

# Deciphering the Role of the Electrostatic Interactions Involving Gly70 in Eglin C by Total Chemical Protein Synthesis

Wei-Yue Lu,<sup>‡,§</sup> Melissa A. Starovasnik,<sup>||</sup> John J. Dwyer,<sup>||,⊥</sup> Anthony A. Kossiakoff,<sup>||,⊥</sup> Stephen B. H. Kent,<sup>‡</sup> and Wuyuan Lu<sup>\*,‡,⊥</sup>

Gryphon Sciences, 250 East Grand Avenue, Suite 90, South San Francisco, California 94080, and Department of Protein Engineering, Genentech, Inc., 1 DNA Way, South San Francisco, California 94080

Received October 4, 1999; Revised Manuscript Received December 14, 1999

**ABSTRACT:** Eglin c from the leech *Hirudo medicinalis* is a potent protein inhibitor of many serine proteinases including chymotrypsin and subtilisins. Unlike most small protein inhibitors whose solvent-exposed enzyme-binding loop is stabilized primarily by disulfide bridges flanking the reactive-site peptide bond, eglin c possesses an enzyme-binding loop supported predominantly by extensive electrostatic/H-bonding interactions involving three Arg residues (Arg48, Arg51, and Arg53) projecting from the scaffold of the inhibitor. As an adjacent residue, the C-terminal Gly70 participates in these interactions via its  $\alpha$ -carboxyl group interacting with the side chain of Arg51 and the main chain of Arg48. In addition, the amide NH group of Gly70 donates an H-bond to the carbonyl C=O groups of Arg48 and Arg51. To understand the structural and functional relevance of the electrostatic/H-bonding network, we chemically synthesized wild-type eglin c and three analogues in which Gly70 was either deleted or replaced by glycine amide ( $\text{NH}_2\text{CH}_2\text{CONH}_2$ ) or by  $\alpha$ -hydroxylacetamide ( $\text{HOCH}_2\text{CONH}_2$ ). NMR analysis indicated that the core structure of eglin c was maintained in the analogues, but that the binding loop was significantly perturbed. It was found that deletion or replacement of Gly70 destabilized eglin c by an average of 2.7 kcal/mol or 20 °C in melting temperature. As a result, these inhibitors become substrates for their target enzymes. Binding assays on these analogues with a catalytically incompetent subtilisin BPN' mutant indicated that loss or weakening of the interactions involving the carboxylate of Gly70 caused a decrease in binding by approximately 2 orders of magnitude. Notably, for all four synthetic inhibitors, the relative free energy changes ( $\Delta\Delta G$ ) associated with protein destabilization are strongly correlated (slope = 0.94,  $r^2 = 0.9996$ ) with the  $\Delta\Delta G$  values derived from a decreased binding to the enzyme.

Most small protein inhibitors of serine proteinases obey a so-called “standard mechanism” by which they interact, like a good substrate, with their cognate enzymes by fitting the exposed reactive-site binding loop into the active-site cleft of the enzyme (1–3). Even though the reactive-site binding loop in inhibitors, upon binding to the enzyme, adopts the same complementary, canonical conformation as it does in substrates, the turnover rate (enzymatic cleavage at the scissile reactive-site peptide bond) for inhibitors is in general extremely slow.

A slow turnover rate has often been suggested to be attributable to the rigidity of the binding loop. This is understandable because a rigid reactive-site peptide bond will effectively elevate the activation energy barrier to the transition state, deterring a productive nucleophilic attack on the peptide bond by the enzyme. On the other hand, since the binding loop of inhibitors becomes more rigid upon complex formation, as observed in high-resolution structural

analyses, certain flexibility in the binding loop seems to be necessary for an inhibitor to be able to target a number of different enzymes (2). It has been suggested that the interplay of rigidity and flexibility of the binding loop plays an important role in defining an inhibitor's binding action (4).

The rigidity of the binding loop in most small protein proteinase inhibitors is provided by covalent interactions between the loop region and the rest of the molecule through disulfide bridges flanking the reactive-site peptide bond. Eglin c of the leech *Hirudo medicinalis* (5, 6), a 70-residue protein inhibitor of many serine proteinases such as human leukocyte elastase, human cathepsin G, bovine  $\alpha$ -chymotrypsin, and subtilisins (7, 8), is clearly an exception. Three-dimensional structures of eglin c (9, 10) and eglin c complexes with a number of serine proteinases have been reported, including bovine  $\alpha$ -chymotrypsin (11), subtilisin Carlsberg (12, 13), subtilisin BPN' (14), and thermitase (15). More recently, a chemically synthesized truncated form of eglin c, (8–70)S41C, has been crystallized in complex with subtilisin BPN', and the crystal structure solved at 2.0 Å resolution (4). As seen in all the crystal structures, the reactive-site binding loop in eglin c is supported primarily by a network of electrostatic/H-bonding interactions between Thr44, Asp46 in the loop region, and Arg51, Arg53 projecting from the hydrophobic core (Figure 1). Additional

\* To whom correspondence should be addressed. Phone: (773) 834-3651. FAX: (773) 834-2777. E-mail: wuyuanlu@midway.uchicago.edu.

<sup>‡</sup> Gryphon Sciences.

<sup>§</sup> Current address: School of Pharmacy, Shanghai Medical University, Shanghai 200032, China.

<sup>||</sup> Genentech, Inc.

<sup>⊥</sup> Current address: Department of Biochemistry and Molecular Biology, The University of Chicago, 920 E. 58th St., Chicago, IL 60637.

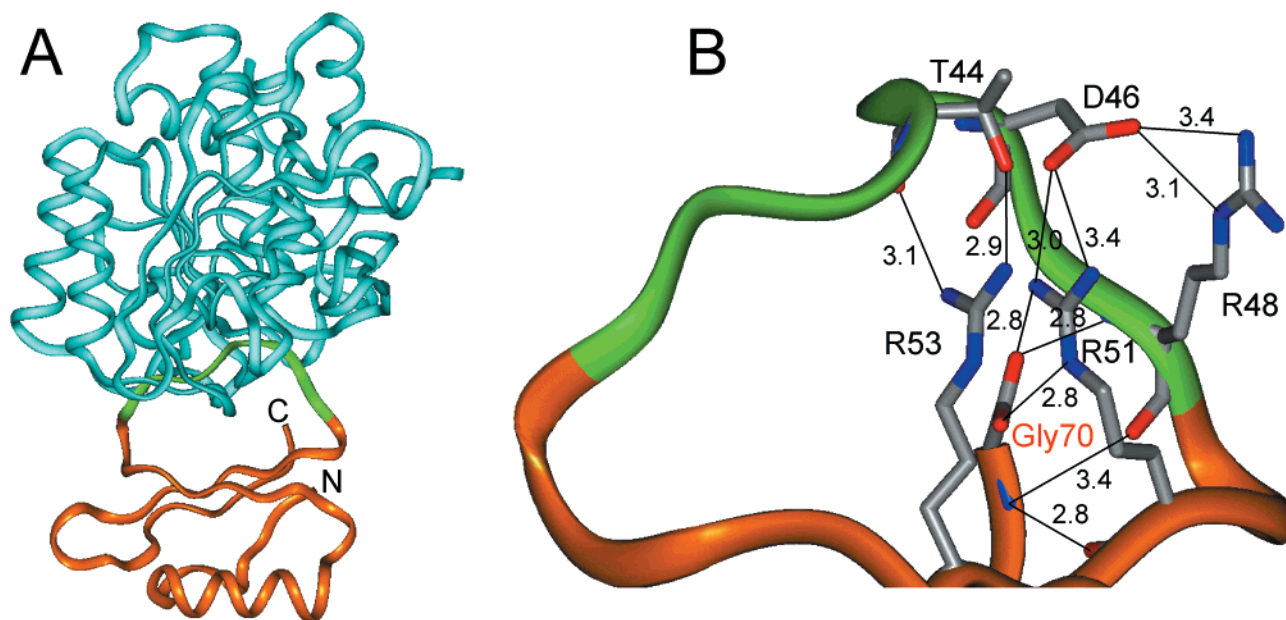


FIGURE 1: Panel A: X-ray crystal structure of synthetic (8–70)S41C eglin c in complex with SBPN (4). The enzyme-binding loop of eglin c is depicted in green. Panel B: Close-up view of the electrostatic interactions that support the binding loop of eglin c, as seen in the crystal structure of the complex. The nitrogen and oxygen atoms are depicted in blue and red, respectively. The C-terminal Gly70 participates in these interactions via its  $\alpha$ -carboxyl group forming a salt bridge with the side chain of Arg51 and accepting an H-bond donated by Arg48N. In addition, the amide NH group of Gly70 donates a bifurcated H-bond to the carbonyl C=O groups of Arg48 and Arg51.

structuring of the binding loop comes from interactions between Asp46 and Arg48. Mutations involving Thr44Pro, Asp46Ser, Arg51Lys, and Arg53Lys cause either a decrease in inhibition or a transformation of the inhibitor into a substrate, presumably due to the weakening of the H-bonding network and a corresponding increase in flexibility of the enzyme-binding loop (16).

