The pepsin residue glycine-76 contributes to active-site loop flexibility and participates in catalysis

Monika OKONIEWSKA, Takuji TANAKA and Rickey Y. YADA[1]
Department of Food Science, University of Guelph, Guelph, ON, N1G 2W1, Canada

Glycine residues are known to contribute to conformational flexibility of polypeptide chains, and have been found to contribute to flexibility of some loops associated with enzymic catalysis. A comparison of porcine pepsin in zymogen, mature and inhibited forms revealed that a loop (a flap), consisting of residues 71–80, located near the active site changed its position upon substrate binding. The loop residue, glycine-76, has been implicated in the catalytic process and thought to participate in a hydrogen-bond network aligning the substrate. This study investigated the role of glycine-76 using site-directed mutagenesis. Three mutants, G76A, G76V and G76S, were constructed to increase conformational restriction of a polypeptide chain. In addition, the serine mutant introduced a hydrogen-bonding potential at position 76 similar to that observed in human renin. All the mutants, regardless of amino acid size and polarity, had lower catalytic efficiency and activated more slowly than the wild-type enzyme. The slower activation process was associated directly with altered proteolytic activity. Consequently, it was proposed that a proteolytic cleavage represents a limiting step of the activation process. Lower catalytic efficiency of the mutants was explained as a decrease in the flap flexibility and, therefore, a different pattern of hydrogen bonds responsible for substrate alignment and flap conformation. The results demonstrated that flap flexibility is essential for efficient catalytic and activation processes.

Key words: aspartic proteinase, hinge, hydrogen bond, pepsinogen activation, substrate alignment.

INTRODUCTION

Loops represent a relatively new category of secondary structures. They reverse the direction of a polypeptide chain and are usually situated at the protein surface. Unlike α-helices and β-sheets, loop structures have no regular patterns of dihedral angles and hydrogen bonds [1,2]. Loop residues that are involved in catalysis may be more highly conserved than amino acids in surface loops whose role is purely structural. This sequence conservation can involve amino acids that interact directly with a substrate as well as amino acids that help stabilize various conformations that the loop may adopt during catalysis [3].

The hydrogen bonds to main-chain polar atoms play an important role in maintaining medium-sized loop structures. They can be formed from the loop atoms to main-chain or side-chain atoms or ligands [4]. It was observed in lactate dehydrogenase that open and closed loop conformations were facilitated by the same pattern of hydrogen bonds made among loop atoms [5]. In contrast, hydrogen bonds in triosephosphate isomerase reorganized during catalysis [6].

Loop structures have been identified in many enzymes where they have been implicated in protein function, stability and possibly in protein folding. In aspartic proteinases, a loop, commonly known as a flap, is found to extend over the active site and has been implicated in the catalytic process [7,8].

Three-dimensional structures at high resolutions have been identified for viral, fungal and mammalian aspartic proteinases. A comparison of the molecules reveals that their tertiary structures have been conserved during evolution. However, the enzymes differ in their primary structures, sequence homologies and substrate preferences. Therefore, aspartic proteinases represent an interesting group in which to study structure and function relationships [9]. Porcine pepsin is a particularly good model for these studies since its structures in zymogen, mature and inhibited forms have been determined [8,10,11].

