged amino acids (Lys, Arg) and those with large hydrophobic side chains (Ile, Leu, Phe, Trp, Tyr) are disfavored. At P3, the only amino acid discriminated against is Pro and all amino acids are tolerated at P4 except for a slight disposition against Trp. Therefore, it is apparent that the catalytic efficiency of trypsin is reduced if certain amino acids occupy specific positions in the substrate. This provides a sensitive assay to verify the structural and catalytic integrity of the trypsin produced by iodine activation. After iodine treatment, if the ^{A}Gly-Tn maintained the same three-dimensional structure as trypsin, one would expect it to display the same specificity profile as wild type rat Tn. To assay for any possible changes in substrate specificity, we assayed the iodine activated sample against the same semi-randomized peptide library. Subjecting the same library to hydrolysis by the iodine activated ^{A}Gly-Tg resulted in a cleavage pattern that is indistinguishable from that of wild type rat trypsin. These results, coupled with the activity measurements against single substrates illustrate that the enzyme produced by iodine treatment of ^{A}G-Tg is identical to wild type rat Tn.

Conclusion

The present results extend earlier findings for site-specific cleavage of dihydrofolate reductase[28] and demonstrate unequivocally that protein backbone cleavage can be effected under conditions compatible with maintenance of protein function. Using similar techniques, the activation of trypsin from an essentially inactive trypsinogen analogue containing the unnatural amino acid allylglycine has been achieved by simple treatment with aqueous iodine.

The initiation of enzyme function by limited proteolysis of inactive precursors is a widespread phenomenon in prokaryotic and eukaryotic organisms. Without having to add an activating protease, this approach provides a mechanism to study such time-dependent processes mediated by trypsin or other proteins that have inactive precursors.

This technique also provides the researcher with absolute control of the activation process. With the incorporated unnatural amino acid, the zymogen form of the enzyme is chemically incapable of premature activation and consequent degradation. The need for an activating protease is eliminated which also eliminates a common source of contamination. Additionally, this technique provides a means to achieve specific cleavage at sites that may not be readily accessible to specific proteases. Several applications of the present strategy can be envisioned such as the division of multi-domain proteins into structural units to permit analysis of individual domain function.

Experimental

General methods

[^{35}S]Methionine (1000 Ci/mmol, 10 μCi/μl) was purchased from Amersham Corporation. Nuclease-treated rabbit reticulocyte lysate system and DNA polymerase I (Klenow fragment) were obtained from Promega Corporation (Madison, WI). Restriction endonucleases, purified acylated bovine serum albumin (BSA), T4 DNA ligase and T4 RNA ligase were obtained from New England Biolabs (Beverly, MA) *E. coli* competent cells were from Stratagene Cloning Systems (La Jolla, CA). Kits for plasmid isolation were purchased from PGC Scientific (Gaithersburg, MD). AmpliScribe transcription kits and T7 RNA polymerase were purchased from Epicentre Technologies (Madison, WI); plasmid pET-14b was from Novagen (Madison, WI). Acrylamide, *N*,*N*-methylenebisacrylamide, urea, Tris base, NTPs, DEAE Sepharose CL-6B, GMP, amino acids, bovine trypsin, bovine trypsinogen, trypsin substrate N^{α}-benzoyl-(L)-arginine 7-amido-4-methylcoumarin and LTI (lima bean trypsin inhibitor)-agarose resin were from Sigma.

Nitroveratryloxycarbonyl (NVOC)-protected pdCpA derivatives of valine and allylglycine were prepared as described previously.[28]

Plasmid DNAs were isolated using a Wizard Miniprep purification system (Promega) or a JetStar plasmid midi kit (PGC Scientific) according to the protocols provided. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis was carried out using the standard Laemmli procedure.[16] Gels were visualized and quantified utilizing a Molecular Dynamics 400E PhosphorImager equipped with ImageQuant version 5.0 software. Sequence analysis was performed using a Sequenase version 2.0 DNA sequencing kit (USBiochemical). All aqueous employed distilled, deionized water from a Milli-Q system.