The C-terminal Gly70 is adjacent to the loop region in the folded state of the protein, and also forms an important element of the structure. Specifically, the Gly residue participates in the extensive electrostatic/H-bonding network by interacting its  $\alpha$ -carboxyl group with Arg51N<sup>6</sup> and Arg51N<sup>72</sup> (a salt bridge), and by accepting an H-bond donated by Arg48N (Figure 1). In addition, the amide NH group of Gly70 forms an H-bond to the carbonyl C=O groups of Arg48 and Arg51. To understand the structural and functional importance of these interactions involving Gly70 in eglin c, we chemically synthesized, via native chemical ligation (17), (1–70)S41C eglin c and three analogues in which Gly70 was either deleted or replaced by glycine amide or by  $\alpha$ -hydroxylacetamide. We report here a characterization of the covalent and three-dimensional structure of the four synthetic proteins by electrospray ionization mass spectrometry and <sup>1</sup>H NMR<sup>1</sup> spectroscopy, and an in-depth evaluation of their kinetic and biophysical properties by inhibitory activity assay, calorimetric thermal denaturation, and GuHCl-induced denaturation monitored by circular dichroism. Our results show that the electrostatic/H-bonding interactions involving Gly70 are critical in the inhibitory activity of eglin c. This supports the premise that the rigidity of the binding loop in proteinase inhibitors plays an important role in turning an inhibitor into a substrate or vice versa.

## EXPERIMENTAL PROCEDURES

**Materials.** Boc-L-amino acids were purchased from Peptide Institute, Inc., Osaka, Japan. MBHA-resin and HBTU were

purchased from Novabiochem; Boc-Gly-OCH<sub>2</sub>-PAM-resin, Boc-Val-OCH<sub>2</sub>-PAM-resin, and DIEA were obtained from Applied Biosystems. Boc-Gly-(thioester linker)-amino-methyl-resin was prepared by Gina Figliozzi according to published procedures (18, 19). Thiophenol, bromoacetic acid, and glycolic acid were purchased from Aldrich Chemical Co. CHYM was purchased from Worthington Biochemical Co. SCAR was obtained from Sigma Chemical Co. Genenase, a catalytically incompetent H64A mutant of SBPN (20, 21), was generously provided by the Wells laboratory at Genentech, Inc. The chromagenic substrate Suc-GGF-pNA used in enzyme activity assays was purchased from Bachem Bioscience Inc. Solvents commonly used in peptide synthesis and purification, namely, TFA, DMF, DCM, and ACN, were obtained in HPLC grade or better.

**Methods.** Analytical and preparative reversed-phase HPLC was performed on a Rainin Dynamax HPLC system equipped with interchangeable pump heads, an analytical Vydac C-18 column (5  $\mu$ m, 4.6  $\times$  150 mm), and a preparative Vydac C-18 column (15–20  $\mu$ m, 22  $\times$  250 mm). Solvent A for HPLC purification was water containing 0.1% TFA, solvent B 100% ACN containing 0.1% TFA. Mass spectrometry analysis was carried out on a PE Sciex API-I electrospray

<sup>1</sup> Abbreviations: 2D, two-dimensional; ACN, acetonitrile; Boc, *tert*-butoxycarbonyl; CD, circular dichroism; CHYM, bovine  $\alpha$ -chymotrypsin; COSY, correlation spectroscopy; DCM, dichloromethane; DIEA, *N,N*-diisopropylethylamine; DMF, dimethylformamide; DMSO, dimethyl sulfoxide; DSC, differential scanning calorimeter; DSS, 3-(trimethylsilyl)-1-propane-1,1,2,2,3,3-*d*<sub>6</sub> sulfonic acid; ESI-MS, electrospray ionization mass spectrometry; HBTU, 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; MBHA, 4-methylbenzhydrylamine; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; PAM, 4-hydroxymethylphenylacetamidomethyl; RP-HPLC, reversed-phase high-pressure/performance liquid chromatography; SBPN, subtilisin BPN'; SCAR, subtilisin Carlsberg; SPPS, solid-phase peptide synthesis; Suc-GGF-pNA, succinyl-Gly-Gly-Phe-*p*-nitroanilide; TFA, trifluoroacetic acid; TOCSY, total correlation spectroscopy.

ionization mass spectrometer (ESI-MS), with a direct sample infusion by a syringe pump at 5  $\mu\text{L}/\text{min}$ . Thermal denaturation was performed in 0.2 M  $\text{KH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$  buffer, pH 5.4, on a Calorimetry Sciences Nano differential scanning calorimeter (DSC) with a scanning rate of 1  $^\circ\text{C}/\text{min}$ .

**Solid-Phase Peptide Synthesis (SPPS).** The strategies for the syntheses of (1–70)S41C eglin c and three analogues [(1–69)S41C, (1–70) $\alpha\text{CONH}_2$ -S41C, and (1-[O]70) $\alpha\text{CONH}_2$ -S41C] were essentially as described in the previous publication (4). The amino acid sequence of naturally occurring eglin c is as follows: TEFGSELKSF PEVVGKTVDQAREY-FTLHPYQYNVYFLPEG SPVTLDLRYN RVRVFNPGT NVNVHVPVHG. All four synthetic proteins contain a Ser41→Cys mutation in order to introduce a ligation site between Gly40 and Cys41.

The peptide segments, (1–40) $\alpha\text{COSH}$ , (41–70)S41C, (41–69)S41C, (41–70) $\alpha\text{CONH}_2$ -S41C, and (41-[O]70)- $\alpha\text{CONH}_2$ -S41C, were all synthesized manually in stepwise fashion using the published in situ neutralization/HBTU activation protocol for Boc chemistry SPPS (22). The two  $\alpha$ -carboxylate-ending peptides, (41–70)S41C and (41–69)-S41C, were synthesized on Boc-Gly-OCH<sub>2</sub>-PAM-resin and Boc-Val-OCH<sub>2</sub>-PAM-resin, respectively. While the thio acid peptide (1–40) $\alpha\text{COSH}$  was made on Boc-Gly-(thioester linker)-aminomethyl-resin, MBHA-resin was used for the two peptides ending with  $\alpha\text{-CONH}_2$ . Couplings of glycolic acid and the subsequent amino acid were carried out according to the procedure published previously (23). The following side chain protection was used: Arg(Tosyl); Asn(Xanthyl); Asp(OcHxl); Cys(4MeBzl); Glu(OcHxl); His(DNP); His-(Bom); Lys(2ClZ); Ser(Bzl); Thr(Bzl); Tyr(BrZ).

After HF deprotection and cleavage from the resin, the crude peptides were purified by preparative RP-HPLC. The thioester peptide (1–40) $\alpha\text{COSCH}_2\text{COOH}$  used in the ligation reactions was obtained by reacting crude (1–40) $\alpha\text{COSH}$  (10 mg/mL) with bromoacetic acid (2 mg/mL) in 0.1 M sodium acetate buffer containing 6 M GuHCl, pH 4.0, for 45 min. The resultant product was subsequently purified by preparative RP-HPLC. Molecular weights of all the peptides were verified by ESI-MS (data not shown).

**Native Chemical Ligation.** Native chemical ligation of (1–40) $\alpha\text{COSCH}_2\text{COOH}$  and the four C-terminal peptides, (41–70)S41C, (41–69)S41C, (41–70) $\alpha\text{CONH}_2$ -S41C, and (41-[O]70) $\alpha\text{CONH}_2$ -S41C, was carried out individually at a total peptide concentration of 10 mg/mL in 0.1 M sodium phosphate buffer containing 2% (v/v) thiophenol and 6 M GuHCl, pH 7.5. The principles of native chemical ligation reactions have been described in detail elsewhere (17, 24). Typically the ligation reaction proceeded to completion within 8 h. The final ligation product was purified by preparative C-18 RP-HPLC, and its molecular weight was determined to be within experimental error of expected values by ESI-MS (Table 1).