Under physiological conditions porcine pepsin is secreted in the stomach as a zymogen, pepsinogen. In pepsinogen, a 44-residue peptide, referred to as a prosegment, blocks the entrance to the active site and is held in place by electrostatic interactions with the pepsin molecule. The acidic environment triggers the activation process. At pH values lower than 5.0, the reduction of electrostatic interactions between the pepsin molecule and the prosegment initiates a series of conformational changes and proteolytic cleavage of the prosegment or its fragments from the N-terminus of pepsin to yield a mature enzyme [10,12]. The final shape of substrate-binding subsites is gained during the activation process and residues lining the binding cleft assume positions that enable hydrogen bonding to a substrate. The hydrogen bonds align the substrate in the position most favourable for the catalysis. The crystal structures of pepsin in free and inhibited forms indicated that three flap residues, Y75, G76 and T77 (one-letter amino acid notation), contribute directly to a particular subsite specificity by hydrogen bonding to substrate residues [10,11]. Y75 was found to participate in substrate alignment; however, its influence was less pronounced than in chymosin and *Rhizomucor* pepsin, which had higher substrate specificities [13]. It was suggested that Y75 anchored the flap in pepsin. In chymosin, the residue was more mobile and therefore able to adopt more conformations that could contribute to high enzyme specificity [14]. An investigation of the role of T77 in pepsin catalysis revealed that the residue provides a catalytically essential hydrogen bond to a substrate. It was concluded that the hydrogen bond in position 77 participates in the substrate alignment and in this way is indirectly responsible for the proper geometry in the transition state [15].

Abbreviation used: Trx-PG, thioredoxin–pepsinogen fusion protein.

[1] To whom correspondence should be addressed (e-mail ryada@uoguelph.ca).

Table 1 Flap sequences of selected aspartic proteinases

Fungal pepsins are represented by endothiapepsin, penicillopepsin and rhizopuspepsin. The numbering is according to porcine pepsin [36,44].

		Loop residue				
Enzyme	Position ...	75	76	77	78	79
Porcine pepsin		Y	G	T	G	S
Human pepsin		Y	G	T	G	S
Bovine chymosin		Y	G	T	G	S
Porcine chymosin		Y	G	T	G	S
Human cathepsin E		Y	G	T	G	S
Human cathepsin D		Y	G	S	G	S
Human gastricsin		Y	G	S	G	S
Human renin		Y	S	T	G	T
Fungal pepsins		Y	G	D	G	S

An examination of G76 atomic contacts in inhibited pepsin structure [11] indicates that main-chain hydrogen of G76 participates in hydrogen bonding to a substrate. It has also been suggested that glycine 76 is involved in the hydrogen bond stabilizing the transition state [16,17]. A comparison of eukaryotic aspartic proteinases (Table 1) in the flap region demonstrates that G76 and G78 are conserved in all enzymes with the exception of renin. However, G78 is in a location too distant to interact directly with a substrate. At the same time, the greatest change in the pepsin flap structure is induced in residues G76–G78 located at the tip of the flap. C_α atoms of G76–T77–G78 move by 4.8 Å (1 Å = 0.1 nm) on average [14]. The observations suggest that G76 participates in the catalysis by direct interaction with a substrate, a transition state and/or by providing essential flap flexibility.

Glycine residues were implicated in providing conformational freedom to the polypeptide chain [18]. They have been shown to act as points of flexibility for hinge-type motions in the active-site loop of glutathione synthetase and tryptophan synthase. Some flexible loop structures identified in other enzymes that are known to change their positions during catalysis have glycines at their terminal positions, e.g. adenylate kinase, lactate dehydrogenase and Rubisco. The residues probably represent flexibility points essential for the flap motions [19–23].

The present study investigated the role of G76 in pepsin's catalytic mechanism by substituting the residue in position 76 with alanine, valine and serine. These amino acids differ in their van der Waals volumes (G, 48 Å^3; A, 67 Å^3; V, 105 Å^3; S, 73 Å^3), accessible surface areas (G, 85 Å^2; A, 113 Å^2; V, 160 Å^2; S, 122 Å^2), polarities and allowable energy levels on Ramachandran plots for individual amino acids [18].

EXPERIMENTAL PROCEDURES

Materials

Restriction enzymes with corresponding buffers were purchased from New England Biolabs (Mississauga, ON, Canada) and Boehringer Mannheim Canada (Laval, QC, Canada). All the chromatography resins were purchased from Pharmacia Biotech (Uppsala, Sweden). Protein assay reagents, SDS/PAGE reagents and equipment were purchased from Bio-Rad (Hercules, CA, U.S.A.). Substrate I, Lys-Pro-Ala-Glu-Phe-Phe(NO_2)$_p$-Ala-Leu, was synthesized at the Institute for Molecular Biology and Biotechnology, McMaster University, Hamilton, ON, Canada. Substrate II, Leu-Ser-Phe(NO_2)$_p$-Nle-Ala-Leu, was purchased from Sigma (St. Louis, MO, U.S.A.). Additional materials were obtained either from Fisher Scientific (Mississauga, ON, Canada) or Sigma.