UV spectral measurements were made using a Perkin–Elmer Lambda 20 UV/VIS spectrometer. Radioactivity measurements were performed with a Beckman LS-100C liquid scintillation counter. Fluorescence spectral measurements were made using a Hitachi F2000 fluorescence spectrophotometer. Radioactivity measurements were made using a calibrated phosphorimager screen, such that the pixel density of the image could be related directly to the amount of radioactivity present in the sample.

Construction of expression plasmid for the in vitro elaboration of rat trypsinogen analogues

Plasmids pGEM-Tg (−1) and pET-14b (Novagen) were each digested with *Nco*I and *Bam*HI. The small fragment isolated from the pGEM-Tg (−1) digestion contained the Tg gene and was subcloned into the digested pET-14b vector to give plasmid pET-Tg (−1) (Fig. 2). The nucleotide sequence of pET-Tg (−1) was verified by restriction analysis and Sanger DNA sequencing.

In vitro transcription of 5′-monophosphate-enriched suppressor tRNA[20]

Plasmid pYRNA8, encoding the yeast suppressor tRNA-$^{Phe}_{CUA}{}^{23}$ was linearized with *Fok*I and then transcribed using an AmpliScribe T7 RNA polymerase transcription kit. The *Fok*I digestion reaction typically contained 100 μg of pYRNA8 plasmid DNA, 20 mM Tris–acetate, pH 7.9, 10 mM magnesium acetate, 50 mM potassium acetate, 1 mM dithiothreitol (DTT), and 200 units of *Fok*I in a total volume of 500 μl. Reactions were incubated at 37°C for 2 h. The digested plasmid DNA was phenol–chloroform extracted, ethanol precipitated, and transcribed for 6 h at 42°C using an AmpliScribe T7 transcription kit in a 1 ml buffered reaction mixture (7.5 mM each of ATP, CTP, and UTP, 5 mM GTP, 20 mM GMP, 10 mM DTT, 20 nM template DNA, 100 μl of the T7 RNA polymerase preparation). The transcribed tRNA was precipitated with 4 volumes of ethanol, collected by centrifugation and dried under vacuum. The crude tRNA was dissolved in loading buffer (80% formamide, 100 mM EDTA, 0.02% bromophenol blue, and 0.02% xylene cyanol), applied to an 8% denaturing polyacrylamide gel (40 cm×22 cm×2 mm), and subjected to electrophoresis at 600 V for 4.5 h. The RNA was visualized by UV shadowing,[9] excised from the gel, and recovered by crush and soak[19] with 100 mM sodium acetate, pH 4.5, 1 mM EDTA and 0.01% SDS at 23°C for 16 h. The tRNA was recovered by ethanol precipitation, dried, and then re-dissolved in RNase–free water prior to use.

Synthesis of mRNA by in vitro transcription

The plasmids were linearized with *Bam*HI and transcribed using an AmpliScribe T7 transcription kit as described previously. The transcribed mRNA was precipitated by the addition of isopropyl alcohol. Unincorporated NTPs were removed chromatographically and the resulting mRNA was extracted using phenol and chloroform successively. The mRNA was precipitated with 2.5 volumes of ethanol, washed with 70% ethanol, dried and re-dissolved in RNase–free water. The sample was stored at −80°C until further use.

Chemical mis-acylation of suppressor tRNA–C_{OH}

Chemical aminoacylation reactions (0.5 mM ATP, 15 mM $MgCl_2$, 50 μg suppressor tRNA–C_{OH}, 1.0 A_{260} unit of NVOC-protected aminoacyl-pdCpA (5–10-fold molar excess), 10% dimethylsulfoxide, 200 units of T4 RNA ligase) were carried out in 100 μl total volume in 50 mM NaOH–Hepes, pH 7.5. After incubation at 37°C for 45 min, the reaction was quenched by the addition of 0.1 vol of 3 M sodium acetate, pH 5.2.The tRNA was precipitated, collected, washed and dried as before, and then re-dissolved in 1 mM KOAc to a final concentration of 3 μg/μl. The efficiency of ligation was estimated by gel electrophoresis at pH 5.0.[26] Deprotection of NVOC-containing aminoacyl-tRNAs was carried out in 1 mM potassium acetate, pH 4.5 with 5 μg/μl tRNA. The aminoacyl-tRNAs were cooled to 2°C and irradiated with a 500 W mercury–xenon lamp using both Pyrex and water filters. Typically, monosubstituted NVOC aminoacyl-tRNAs were irradiated for 2 min. The deblocked aminoacylated suppressor tRNAs were used in in vitro suppression experiments immediately following deprotection.