**Protein Folding and Purification.** The ligated synthetic polypeptide was dissolved at 1.2 mg/mL in 6 M GuHCl, to which 12 mM MES buffer, pH 6.0, was added to make a rapid 6-fold dilution. After 2 h, the folding solution was then loaded onto a methyl-CHYM-Sepharose FF affinity column prepared according to Ryan and Feeney et al. (25) and the coupling procedures provided by Pharmacia. Following washes by 0.2 M NaCl, 0.1 M HCl containing 0.2 M NaCl was used to elute the active, folded protein. The resultant

fractions were dialyzed against 1 mM HCl and lyophilized. The overall yield is 5%; i.e., about 70 mg of purified, folded protein can be obtained from a typical 0.2 mmol scale synthesis.

**NMR Spectroscopy.** All NMR experiments were acquired on a Bruker AMX-500 spectrometer equipped with a 5 mm inverse triple resonance probe with three-axis gradients at 27  $^\circ\text{C}$ . Standard pulse sequences (26) were used to collect 2D homonuclear COSY, TOCSY (83 ms mixing time), and NOESY (100 ms mixing time) spectra of wild type and all three eglin c analogues. Samples were prepared at a concentration of 1–2 mM by dissolving each protein in 92% H<sub>2</sub>O, 8% D<sub>2</sub>O, pH 5.7. All spectra were processed and analyzed using FELIX 97.0 (Molecular Simulations, Inc.).  $^3J_{\text{H}^{\text{N}}-\text{H}^{\alpha}}$  coupling constants were determined using a Levenberg–Marquardt line-fitting algorithm (27) to measure splittings from the antiphase double-Lorentzian cross sections of  $\text{H}^{\text{N}}-\text{H}^{\alpha}$  cross-peaks in COSY spectra processed with high digital resolution in  $F_2$ . Chemical shifts were referenced to internal DSS.

**Inhibitory Activity Assays.** Inhibition assays were conducted on a Hewlett-Packard HP8453 spectrophotometer at room temperature in 50 mM Bis-Tris propane buffer containing 20 mM CaCl<sub>2</sub>, 0.005% Triton X-100, 0.01%  $\beta$ -mercaptoethanol, pH 7.0. Equal molar amounts of enzyme and inhibitor were mixed in 3 mL cuvettes with a final enzyme or inhibitor concentration of  $2.1 \times 10^{-7}$  M. After an incubation of different time lengths (0, 30 s, 1 min, 2 min, 4 min, 6 min, 10 min, 30 min, 1 h, 2 h, 4 h, 6 h, 12 h, 18 h, and 24 h), 10  $\mu\text{L}$  of Suc-GGF-pNA dissolved in DMSO at 25 mg/mL was added. Enzymatic activity was then monitored by following the hydrolysis of Suc-GGF-pNA at  $\lambda = 380$  nm with a background subtraction averaging over the range from 650 to 700 nm. Initial slopes of the substrate hydrolysis curves were converted to percent residual enzyme activity, and were plotted versus incubation time.

**Binding Assay with Genenase.** Association equilibrium constants of eglin c analogues were determined by surface plasmon resonance on a BIAcore 2000 instrument (BIAcore, Inc.). Analogues (50  $\mu\text{g}/\text{mL}$  in 10 mM sodium acetate buffer, pH 4.5) were covalently coupled to a CM5 carboxylated dextran chip using *N*-hydroxysuccinimide (NHS) and *N*-ethyl-*N'*-(3-dimethylamino)propyl]carbodiimide hydrochloride (EDC) chemistry. Uncoupled NHS-ester groups were blocked with 1 M ethanolamine, and about 100–200 resonance units (RU's) of each eglin c analogue were coupled to the sensor chip by this procedure. Interaction analysis was carried out at 25  $^\circ\text{C}$  in 10 mM HEPES, 150 mM NaCl, 0.005% surfactant P20, pH 7.4 (HBS-P Buffer, BIAcore Inc.). Concentrations of Genenase ranging from 0.5 to 100 nM were injected at a flow rate of 20  $\mu\text{L}/\text{min}$  over each flow cell, and binding was experimentally determined by comparing the number of RU's observed as a function of time to an activated and blocked flowcell containing no covalently bound eglin c.

**GuHCl-Induced Denaturation.** Protein denaturation in guanidine hydrochloride (GuHCl) was carried out at 25  $^\circ\text{C}$  using an Aviv 62A DS circular dichroism spectrometer equipped with a Hamilton Microlab 500 dual-syringe liquid diluter/dispenser. An initial 1.8 mL of protein solution prepared in 10 mM  $\text{KH}_2\text{PO}_4/\text{NaOH}$ , pH 5.4, with an absorbance of 0.8 at 222 nm (approximately 7.5  $\mu\text{M}$ ) was



Table 1: Calculated Average Isotopic Masses and Observed Masses Determined by ESI-MS

	(1–70)S41C eglin c	(1–70) $\alpha$ CONH <sub>2</sub> -S41C eglin c	(1–[O]70) $\alpha$ CONH <sub>2</sub> -S41C eglin c	(1–69)S41C eglin c
calculated	8106.1	8105.2	8106.1	8049.1
observed	8106.5 $\pm$ 0.5	8105.8 $\pm$ 0.6	8106.7 $\pm$ 0.5	8049.6 $\pm$ 0.5

aliquoted into a 3 mL cuvette. Under constant stirring, an increasing amount of aliquot was withdrawn, followed immediately by an injection of an equal amount of protein solution of the same concentration, prepared in 10 mM KH<sub>2</sub>-PO<sub>4</sub>/NaOH, 7.2 M GuHCl, pH 5.4. After each withdrawal/injection, the solution in the cuvette was mixed for 30 s before signals were sampled over a 40 s period. A typical run performs 30 withdrawals/injections, generating a linear gradual increase in the concentration of GuHCl in the cuvette from 0 to 7 M. Aliquot withdrawals and injections as well as data acquisition were fully automated with the control software Star 3.0 and Igor Pro provided by Aviv Instruments, Inc.

For a two-state protein denaturation process:



$$\Delta G = -RT \ln \frac{[D]}{[N]} = -RT \ln \frac{f_D}{1 - f_D}$$

where  $f_D$  stands for the fraction denatured ( $0 \leq f_D \leq 1$ ). If  $\theta_N$  and  $\theta_D$  represent the CD signal (ellipticity at 222 nm) of the native form and of the denatured form, respectively, then at a given concentration of GuHCl ( $m$ ) and the correspondent CD signal ( $\theta$ ),  $f_D$  can be expressed as

$$f_D = \theta - \theta_N / \theta_D - \theta_N$$

Therefore

$$\Delta G^m = -RT \ln \frac{\theta - \theta_N}{\theta_D - \theta}$$

The equation can be rearranged as follows:

$$\theta = \frac{\theta_N + \theta_D \exp(-\Delta G^m/RT)}{1 + \exp(-\Delta G^m/RT)}$$

Since,  $\Delta G^m = \Delta G^0(1 - m/m_{1/2})$  where  $\Delta G^0$  is the free energy change of unfolding at  $m = 0$  and  $m_{1/2}$  is the characteristic concentration of GuHCl when  $[D] = [N]$ ,  $\theta$  can be therefore defined by the following equation:

$$\theta = \frac{\theta_N + \theta_D \exp[-\Delta G^0(1 - m/m_{1/2})/RT]}{1 + \exp[-\Delta G^0(1 - m/m_{1/2})/RT]} = \frac{(a + bm) + (c + dm) \exp[\Delta G^0(m/m_{1/2} - 1)/RT]}{1 + \exp[\Delta G^0(m/m_{1/2} - 1)/RT]}$$

where  $R = 1.987 \text{ cal} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$ ,  $T = 298.2 \text{ K}$ .

## RESULTS

All the synthetic eglin c constructs described here contain a Ser to Cys mutation at position 41 (P5). This mutation was used to introduce a ligation site at Gly40-Cys41. Since the side chain of the P5 residue makes no direct contact with

the enzyme, as indicated by high-resolution structural analysis, the Ser-to-Cys mutation was expected to have no functional consequences. This was confirmed experimentally (4).