Energy minimization and molecular dynamics

Calculations were performed using Discover 2.9.5 (Biosym Technologies, San Diego, CA, U.S.A.) on an IBM Risc System/6000 computer. The co-ordinates of porcine pepsin (PDB entry 4PEP by Sielecki et al. [8]) from the Protein Data Bank were used as initial data for the calculations. The residues of the wild-type protein were replaced with corresponding mutant residues and energy minimization was performed in a consistent valence force field by the steepest-descents method for 2000 iterations with 24-Å cut-off and a maximum derivative of 0.001 kcal (1 cal ≡ 4.184 J)/mol. The wild-type was minimized as the reference. Molecular-dynamics simulations were performed on the energy-minimized molecule models for 11 fs at 300 K. Co-ordinates were collected every 20 fs. Flap flexibility was determined based on the mean square deviation of the α-carbon atom displacements during the dynamics calculations.

Plasmid construction and cloning

Mutations of pepsinogen were performed according to the method of Kunkel [24]. Site-directed mutagenesis was performed using bacterial strains of *Escherichia coli* CJ236 (*dut*$^-$ *ung*$^-$), JM109 (wild-type) and bacteriophage M13mp19 RF (replicative form) DNA. The primers that introduced desired mutations and new restriction enzyme sites were for G76A, CCATCA-CA*TATGC*CACCGGTAG (*Nde*1); for G76V, CCATCA-CA*TATGT*CACCGGTAG (*Nde*1); and for G76S, CACC-TATA*GT*ACT*GGTAGCATG (*Sca*1). Base replacements are indicated by asterisks, and restriction-enzyme sites are underlined. The mutations were confirmed on agarose gel after digestion with appropriate restriction enzymes and DNA sequencing by the automated M13-dideoxy method. The mutated pepsinogen genes were expressed as fusion proteins with thioredoxin (fusion protein, Trx-PG) in *E. coli* GI724 using pTFP1000 plasmid [25] by introducing a short fragment containing the mutation in place of the wild-type counterpart. Fusion-protein production was verified using Western blotting with rabbit anti-pepsinogen polyclonal antibodies and goat anti-rabbit-IgG–alkaline phosphatase conjugate. The blots were visualized using Nitro Blue Tetrazolium (NBT) and 5-bromo-4-chloroindo-3-yl phosphate toluidine (BCIP) solutions in 0.1 M Tris/50 mM $MgCl_2$/0.1 M NaCl, pH 9.5.

Protein expression and purification

Transformed cells were pre-cultured in RM media [0.8 % (w/v) casamino acids, 1 % (v/v) glycerol, 1 × M9 salt and 0.1 mM $MgCl_2$] at 30 °C for 18 h, and grown in Luria–Bertani media (1 % Bacto Tryptone, 0.5 % Bacto Yeast Extract, 1 % NaCl and 0.15 μg/ml ampicillin, pH 7.5) at 30 °C for 10 h. Cells were collected by centrifugation at 15000 ***g*** for 10 min, resuspended in 20 mM Tris/HCl, pH 8.0, and disrupted by sonication. Cell lysate was collected by centrifugation at 20000 ***g*** for 20 min. The purification procedure was performed using three ion-exchange columns. Nucleic acids were precipitated from the cell lysate with protamine sulphate (10 % of protein concentration) and removed by centrifugation (20000 ***g***, 20 min). Urea was added to the supernatant at 2 M. The supernatant was applied at a flow rate of 1 ml/min on a CM-Sepharose column (2.5 × 22 cm) connected directly to a DEAE-Sephacel column (2.5 × 30 cm) equilibrated with 0.1 M NaCl in 20 mM Tris/HCl/2 M urea (pH 8.0) buffer.