In vitro synthesis of trypsinogen analogues

Trypsinogen analogues were synthesized in 10–5000 μl reaction mixtures that contained the following per 100 μl: 70 μl of methionine-depleted, nuclease-treated rabbit reticulocyte lysate, 80 μCi of [^{35}S]-*S*-methionine (1000 Ci/mmol), 2 μl of a solution 1 mM in 19 amino acids used in ribosomal protein synthesis (but lacking methionine), 10 μg of the appropriate mRNA, and 20 μg (~0.8 nmol) of deprotected mis-acylated $tRNA_{CUA}$ or deacylated suppressor $tRNA_{CUA}$ as a control. Each reaction mixture was incubated at 30°C for 1.5 h. In vitro translation was also carried out without added mis-acylated $tRNA_{CUA}$ as a control. Aliquots (typically 1 μl) were taken for analysis by 20% SDS–PAGE. Autoradiography of the gels was carried out to determine the location of ^{35}S-labeled protein; quantification of the bands was carried out using a phosphorimager.

Purification of in vitro synthesized trypsinogen analogues

In vitro translation mixtures (200 μl) containing ^{35}S-labeled protein were incubated with 1 μg of RNase A for 30 min at

room temperature and then applied to a 200-μl DEAE Sepharose CL-6B column that had been equilibrated with 5 mM potassium phosphate buffer, pH 7.0. The column was washed sequentially with 4.5 volumes of the following in 5 mM potassium phosphate buffer, pH 7.0: 10 mM β-mercaptoethanol, 5 mM potassium phosphate and 50 mM KCl. The remaining proteins were washed from the column with 3 volumes of 1 M KCl in 5 mM potassium phosphate buffer.

The fractions containing radioactivity were applied to a 100-μl Ni–NTA agarose column equilibrated with 300 mM NaCl, 10 mM imidazole, and 100 μg/ml BSA in 50 mM sodium phosphate, pH 8.0. After washing the column with five volumes of equilibration buffer, the protein was eluted with eight volumes of 300 mM NaCl, 250 mM imidazole, and 10 μM BSA in 50 mM sodium phosphate, pH 8.0. The radioactive fractions were combined and dialyzed (Spectra/Por 2, MW cutoff 12–14 kDa) against 50 mM NaCl in 20 mM sodium phosphate, pH 8.0, followed by dialysis against H_2O and finally against 1 mM HCl. Each dialysis was carried out at 4°C for 4 h. The samples were then concentrated to ~20 ng/μl using a microconcentrator (Microcon YM-3, Millipore Corporation). The amount of ^{35}S-labeled protein in each fraction was determined by liquid scintillation counting of a portion of each.

Binding of trypsinogen analogues to an LTI–agarose resin

A 100 μl LTI–agarose column was equilibrated with 1 mM benzamidine in 5 mM morpholinoethanesulfonic acid (Mes), pH 6.0. The buffer for the purified protein was exchanged with the equilibration buffer by dialysis. Aliquots (20–100 μl) of ^{35}S-labeled protein were then combined with equilibrated LTI–agarose. The radioactivity was determined prior to combination. After incubation at 25°C for 0.5 h, the resin was washed with either high ionic strength buffer (0.5 M NaCl, 1 mM benzamidine, 5 mM Mes, pH 6.0) or low ionic strength buffer (1 mM benzamidine, 5 mM Mes, pH 6.0) until the radioactivity in the supernatant reached a baseline value. The resin was then combined with 20 μl of 0.1 M sodium formate, pH 5.0 and incubated at 25°C for 0.5 h. This process was repeated until no radioactivity was detected in the supernatant.