Previously, we synthesized a truncated form, (8–70)S41C eglin c, using the native chemical ligation technique. The X-ray crystal structure of (8–70)S41C eglin c in complex with SBPN has been solved at 2.0 Å (4). The overall structure of the synthetic protein was identical to that of recombinant and naturally occurring forms. In this work, we used native chemical ligation to form a full-length synthetic (1–70)-S41C eglin c from (1–40) $\alpha$ COSR and (41–70)S41C. Analogues were simply made by syntheses of the appropriate segments as described under Experimental Procedures.

**ESI-MS.** All four synthetic proteins were analyzed on RP-HPLC to ensure chromatographic purity. The molecular masses and purity of the synthetic proteins were also determined by ESI-MS, and the results are tabulated in Table 1 along with calculated average isotopic masses. The observed masses, though well within the experimental error ( $\leq 0.01\%$ ), are consistently higher by half an atomic mass unit (amu) than the expected values, indicating a systematic deviation resulting probably from a small calibration error of the mass spectrometer used. Nevertheless, the ESI-MS results corroborate well with the modifications made at the position of Gly70, attesting to the quality of the synthetic products.

**NMR Analysis.** 2D NMR spectroscopy was used to assess the structural consequences of modifying the C-terminus of eglin c. Initial comparison of even the simplest modification, that of changing the C-terminal carboxylate to an amide, revealed surprisingly large changes in chemical shifts (Figure 2). The spectra of the three analogues, while quite different from wild type, are all rather similar to each other (Figure 3). Therefore, detailed analyses of only wild type (1–70)-S41C and (1–70) $\alpha$ CONH<sub>2</sub>-S41C are presented. The backbone proton resonances of wild-type (1–70)S41C eglin c were assigned primarily by inspection from resonance assignments reported previously (28); assignments were confirmed by sequential NOE analysis (29). Backbone assignments for the (1–70) $\alpha$ CONH<sub>2</sub>-S41C analogue were obtained by standard 2D analysis relying on sequential NOEs for sequence-specific assignments (29).

The chemical shift differences for the amide and  $\alpha$  protons of each residue observed between wild-type and (1–70) $\alpha$ CONH<sub>2</sub>-S41C eglin c are shown in Figure 4A; those residues that undergo a large change in chemical shift (i.e., a combined H $^{\alpha}$  and H $^N$  chemical shift change of greater than 0.15 ppm) are colored in Figure 4B. Residues exhibiting large chemical shift differences are localized to the binding loop and the four strands of the  $\beta$ -sheet, including primarily residues proximal to Gly70 (Figure 4B). Interestingly, the residues that have the largest change in chemical shift between wild type and the analogue are Thr44, Arg48, Asn50, Arg51, and Arg53. These residues make intimate

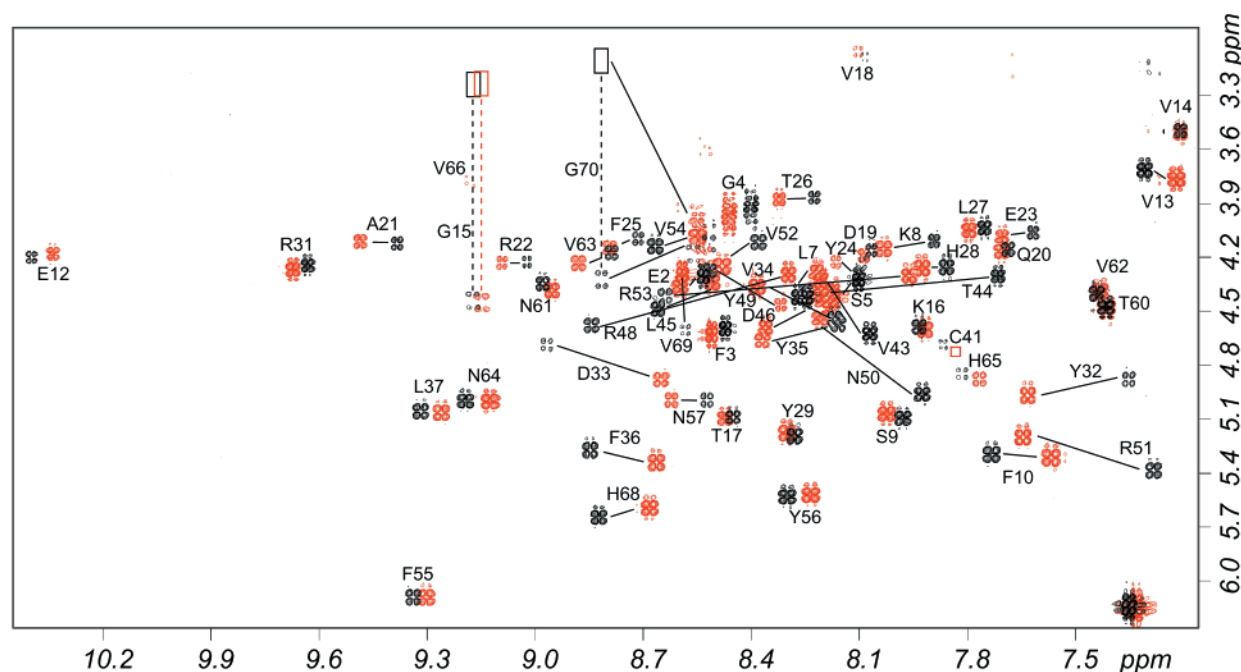


FIGURE 2: Fingerprint region of the 2D COSY spectra for wild-type (1–70)S41C (black) and (1–70) $\alpha$ CONH<sub>2</sub>-S41C (red) eglin c. Spectra were acquired under identical conditions as indicated under Experimental Procedures. Lines are shown to indicate the change in resonance positions in the two spectra. Boxes indicate peak positions for resonances that are too weak to be visible in this spectrum.

contact with the C-terminal carboxylate in the structure of enzyme-bound eglin c either directly through hydrogen-bonding and/or van der Waals interactions or indirectly through the extended hydrogen-bonding network (4, 11–15) (Figure 1).

Despite the significant chemical shift changes in the  $\beta$ -sheet, all the cross-strand NOEs expected from the wild-type structure are observed in the analogue, indicating that the secondary structure has been preserved. Furthermore, analysis of backbone  $^3J_{\text{H}^{\text{N}}-\text{H}^{\alpha}}$  coupling constants indicates that the only residues (that are sufficiently resolved in both spectra) that have significantly different values between wild type and (1–70) $\alpha$ CONH<sub>2</sub>-S41C are located within the binding loop. Val43 has a 9.0 Hz coupling constant in wild type that is only 7.7 Hz in (1–70) $\alpha$ CONH<sub>2</sub>-S41C; Arg48 has a 9.1 Hz coupling constant that is only 7.0 Hz in the analogue. Thus, residues in the binding loop that are fully extended in wild-type eglin c exhibit signs of conformational averaging in the analogue. There is the formal possibility that the different coupling constants observed for the analogue reflect an altered, but still rigid, conformation; however, this is less likely, particularly given that the analogue has decreased stability (see below). Hence, our analysis indicates that the structure of the core of the protein is relatively unperturbed, but suggests that the binding loop has changed significantly and is less rigid in molecules where the C-terminal carboxylate group has been modified.

**Inhibition Assays.** Wild-type (1–70)S41C eglin c is a potent inhibitor of both CHYM and SCAR. The inhibition constants ( $K_i$ ) of the wild-type inhibitor with respect to CHYM and SCAR have been determined to be in the  $10^{-12}$  M range at pH 7.0 (4). Attempts to measure the binding constants for the three eglin c analogues, nonetheless, failed due to hydrolysis of the inhibitors by the enzymes during incubation. For this reason, hydrolysis kinetics were used to compare the relative resistance of the eglin c analogues to

proteolytic cleavage by CHYM and SCAR. As illustrated in Figures 5 and 6, the wild-type inhibitor (1–70)S41C eglin c formed a stable complex with both enzymes over a 24 h incubation period. However, after an initial stoichiometric inhibition by the three analogues, CHYM regained 100% activity within 24 h, and SCAR within 6 h, indicating that the inhibitors were hydrolyzed and subsequently dissociated from the inhibitor–enzyme complexes initially formed. The half-lives ( $t_{1/2}$ ) for (1–69)S41C and (1–70) $\alpha$ CONH<sub>2</sub>-S41C with respect to CHYM were about 4 h, whereas the  $t_{1/2}$  for (1–[O]70) $\alpha$ CONH<sub>2</sub>-S41C with CHYM was only 2 h. Notably, hydrolysis of the analogues by SCAR is much faster than that by CHYM. While half of (1–70) $\alpha$ CONH<sub>2</sub>-S41C was hydrolyzed by SCAR within 30 min, the  $t_{1/2}$  for both (1–69)S41C and (1–[O]70) $\alpha$ CONH<sub>2</sub>-S41C was merely 5 min. It should be pointed out that in the absence of inhibitors, both enzymes, under the same conditions used above, showed full activity after 24 h.