Urea was necessary to increase resolution of the fusion protein. Without urea, the fusion protein eluted as a very broad band spreading over approx 0.15–0.35 M. The two columns were disconnected after washing with 0.1 M NaCl in 20 mM Tris/HCl/2 M urea, pH 8.0, and the second column was subsequently washed with a stepwise gradient of NaCl (0.15, 0.2, 0.35 M) in 20 mM Tris/HCl/2 M urea, pH 8.0, at 1 ml/min and 4 °C. The purest fractions of the fusion protein, as detected by SDS/PAGE, were eluted with 0.35 M NaCl, pooled, and diluted three to four times to decrease the ionic strength. This preparation was applied to a DEAE-Sepharose column (2.5 × 25 cm) equilibrated with 50 mM phosphate buffer, pH 6.5, and washed with a stepwise gradient of NaCl (0.1, 0.2, 0.25, 0.3 M), at 1 ml/min and 4 °C. The Trx-PG fractions eluted with 0.3 M NaCl were subjected to further concentration on an Amicon concentrator model 8200 using a 10-kDa cut-off refined cellulose membrane (Amicon, Beverly, MA, U.S.A.) to a final volume of approx. 10 ml. The sample was then dialysed against 20 mM Tris/HCl, pH 8.0, filtered through a 0.22-μm nitrocellulose filter, and stored at 4 °C.

Protein concentration

Protein concentration was determined using two methods: (i) absorbance at 280 nm using an absorption coefficient of 27 500 l/mol·cm [26]; and (ii) Bio-Rad D_c protein assay based on the Lowry method. Standard curves were generated using BSA.

SDS/PAGE and Western-blot analysis

SDS/PAGE was performed using a Bio-Rad Mini-Protean cell on a 12% acrylamide gel at 30–40 mA. Proteins were stained with Coomassie Brilliant Blue R-250. Western blots were run on a Bio-Rad Mini-Trans Blot Cell at 100 V. Proteins were transferred to a 0.2 μm PVDF membrane.

Activation of Trx-PG and purification of pepsin

Fusion protein was incubated with 0.09 M HCl for 1 h at room temperature. Activated pepsin samples were then purified from thioredoxin-activation fragments by gel-filtration chromatography on Sephadex G50 [25]. Two buffers were used: 20 mM sodium acetate, pH 5.3, for kinetic assays and 10 mM sodium acetate, pH 5.3, for far-UV CD spectroscopy analysis. Active samples were identified by milk-clotting assay [27] and only the first active fraction that eluted in the void volume was used for further pepsin analysis. For near-UV spectroscopy analysis, fusion proteins were acidified as described above, neutralized with 1 M sodium acetate buffer, pH 5.3, and then pepsins were purified from activation fragments in Centricon filtration units with a 30-kDa molecular-mass cut-off (Amicon) at 5000 ***g***. The buffer was exchanged to 10 mM sodium acetate, pH 5.3, during filtration in the Centricon units.

CD spectroscopy

Far-UV CD spectra were generated using 800 μl of 0.1 mg/ml purified enzyme solutions with a Jasco 600 spectropolarimeter (Japan Spectroscopic Co., Tokyo, Japan). Pepsin samples in a 0.1-cm pathlength cuvette were scanned six times from 190 to 250 nm at room temperature under continuous nitrogen flush. The buffer blank and enzyme solutions were degassed prior to the analyses. Secondary structures were determined using the Jasco Protein Secondary Structure Estimation Program based on the method of Chang et al. [28].