Iodine treatment of proteins

I_2 was dissolved in 50/50 THF (tetrahydrofuran)/H_2O and added to a final concentration of 0.5–2.5 mM (10–15% THF) to 100–400 μl of 2–10 pmol of target protein. The solution was incubated at 25°C for 30–60 min with constant mixing. The excess iodine was removed by dialysis into 1 mM HCl at 4°C for 2–4 h or until the iodine was completely removed. After dialysis, the sample was stored at 4°C until further use.

Determination of active ^{A}Gly-Tn concentration

Wild type rat trypsinogen (Tg) was expressed using a *P. pastoris* expression system and activated with enteropeptidase (1/20 w/w) purchased from Biozyme. Purification was carried out using hydrophobic and affinity column chromatography sequentially. The concentration of active trypsin was determined by active site titration with the fluoreginic substrate 4-methylumbelliferyl guanidine benzoate (λ_{ex}=360, λ_{em}=450) in an ISA SPEX Fluorolog-3 fluorescence detector. An activity standard curve was constructed to determine the specific activity of wild type Tn against the fluoregenic substrate Z-GPR-AMC (*N*-α-benzyloxycarbonyl-L-glycylprolylarginine 7-amido-4-methylcoumarin) (λ_{ex}=380, λ_{em}=460). The hydrolysis activity of 1–100 pM Tn was measured against 100 μM Z-GPR-AMC in 50 mM HEPES, 100 mM NaCl, 20 mM $CaCl_2$, pH 8.0 at 25°C. The concentration of active ^{A}Gly-Tn was then determined by measuring the activity of 5–20 μl of the I_2 treated sample against 100 μM Z-GPR-AMC and calculating the activity present from the previously determined specific activity of wild type rat Tn.

Assay of trypsin activity

Stock solutions of the substrates were prepared in DMSO (final concentration <3%). Initial detection of activity from the iodine treated samples was assayed as a function of *N*-benzoyl-L-arginine 7 amido-4-methylcoumarin (*N*-Bz-R-AMC, 200 μM) hydrolysis in 50 mM Tris–HCl, pH 8.0, 10 mM $CaCl_2$, 100 mM NaCl. The fluorescence of the free AMC hydrolysis product (λ_{ex}=380, λ_{em}=460) was monitored continuously using a Hitachi F2000 fluorescence spectrophotometer. For kinetic profile acquisition, the hydrolytic activity of 10 pM activated ^{A}Gly-Tn sample was assayed against 1–100 μM Z-GPR-AMC at 25°C in 50 mM HEPES, 0.1 M NaCl, 20 mM $CaCl_2$, pH 8.0 using an ISA Spex Fluorog-3 fluorometer. The catalytic rate constants were determined by plotting the initial rates as a function of substrate concentration and fitting the data to the Michaelis–Menten equation using the curve-fitting program Kaleidagraph (Synergy Software).

Assay of protection against iodine inactivation

Wild type rat trypsin was added to a rabbit reticulocyte lysate mixture similar to that used for in vitro translation of ^{A}Gly-Tg and subjected to the purification scheme previously outlined. The sample was then subjected to iodine treatment, quantified by active site titration and assayed for activity using Z-GPR-AMC as described.

Substrate specificity profile

The activity of wild type Tn or ^{A}Gly-Tn was assayed against a semi-randomized tetrapeptide substrate library in a 96-well plate. Each well contained 100 μl of 10 pM ^{A}Gly-Tn or 10–50 pM wild type Tn and 361 substrates per well at concentration of 250 nM/substrate. Each substrate had two positions fixed; P1 was arginine in all samples while P2, P3, or P4 were fixed in one well with one of 19 standard amino acids (all standard amino acids were used except cysteine and methionine. Norleucine was added to approximate methionine). The samples were monitored fluorometrically (λ_{ex}=380, λ_{em}=460) at 25°C for 90 min in a Molecular Devices SpectrMax Genini microtiter plate reader.

Acknowledgements

This work was supported at the University of Virginia by Research Grant CA77359 from the National Cancer Institute (B. W., M. L., and S. M. H.) and at UC San Francisco by National Science Foundation Research Grant MCB9604379 (C. S. C.) and by a fellowship from the National Science Foundation (T. T. B.). The authors would also like to thank Jennifer L. Harris, Brad Backes, and Jonathan A. Ellman for providing the combinatorial libraries.