Hydrolysis analysis was also carried out using analytical RP-HPLC and ESI-MS to follow cleavage products (data not shown). The results were in excellent agreement with the kinetic data. Interestingly, both HPLC and ESI-MS results unambiguously demonstrated that the first three residues, Thr-Glu-Phe, of the full-length wild type (1–70)S41C eglin c were slowly cleaved by CHYM during incubation (80% cleaved in 24 h). On the contrary, no cleavage at any position was detected with SCAR during the same period of time. Similar results were also observed with (1–69)S41C eglin c. While the first three residues, Thr-Glu-Phe, in (1–69)-S41C eglin c were removed by CHYM within 10 min of incubation, no cleavage by SCAR other than hydrolysis at the reactive-site peptide bond (P1–P1') was detected during the first 5 min, even though 50% of the protein was shown to have already been hydrolyzed at the reactive site. These observations are in conflict with earlier suggestions that the absence of electron density for the first seven residues of

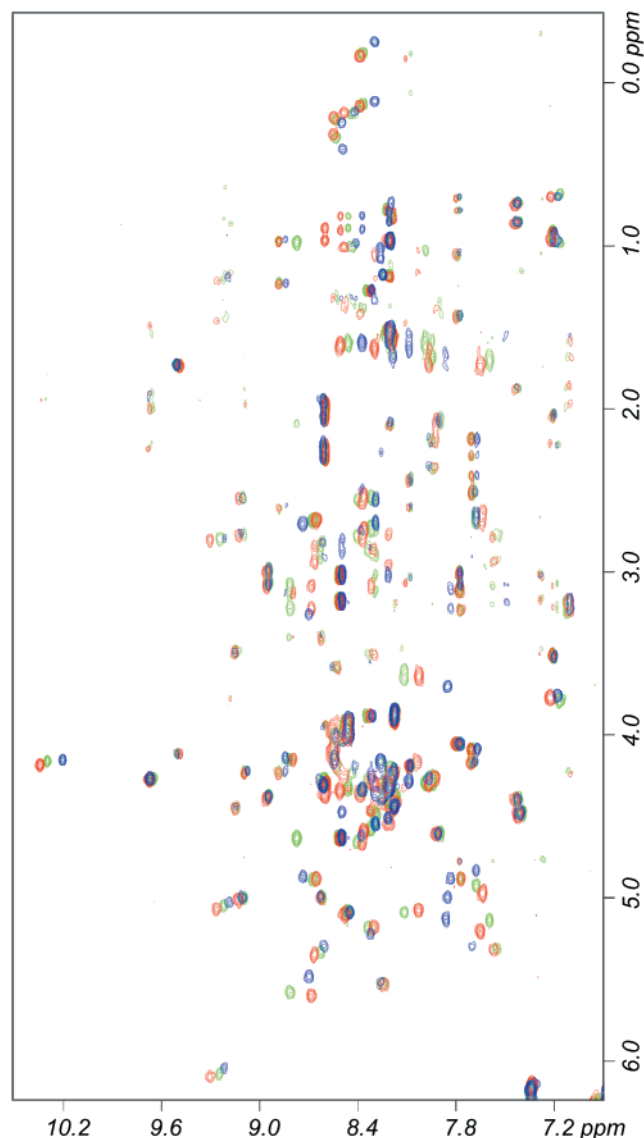


FIGURE 3: Amide to aliphatic proton region of the 2D TOCSY spectra for (1–70) $\alpha$ CONH<sub>2</sub>-S41C (red), (1–69)S41C (blue), and (1–[O]70) $\alpha$ CONH<sub>2</sub>-S41C (green) eglin c. Spectra were acquired under identical conditions as indicated under Experimental Procedures.

eglin c in the crystallographic work was due to enzymatic cleavage between residues 7 and 8 by CHYM, SCAR, or SBPN during the complex crystallization (11, 12, 14). More likely the enhanced mobility of the N-terminal residues (Thr1–Lys8), as shown from NMR dynamic analyses of free eglin c, is responsible (30, 31).

**Binding Assays with Genenase.** The fact that the three inhibitor analogues were not resistant to proteolysis by CHYM, SCAR, and most likely other serine proteinases effectively prevented us from quantitatively investigating effects of the Gly70 modifications on the binding to native enzymes. To circumvent this problem, Genenase, a catalytically incompetent mutant SBPN whose catalytic His64 was replaced by Ala (21), was used. This engineered enzyme becomes partially active in the presence of a proximal His residue at the P2 position of substrates, a phenomenon termed “substrate-assisted catalysis” (20). Since the synthetic inhibitors do not contain His residues in the binding loop region, no cleavage was expected with Genenase during binding assays, and none was observed.

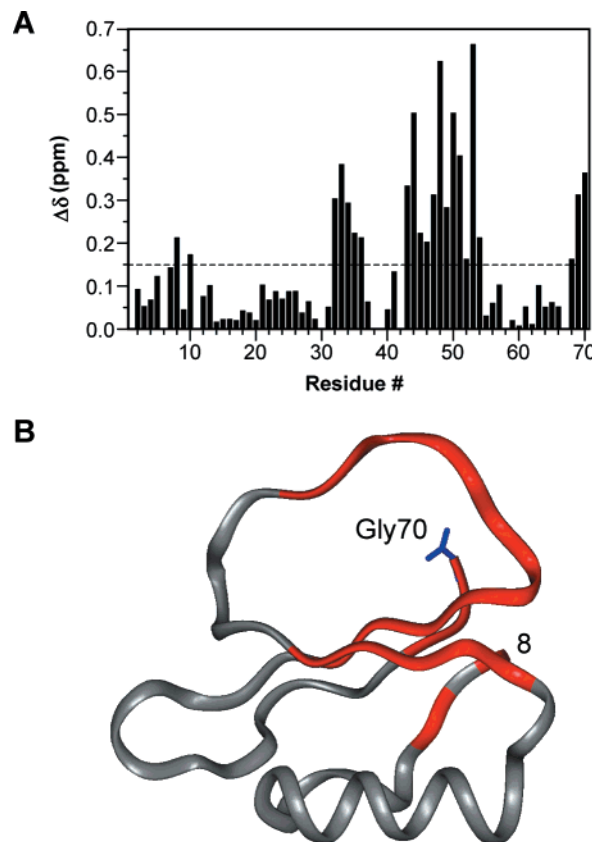


FIGURE 4: Chemical shift differences between wild-type (1–70)-S41C and (1–70) $\alpha$ CONH<sub>2</sub>-S41C eglin c. The combined change in backbone chemical shift ( $\Delta\delta$ ) is defined as  $[(\Delta\delta_{H^N})^2 + (\Delta\delta_{H^\alpha})^2]^{1/2}$  and plotted on a per residue basis (A). Those residues with  $\Delta\delta$  values greater than 0.15 ppm (horizontal line) are arbitrarily considered to be residues that exhibit a large change in chemical shift due to modification of the C-terminal carboxylate. Such residues are colored red on the ribbon diagram of the structure of wild type (8–70)S41C (4). (B) The C-terminal carboxylate group is colored blue. Panel B was generated using the program INSIGHT 97.0 (Molecular Simulations, Inc.).