Near-UV CD spectra were obtained from 1 mg/ml degassed pepsin solutions in 10 mM sodium acetate buffer, pH 5.3. Pepsin samples of 1 ml were placed in a 1-cm pathlength cuvette and scanned eight times from 320 to 240 nm at room temperature under continuous nitrogen flush. Tanaka and Yada [25] showed that properties of the wild-type and native pepsin(ogens) were the same (including N-terminal sequences). In addition, preliminary studies demonstrated that structures of the wild-type and native pepsins were the same (results not shown). Therefore, comparison of structural results for G76 mutants was performed relative to the wild-type pepsin obtained from the activation of the fusion protein in the same way as all the mutants.

Kinetic analyses

Kinetic measurements were performed on a DU 640 spectrophotometer (Beckman Instruments, Fullerton, CA, U.S.A.) with a temperature-control unit. Two synthetic substrates, as defined in the Materials section, were used. For substrate I, a change in absorbance was measured at 300 nm, 37 °C, in 100 mM sodium citrate buffer, pH 2.1 [29]. The substrate range was 2–200 μM. For substrate II, a change in absorbance was measured at 310 nm, 37 °C, in 100 mM sodium citrate buffer, pH 3.95. The second substrate range was 2–300 μM. A minimum of six concentrations for each substrate range was used to determine initial rates. All the kinetic experiments were performed within the first 3 days after the final enzyme concentration. Enzymes were activated as described in the section on activation of Trx-PG and purification of pepsin (see above). The initial slopes of progress curves were measured to give ΔA/min. The non-linear least-squares method was used to fit data and to calculate the kinetic constants K_m and k_{cat} [30]. The results were calculated based on a minimum of two measurements for each enzyme.

Trx-PG activation kinetics, and pH dependency of activation were performed as described by Okoniewska et al. [15]. Starting fusion-protein concentrations were 0.2 mg/ml for Trx-PG activation kinetics and 0.3 mg/ml for pH dependency of activation studies. In Trx-PG activation kinetics, samples were taken as follows: for the wild-type protein, every 15 s at pH 1.1 and 2.0, and every minute at pH 3.0; for G76A, every 30 s at pH 1.1 and 2.0, and every 2 min at pH 3.0; and for G76V and G76S every 2 min at pH 1.1 and 2.0, and every 5 min at pH 3.0. The activation reactions were conducted at 14 °C. Activation rate constants were calculated based on the hydrolysis of synthetic substrate I, using integrated rate law. pH dependency of activation was determined using milk-clotting assay [27]. The activity of samples was tested over ten pH values starting at pH 1.1, then 1.5, and in 0.5-pH-unit increments to 5.5, after 1 h of activation. The results were calculated based on three measurements for each enzyme.

RESULTS AND DISCUSSION

Molecular-dynamics calculations (Figure 1) indicated that all the mutants in position 76 resulted in altered flexibility of the flap. G76A mutant reduced the flexibility (increased the rigidity) in position 75 less than G76S and G76V mutants. The G76S mutant reduced flexibility of the flap in position 76 slightly less than the G76A mutant. The G76V mutant resulted in the greatest reduction in flexibility in position 76 and overall flap flexibility. At the same time, the G76V mutation resulted in increased flexibility of T77 and G78. This increased flexibility could have resulted from the inability of T77 to participate in certain hydrogen-bond interactions due to increased rigidity in the neighbouring position 76. Substitution of glycine 76 with serine increased flap rigidity; however, less so than in the G76V mutant. The G76A mutation resulted in increased rigidity at the tip of the flap and greater overall flap flexibility. Therefore,