References

1. Backes, B. J.; Harris, J. L.; Leonetti, F.; Craik, C. S.; Ellman, J. A. *Nat. Biotechnol.* **2000**, *18*, 187–193.
2. Bain, J. D.; Switzer, C.; Chamberlin, A. R.; Benner, S. A. *Nature* **1992**, *356*, 537–539.
3. Bode, W.; Schwager, P.; Huber, R. *J. Mol. Biol.* **1978**, *118*, 99–112.
4. Brodrick, J. W.; Largman, C.; Johnson, J. H.; Geokas, M. C. *J. Biol. Chem.* **1978**, *253*, 2732–2736.
5. Devillers-Thiery, A.; Kindt, T.; Scheele, G.; Blobel, G. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 5016–5020.
6. Fehlhammer, H.; Bode, W.; Huber, R. *J. Mol. Biol.* **1977**, *111*, 415–438.
7. Harris, J. L.; Backes, B. J.; Leonetti, F.; Mahrus, S.; Ellman, J. A.; Craik, C. S. *PNAS* **2000**, *97*, 7754–7759.
8. Harris, J. L.; Peterson, E. P.; Hudig, D.; Thornberry, N. A.; Craik, C. S. *J. Biol. Chem.* **1998**, *273*, 27364–27373.
9. Hassur, S. M.; Whitlock Jr., H. W. *Anal. Biochem.* **1974**, *59*, 162–164.
10. Heckler, T. G.; Zama, Y.; Naka, T.; Hecht, S. M. *J. Biol. Chem.* **1983**, *258*, 4492–4495.
11. Higaki, J. N.; Evnin, L. B.; Craik, C. S. *Biochemistry* **1989**, *28*, 9256–9263.
12. Huber, R.; Bode, W. *Acc. Chem. Res.* **1978**, *11*, 114.
13. Janknecht, R.; de Martynoff, G.; Lou, J.; Hipskind, R. A.; Nordheim, A.; Stunnenberg, H. G. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 8972–8976.
14. Karginov, V. A.; Mamaev, S. V.; An, H.; Van Cleve, M. D.; Hecht, S. M.; Komatsoulis, G. A.; Abelson, J. N. *J. Am. Chem. Soc.* **1997**, *119*, 8166–8176.
15. Kossiakoff, A. A.; Chambers, J. L.; Kay, L. M.; Stroud, R. M. *Biochemistry* **1977**, *16*, 654–664.
16. Laemmli, U. K. *Nature* **1970**, *227*, 680–685.
17. Louvard, M. N.; Puigserver, A. *Biochim. Biophys. Acta* **1974**, *371*, 177–185.
18. Maroux, S.; Baratti, J.; Desnuelle, P. *J. Biol. Chem.* **1971**, *246*, 5031–5039.
19. Maxam, A. M.; Gilbert, W. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 560–564.
20. Nazarenko, I. A.; Harrington, K. M.; Uhlenbeck, O. C. *EMBO J.* **1994**, *13*, 2464–2471.
21. Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Science* **1989**, *244*, 182–188.
22. Perona, J. J.; Craik, C. S. *Protein Sci.* **1995**, *4*, 337–360.
23. Robertson, S. A.; Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Nucleic Acids Res.* **1989**, *17*, 9649–9660.
24. Robinson, N. C.; Neurath, H.; Walsh, K. A. *Biochemistry* **1973**, *12*, 420–425.
25. Sampson, J. R.; Uhlenbeck, O. C. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 1033–1037.
26. Varshney, U.; Lee, C. P.; RajBhandary, U. L. *J. Biol. Chem.* **1991**, *266*, 24712–24718.
27. Walter, P.; Blobel, G. *J. Cell Biol.* **1981**, *91*, 557–561.
28. Wang, B.; Lodder, M.; Baird Jr., T.; Brown, K. C.; Craik, C. S.; Hecht, S. M. *J. Am. Chem. Soc.* **2000**, *122*, 7402–7403.