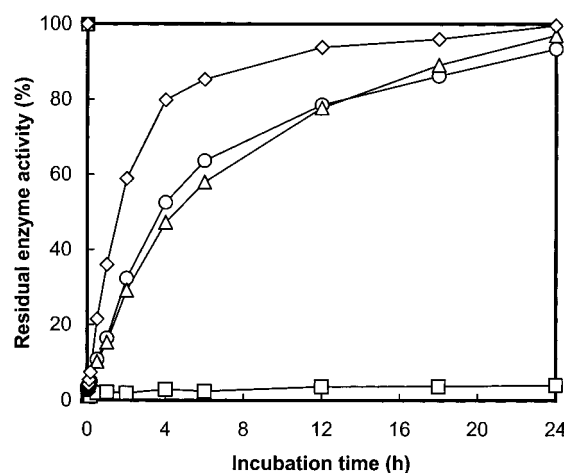


FIGURE 5: Resistance of eglin c analogues to proteolytic cleavage by CHYM. At time zero, the enzyme was stoichiometrically inhibited, showing a residual enzyme activity close to 0%. The enzyme activity fully recovered within 24 h, indicating that the analogue inhibitors were hydrolyzed and subsequently dissociated from the enzyme–inhibitor complexes formed during initial incubation. In contrast, the enzyme remained inhibited by the wild-type inhibitor. (□) Wild-type (1–70)S41C eglin c; (○) Gly70 to glycine amide; (◇) Gly70 to  $\alpha$ -hydroxyacetamide; (Δ) Gly70 truncated.

Table 2: Association Equilibrium Constants for Eglin c Analogues Interacting with Genenase

	(1-70)S41C eglin c	(1-70) $\alpha$ CONH <sub>2</sub> -S41C eglin c	(1-[O]70) $\alpha$ CONH <sub>2</sub> -S41C eglin c	(1-69)S41C eglin c
$k_{on}$ (M <sup>-1</sup> s <sup>-1</sup> )	$(3.5 \pm 1.1) \times 10^7$	$(6.0 \pm 0.3) \times 10^5$	$(4.9 \pm 0.5) \times 10^5$	$(3.0 \pm 0.4) \times 10^5$
$k_{off}$ (s <sup>-1</sup> )	$(2.5 \pm 0.2) \times 10^{-3}$	$(2.4 \pm 0.1) \times 10^{-3}$	$(4.5 \pm 0.1) \times 10^{-3}$	$(4.6 \pm 0.2) \times 10^{-3}$
$K_a$ (M <sup>-1</sup> )	$1.4 \times 10^{10}$	$2.5 \times 10^8$	$1.1 \times 10^8$	$6.6 \times 10^7$

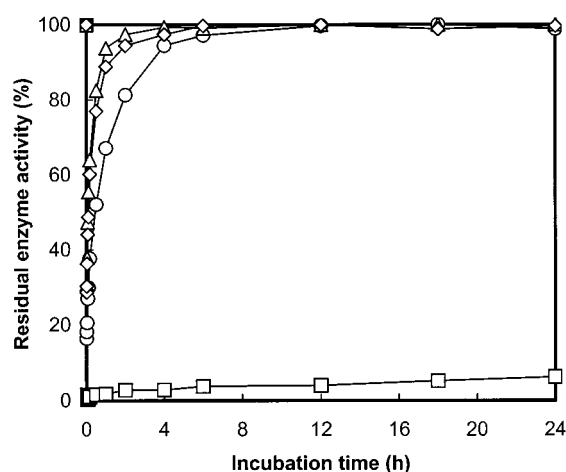


FIGURE 6: Resistance of eglin c analogues to proteolytic cleavage by SCAR. In contrast to what was observed with respect to CHYM (Figure 5), the hydrolysis of the eglin c analogues by SCAR was substantially faster. The enzyme regained its activity within 6 h, even though it remained inhibited by wild type within 24 h. (□) Wild-type (1-70)S41C eglin c; (○) Gly70 to glycine amide; (◇) Gly70 to  $\alpha$ -hydroxyacetamide; (△) Gly70 truncated.

Association equilibrium constants ( $K_a$ ) determined by BIAcore were obtained from the ratio of the kinetic association and dissociation rates ( $k_{on}$  and  $k_{off}$ ) (Table 2). Association and dissociation rate constants were determined using BIAevaluation 3.0.2 (BIAcore Inc.) software, and experimental data were fit globally to a 1:1 (Langmuir) binding model. This model fit the entire sensorgram well, suggesting a homogeneous 1:1 binding event. A low-density surface was used to avoid mass transport limitations and to reduce the effect of rebinding. The equilibrium constants were found to be independent of surface density (up to 500 RU), Genenase concentration, and the flow rate of the injection. For wild-type eglin c, the association rate constant was very fast ( $3.5 \times 10^7$  M<sup>-1</sup> s<sup>-1</sup>); thus, even at very low Genenase concentrations, mass transport limitations were a concern. However, it was found that for wild-type eglin c, a 1:1 binding model that incorporates a correction for mass transport produced  $K_a$  values that differed by only a factor of 2. Similarly, for the other eglin c analogues, there was very little change in the observed fit. This indicates that mass transport limitations do not seriously affect the binding data at the surface density and Genenase concentrations used for these measurements.

It is worth noting that the  $k_{off}$  rates for wild-type eglin c and its three analogues vary by only a factor of 2, whereas the  $k_{on}$  rates span 2 orders of magnitude. Apparently, large changes in the  $K_a$  values for the binding to Genenase are a direct consequence of changes in the association rate constants.

**Protein Stability.** Native eglin c has been shown to be a very stable small protein with a melting temperature ( $T_m$ ) of 86 °C and an enthalpy change of 71 kcal/mol at pH 7.0 (32).

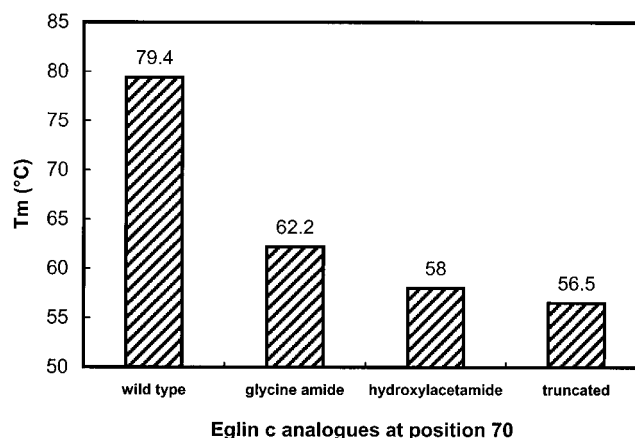


FIGURE 7: Melting temperature values ( $T_m$ ) measured by differential scanning calorimetry at pH 5.4.

Calorimetric studies of thermal denaturation at pH 5.4 on the synthetic wild type (1-70)S41C eglin c and three other analogues have clearly demonstrated that replacement or deletion of Gly70 significantly compromises the stability of the protein. As seen in Figure 7, substitution of the  $\alpha$ -carboxylate of Gly70 with an amide decreased the  $T_m$  value by 17 °C, whereas deletion of Gly70 destabilized the protein by as much as 23 °C. Since thermal denaturation of eglin c at pH 5.4 was found to cause irreversible protein precipitation, GuHCl-induced reversible denaturation was used instead to quantify the free energy changes associated with protein unfolding.

The experimental data were subjected to a six-parameter ( $a$ ,  $b$ ,  $c$ ,  $d$ ,  $m_{1/2}$ , and  $\Delta G^\circ$ ) nonlinear regression analysis (Figure 8) using the computer program Enzfitter (33). The results are listed in Table 3. For wild-type (1-70)S41C eglin c at 25 °C, pH 5.4, the native state is stabilized by 6.5 kcal/mol as compared with the denatured state.