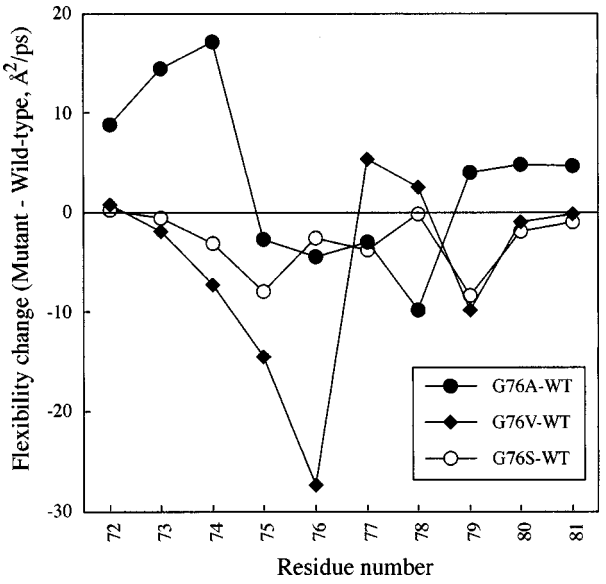

Figure 1 Evaluation of flap flexibility in pepsin G76 mutants and the wild-type (WT) using molecular-dynamics simulations

Molecular-dynamics simulations were performed on the energy-minimized molecule models. Flap flexibility was determined based on the mean square deviation of the α-carbon atom displacements during the dynamics calculations.

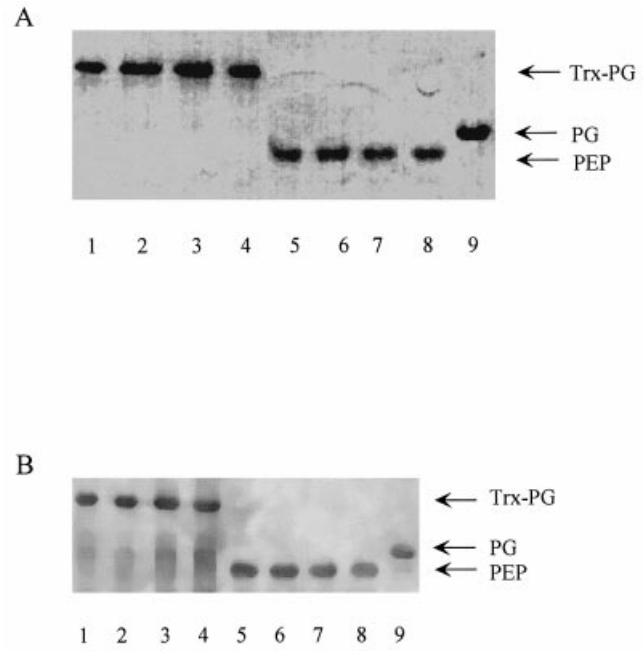

Figure 2 Electrophoresis and Western blot of mutant fusion proteins and pepsins

(**A**) SDS/PAGE (12% gel) stained with Coomassie Brilliant Blue R-250. (**B**) Western blot of the same gel. Fusion protein (Trx-PG) lanes: 1, wild-type; 2, G76A; 3, G76V; 4, G76S. Active pepsin (PEP) lanes: 5, wild-type; 6, G76A; 7, G76V; 8, G76S; lane 9, commercial pepsinogen (PG). In comparison with the results from SDS/PAGE (**A**), the Western blot (**B**) indicated that there were additional diffuse bands with mobilities similar to pepsinogen in the wild-type, G76A, G76V and G76S fusion proteins. These bands were artifacts resulting from storage. Freshly prepared protein samples did not show any of these bands. However, to run all samples on a single gel required that certain samples be stored for some months. In addition it was estimated that the amount of these faint bands was less than 0.1% of the fusion protein, and that no difference in activity should be observed.

the overall increase in flexibility could compensate for the more rigid tip of the flap. The molecular-dynamics calculations indicated that the greatest reduction of flexibility in position 76 resulted in the highest increase in flexibility of T77. It was determined previously [15] that the hydrogen bond from T77 participates in substrate alignment and indirectly positions the transition state. Consequently, the G76V mutation would be expected to have the greatest effect on enzyme kinetic parameters.

All the enzymes had similar elution profiles (results not shown) and were purified to homogeneity, as indicated by single bands on SDS/PAGE (Figure 2A). The identity of all enzyme preparations in zymogen and active forms was determined by immunoblotting (Figure 2B).