## DISCUSSION

It has long been realized that the presence of a structured scaffold supporting the binding loop in a protein proteinase inhibitor is a necessity for potent inhibition. Efforts to design strong inhibitors based on small peptide fragments comprising the sequence of the binding loop of a protein inhibitor have not been successful. The problem is 2-fold. There is an entropy loss for an otherwise disordered small peptide to adopt a fixed canonical conformation upon binding to the enzyme. Second, due to added flexibility in the backbone of small peptides, the energy barrier to the transition state of the peptide-enzyme complex is low enough that a productive proteolytic hydrolysis often occurs even before any inhibition can be detected. Strategies such as peptide cyclization have often been used to stabilize peptide inhibitors. While cyclization does improve inhibition potency and alleviate proteolysis to a certain extent, cyclized peptides are generally weak inhibitors. This has been demonstrated in



Table 3:  $m_{1/2}$  and  $\Delta G^0$  Values As Determined by GuHCl Titration at pH 5.4

	(1-70)S41C eglin c	(1-70) $\alpha$ CONH <sub>2</sub> -S41C eglin c	(1-[O]70) $\alpha$ CONH <sub>2</sub> -S41C eglin c	(1-69)S41C eglin c
$m_{1/2}$ (M)	$3.87 \pm 0.01$	$2.56 \pm 0.02$	$2.36 \pm 0.02$	$2.18 \pm 0.02$
$\Delta G^0$ (kcal/mol)	$6.50 \pm 0.15$	$4.20 \pm 0.18$	$3.83 \pm 0.15$	$3.51 \pm 0.17$

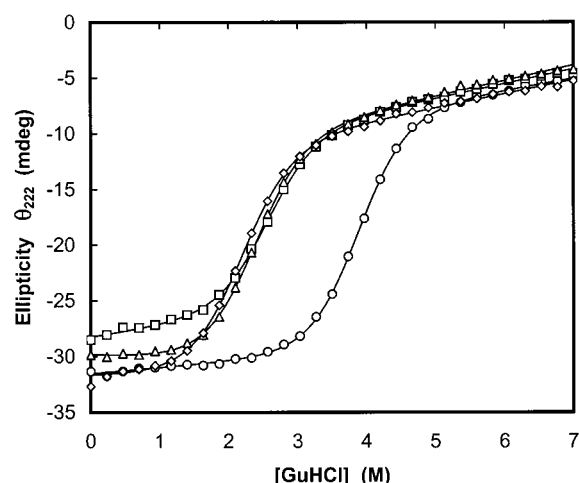


FIGURE 8: GuHCl-induced protein denaturation monitored at 222 nm by circular dichroism. The curves were subjected to a six-parameter, nonlinear regression analysis according to the equation described under Experimental Procedures. (○) Wild-type (1-70)S41C eglin c; (□) Gly70 to glycine amide; (△) Gly70 to  $\alpha$ -hydroxylacetamide; (◇) Gly70 truncated.

bovine pancreatic trypsin inhibitor (BPTI) (34) and eglin c (35), where cyclized peptide inhibitors derived from the sequence of the binding loop are approximately 6 orders of magnitude weaker than their parent inhibitors.

Based on the importance of the scaffold present in a protein inhibitor for its proper function, we were interested in the question of how will the function of a protein inhibitor be adversely affected by a destabilized binding loop and/or by a destabilized scaffold. This question was addressed by the Laskowski group in their early work on BPTI of the Kunitz family of serine proteinase inhibitors (36). They made the observation that a single peptide bond cleavage on the backside of the molecule in BPTI destabilized the protein to such an extent that the clipped BPTI was partially unfolded at room temperature. When incubated with CHYM, the clipped inhibitor was quickly hydrolyzed, which effectively led to a total loss of inhibition against the enzyme. However, the inhibition was fully restored when binding assays were conducted at 5 °C. This observation suggests that destabilization of the scaffold of a protein inhibitor such as BPTI does not automatically contribute to a decrease in inhibition, so long as the functional assay is performed under conditions that the protein folds predominantly.

Unlike BPTI and most small protein proteinase inhibitors, eglin c, which has no disulfide bridges, possesses a binding loop solely supported by electrostatic interactions. NMR studies on free eglin c indicate that the binding loop residues (Val43-Leu47) exhibit greater motion than those from the core of the domain; the data are consistent with a hinge-bending motion where the loop has internal rigidity but exhibits a rigid body motion with respect to the rest of the molecule (10, 30). Crystallographic analyses on free and bound eglin c further suggest that the wedge-shaped inhibitor undergoes a conformational adjustment upon complex for-

mation with target enzymes (9, 37). It is conceivable that destabilization in the scaffold of eglin c may seriously compromise the internal rigidity in the binding loop, leading to deleterious functional effects. The functional consequences are likely to include weakened inhibition due to an excessive entropic cost upon binding, and susceptibility of being hydrolyzed by target enzymes due to a lowered activation energy barrier to the transition state in the complexes. Results of current work clearly demonstrated that whenever the electrostatic/H-bonding network that directly stabilizes the binding loop in eglin c is weakened, deleterious functional effects follow.

The C-terminal residue of Gly70, even though a peripheral element to the main binding interface, plays an important role in supporting the binding loop of eglin c, by making a critical salt bridge with Arg51, which in turn interacts with the P1' Asp46. Compromising the integrity of the interaction between Asp46 and Arg51 invariably results in weak binding and/or temporary inhibition to target enzymes (16). It is noteworthy that disruption of the salt bridge between Arg51 and Gly70 as a result of the Gly-to-glycine amide replacement, i.e., loss of a single negative charge, caused the protein inhibitor to destabilize by 2.3 kcal/mol (17.2 °C in  $T_m$ ). NMR analysis shows surprisingly large changes in chemical shifts for residues surrounding Gly70 in the binding loop. The secondary structure and overall fold of the analogue appear similar to wild-type eglin c, but backbone coupling constants suggest the binding loop is significantly less rigid. Moreover, the resultant analogue (1-70) $\alpha$ CONH<sub>2</sub>-S41C became a temporary inhibitor against CHYM ( $t_{1/2}$  = 4 h) and SCAR ( $t_{1/2}$  = 30 min), and its binding to Genenase was weakened by 56-fold (2.4 kcal/mol) due solely to a decreased  $k_{on}$  rate. When Gly70 was replaced by  $\alpha$ -hydroxylacetamide, in addition to the breakage of the Gly70-Arg51 salt bridge, a bifurcated H-bond donated by Gly70N to Arg48O and Arg51O was also lost. The resultant protein inhibitor analogue (1-[O]70) $\alpha$ CONH<sub>2</sub>-S41C was further destabilized [ $\Delta T_m$  = -21.4 °C and  $\Delta\Delta G_{folding}$  = +2.7 kcal/mol compared with wild-type (1-70)S41C], and became more susceptible to proteolytic cleavage by CHYM ( $t_{1/2}$  = 2 h) and SCAR ( $t_{1/2}$  = 5 min). Expectedly, (1-[O]70) $\alpha$ CONH<sub>2</sub>-S41C was a weaker Genenase binder than (1-70) $\alpha$ CONH<sub>2</sub>-S41C ( $K_a$  decreased by 130-fold, i.e., 2.9 kcal/mol).

Deletion of Gly70 led only to a marginal further destabilization of the protein with respect to (1-[O]70) $\alpha$ CONH<sub>2</sub>-S41C ( $\Delta T_m$  = -22.9 °C and  $\Delta\Delta G_{folding}$  = +3.0 kcal/mol relative to wild type). While the binding to Genenase was weakened by 210-fold (3.2 kcal/mol), the resistance of (1-69)S41C eglin c to proteolytic cleavage remained unchanged with SCAR ( $t_{1/2}$  = 5 min) and actually improved with respect to CHYM ( $t_{1/2}$  = 4 h). This observation seems to indicate that resistance of the inhibitor analogues to proteolytic cleavage by their target enzymes is not necessarily correlated to the stability of the proteins. With the deletion of Gly70, the contacts made previously by the main-chain atoms of



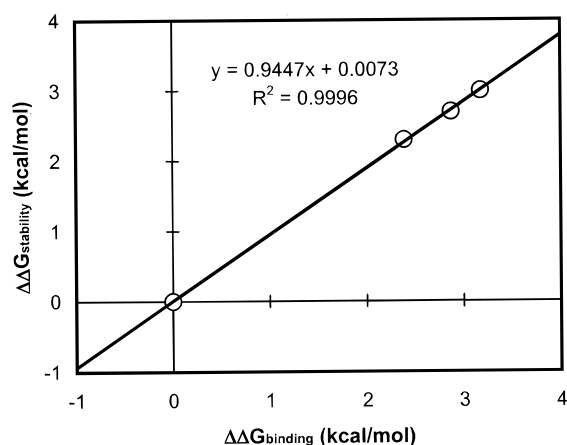


FIGURE 9: Correlation between the relative free energy changes ( $\Delta\Delta G_{\text{stability}}$ ) associated with protein destabilization and the  $\Delta\Delta G_{\text{binding}}$  values derived from a decreased binding to Genenase. The data were subjected to a linear regression analysis, yielding a slope of 0.94 and an  $r^2$  value of 0.9996.

the residue with Arg51 and Arg48 were all lost. It is plausible that a newly generated  $\alpha$ -carboxyl group of Val69 might provide additional interactions, favorable or unfavorable, with atoms or residues in proximity. Nevertheless, it is not clear what effects these potential contacts involving the carboxylate of Val69 in (1–69)S41C eglin c exerted on protein stability and resistance to proteolysis.

It is worth noting that for all three analogue inhibitors the adverse functional effects reflected by a weakened binding to Genenase are directly correlated to protein destabilization as measured by GuHCl-induced CD titration. As illustrated in Figure 9, the plot of the relative free energy changes associated with protein destabilization,  $\Delta\Delta G_{\text{stability}}$ , versus the  $\Delta\Delta G_{\text{binding}}$  values derived from the binding assays gives rise to a slope of 0.94 and an  $R^2$  value of 0.9996. It is conceivable that the weakening of the binding to Genenase was, at least in part, due to an entropic penalty associated with an increase in flexibility of the binding loop in eglin c, as a result of lost electrostatic interactions involving Gly70. The rationale is in fact consistent with the experimental observation that large changes in the  $K_a$  values for the binding to Genenase are caused predominantly by changes in the  $k_{\text{on}}$  rates. The data suggest that the added flexibility in the binding loop, caused by removal of a stabilizing H-bonding element, plays a key role in turning the protein inhibitor into a substrate, which concurs with earlier observations made by Heinz et al. (16). It should be pointed out that destabilization of the scaffold of eglin c alone is not sufficient to exert a pronounced change in binding. This has been demonstrated in a separate experiment where the buried residue Phe25 in the hydrophobic core of eglin c was replaced by norvaline. The Phe-to-Nva substitution destabilized the protein by 3.2 kcal/mol, but caused no noticeable change in its binding to Genenase as compared with wild type (unpublished results).

In conclusion, we have applied native chemical ligation to synthesize wild-type eglin c and three analogues with good yields and high efficiency. The chemical protein engineering approach, combined with “unnatural” backbone substitutions, has provided quantitative structural and functional information concerning the role of the charge–charge interactions involving Gly70 and Arg51 in eglin c. A strong correlation

between protein stability and protein binding has been demonstrated in this study. While this might not be applicable in every protein–protein interacting system, the correlation has some general implications with respect to protein recognition, particularly when protein stabilization or destabilization directly involves the binding loop studied. In addition, our results reported here are likely to shed light on how protein recognition, in general, can be regulated through modulating rigidity and/or flexibility of the binding loop involved.

## ACKNOWLEDGMENT

We are grateful to Dr. Michael Laskowski, Jr., for sharing unpublished results with us. We thank Dr. Wayne Fairbrother and Dr. Nicholas Skelton for many helpful discussions. We also thank Dr. Michael Weiss for providing us with access to his CD spectrometer, and Ms. Wenhua Jia for technical assistance on the GuHCl titration experiment.

## REFERENCES

1. Laskowski, M., Jr., and Kato, I. (1980) *Annu. Rev. Biochem.* 49, 593–626.
2. Read, R. J., and James, M. N. G. (1986) in *Proteinase Inhibitors* (Barrett, A. J., and Salvesen, G., Eds.) pp 301–336, Elsevier, Amsterdam.
3. Bode, W., and Huber, R. (1992) *Eur. J. Biochem.* 204, 433–451.
4. Lu, W., Randal, M., Kossiakoff, A., and Kent, S. B. H. (1999) *Chem. Biol.* 6, 419–427.
5. Seemuller, U., Meier, M., Ohlsson, K., Muller, H. P., and Fritz, H. (1977) *Hoppe-Seyler's Z. Physiol. Chem.* 358, 1105–1107.
6. Seemuller, U., Eulitz, M., Fritz, H., and Strobl, A. (1980) *Hoppe-Seyler's Z. Physiol. Chem.* 361, 1841–1846.
7. Braun, N. J., Bodmer, J. L., Virca, G. D., Metz-Virca, G., Maschler, R., Bieth, J. G., and Schnebli, H. P. (1987) *Biol. Chem. Hoppe-Seyler* 368, 299–308.
8. Ascenzi, P., Amiconi, G., Menegatti, E., Guarneri, M., Bolognesi, M., and Schnebli, H. P. (1988) *J. Enzyme Inhib.* 2, 167–172.
9. Hippler, K., Priestle, J. P., Rahuel, J., and Grutter, M. G. (1992) *FEBS Lett.* 309, 139–145.
10. Hyberts, S. G., Goldberg, M. S., Havel, T. F., and Wagner, G. (1992) *Protein Sci.* 1, 736–751.
11. Frigerio, F., Coda, A., Pugliese, L., Lionetti, C., Menegatti, E., Amiconi, G., Schnebli, H. P., Ascenzi, P., and Bolognesi, M. (1992) *J. Mol. Biol.* 225, 107–123.
12. Bode, W., Papamokos, E., and Musil, D. (1987) *Eur. J. Biochem.* 166, 673–692.
13. McPhalen, C. A., and James, M. N. (1988) *Biochemistry* 27, 6582–6598.
14. Heinz, D. W., Priestle, J. P., Rahuel, J., Wilson, K. S., and Grutter, M. G. (1991) *J. Mol. Biol.* 217, 353–371.
15. Gros, P., Betzel, C., Dauter, Z., Wilson, K. S., and Hol, W. G. (1989) *J. Mol. Biol.* 210, 347–367.
16. Heinz, D. W., Hyberts, S. G., Peng, J. W., Priestle, J. P., Wagner, G., and Grutter, M. G. (1992) *Biochemistry* 31, 8755–8766.
17. Dawson, P. E., Muir, T. W., Clark-Lewis, I., and Kent, S. B. (1994) *Science* 266, 776–779.
18. Mitchell, A. R., Kent, S. B. H., Engelhard, M., and Merrifield, R. B. (1978) *J. Org. Chem.* 43, 2845–2852.
19. Canne, L. E., Walker, S. M., and Kent, S. B. H. (1995) *Tetrahedron Lett.* 36, 1217–1220.
20. Carter, P., and Wells, J. A. (1987) *Science* 237, 394–399.
21. Carter, P., Nilsson, B., Burnier, J. P., Burdick, D., and Wells, J. A. (1989) *Proteins: Struct., Funct., Genet.* 6, 240–248.

22. Schnölzer, M., Alewood, P., Jones, A., Alewood, D., and Kent, S. B. (1992) *Int. J. Pept. Protein Res.* 40, 180–193.
23. Lu, W., Qasim, M. A., Laskowski, M., Jr., and Kent, S. B. H. (1997) *Biochemistry* 36, 673–679.
24. Dawson, P. E., Churchill, M. J., Ghadiri, M. R., and Kent, S. B. H. (1997) *J. Am. Chem. Soc.* 119, 4325–4329.
25. Ryan, D. S., and Feeney, R. E. (1975) *J. Biol. Chem.* 250, 843–847.
26. Cavanagh, J., Fairbrother, W., Palmer, A., III, and Skelton, N. (1995) *Protein NMR spectroscopy: principles and practice*, Academic Press, San Diego.
27. Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1986) *Numerical Recipes. The Art of Scientific Computing*, Cambridge University Press, Cambridge.
28. Hyberts, S. G., and Wagner, G. (1990) *Biochemistry* 29, 1465–1474.
29. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, J. Wiley and Sons, New York.
30. Peng, J. W., and Wagner, G. (1992) *Biochemistry* 31, 8571–8586.
31. Peng, J. W., and Wagner, G. (1995) *Biochemistry* 34, 16733–16752.
32. Bae, S. J., and Sturtevant, J. M. (1995) *Biophys. Chem.* 55, 247–252.
33. Leatherbarrow, R. J. (1987) *Enzfitter*, Biosoft, Cambridge, U.K.
34. Tan, N. H., and Kaiser, E. T. (1976) *J. Org. Chem.* 41, 2787–2793.
35. Shao, Y., Lu, W., and Kent, S. B. H. (1998) *Tetrahedron Lett.* 39, 3911–3914.
36. Kelly, C. (1989) in Ph.D. Thesis, Department of Chemistry, Purdue University, West Lafayette, IN.
37. Hipler, K., Priestle, J. P., Rahuel, J., and Grutter, M. G. (1996) *Adv. Exp. Med. Biol.* 379, 43–47.

BI992292Q