The secondary structure of all the mutants was analysed by far-UV CD spectroscopy. The results are presented in Figure 3. All the mutants had spectra overlapping with the wild-type, and their calculated secondary-structure contents were virtually the same as the wild-type (Table 2), indicating that no global secondary-structure change occurred as a result of the mutations.

Tertiary-structure analysis by near-UV CD spectroscopy (Figure 4) revealed that there were differences between mutants and the wild-type protein. The G76A mutant had a profile very similar to the wild-type. However, profiles for G76S and G76V mutants were different than the wild-type, with the greatest difference in spectral pattern seen in the G76V mutant. The near-UV CD profile differences reflect changes in the environment of aromatic residues. In the pepsin molecule, G76 is in the direct vicinity of Y75. In addition, Y75 is close to other aromatic residues in pepsin's active site, e.g. W39, F111 and F117 [8,31]. The change in environment of the active-site residues could contribute to observed differences in near-UV CD profiles. Therefore, the results indicated that mutations G76S and G76V changed the environment of aromatic residues in the vicinity of position 76. All the CD results combined suggested that the mutations did not induce dramatic conformational change and that the observed differences were caused by changes in the local conformation of the loop region (residues 71–80).

The mutants and wild-type protein were activated in ten different pH conditions from 1.1 to 5.5 over the same time of acidification to determine the pH dependence of the activation process. The amounts of pepsins formed were determined by milk-clotting assay. The activation profiles were plotted as relative amount of pepsin formed during the acidification time against pH of activation (Figure 5). The amounts of pepsin were the highest at the most acidic pH, demonstrating that all the enzymes were activated fastest in a highly acidic environment. Minimum milk-clotting activity was still observed at pH 4.5. The enzymes did not activate at pH 5.0 and 5.5. The activation profiles were similar for all the mutants and the wild-type protein, demonstrating that all the proteins reacted similarly to a change in pH environment. The mutations of G76 did not introduce any additional ionizable groups to the flap. However, the flap in pepsinogen is in a different position to that in pepsin and interacts with charged residues in the zymogen. The maximum change of the flap position is induced by residues G2-D3-E4-P5-L6-E7-N8 that insert between the flap and the E107–S110 strand in the zymogen [10,14]. In addition, Y9 has been proposed to play a critical role in the stabilization of pepsinogen [32]. If any structural changes occurred in the zymogen they should have been reflected in the interaction pattern between the flap and the G2–Y9 strand and, therefore, in the pH dependency of the activation process. The results showed that there was no change in pH dependency of activation despite the mutations.

All the mutants had similar pH activation profiles and purification chromatogram profiles (results not shown), and

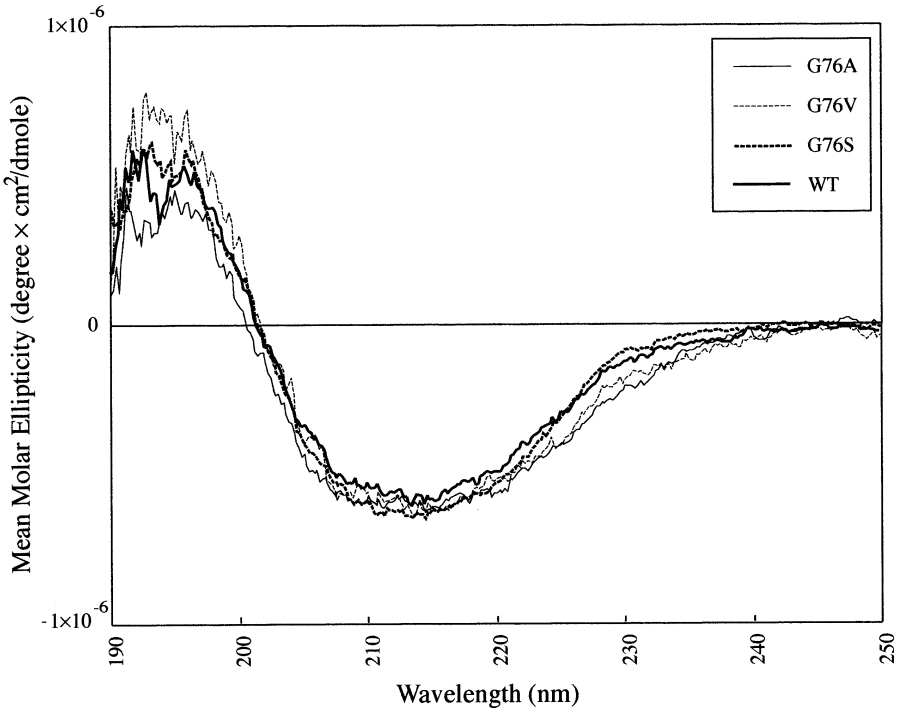

Figure 3 Far-UV CD spectra of the wild-type and mutant pepsins

Experiments were carried in 10 mM sodium acetate buffer, pH 5.3. Spectra were generated using 800 μl of 0.1 mg/ml purified enzyme solutions in a 0.1-cm pathlength cuvette after six scans from 190 to 250 nm at room temperature under continuous nitrogen flush.

Table 2 Secondary-structure content of wild-type pepsin and G76 mutants determined by far-UV CD

RMS, root mean square.

	Secondary-structure content (%)				
Enzyme	α-Helix	β-Sheet	β-Turn	Random coil	RMS
Wild-type	3.9	56.8	12.5	26.8	3.91
G76A	2.5	63.2	12.2	22.1	4.24
G76V	4.1	57.5	11.3	27.0	2.87
G76S	3.7	56.2	11.9	28.1	2.20

responded in the same manner to antibody binding, indicating that spatial conformation of the proteins was similar. Additionally, structural analyses indicated that the conformation of the mutants in the amide region was the same as the wild-type. Structural perturbations were observed in the aromatic region and were associated with local differences of the flap. Therefore, it was concluded that mutations in position 76 did not change overall enzyme conformations and had only regional effects.

Rate constants of the activation process were calculated for the mutants and the wild-type under three different pH conditions: pH 1.1, 2.0 and 3.0. Samples were taken at different activation times, quenched, and the amount of pepsin formed was determined with synthetic substrate I, as described in the Experimental procedures section. The activation process of pepsinogen fused with thioredoxin was shown to be unimolecular, contrary to the bimolecular activation process of native and recombinant pepsinogens [33]. Therefore, the activation rate constants, presented in Table 3, correspond to first-order reaction constants. All the mutants activated more slowly than the wild-type, regardless of amino acid size and polarity. Under all pH conditions the activation reactions were the slowest for valine and serine mutants, which had comparable reaction rates. The alanine mutant activated slower than the wild-type protein but faster than the two other mutants.

Kinetic properties of the wild-type and the mutant enzymes were determined with two synthetic substrates and the kinetic constants, K_m and k_{cat}, were calculated using the non-linear least-squares method [30]. The kinetic parameters are presented in Table 4. All the mutants, regardless of amino acid size and polarity, had lower substrate affinity and turnover numbers than the wild-type enzyme. The differences in the catalytic parameters were similar for the two synthetic substrates, indicating that the changes were not substrate-specific and that the results could, therefore, be attributed to differences in the catalytic environments of the mutants. The kinetic experiments were performed at two different pH values: pH 2.1 for synthetic substrate I and pH 3.95 for synthetic substrate II. The kinetic constants showed the same tendencies under the two different pH conditions. Therefore, it was concluded that the mutations did not change the pH dependencies of mature enzymes. The above results indicated that the mutations had minimal effect on substrate binding and mainly influenced the properties of enzyme-bound species in the reaction pathway. In addition, stereochemical analysis of peptide-bond hydrolysis [16] and pepsin crystal structure [8] indicated that G76 is in a position most favourable for interactions with reaction intermediates. A possible involvement of G76 in stabilizing the transition state was indicated by the authors of the hypothetical catalytic mechanism for pepsin [17].

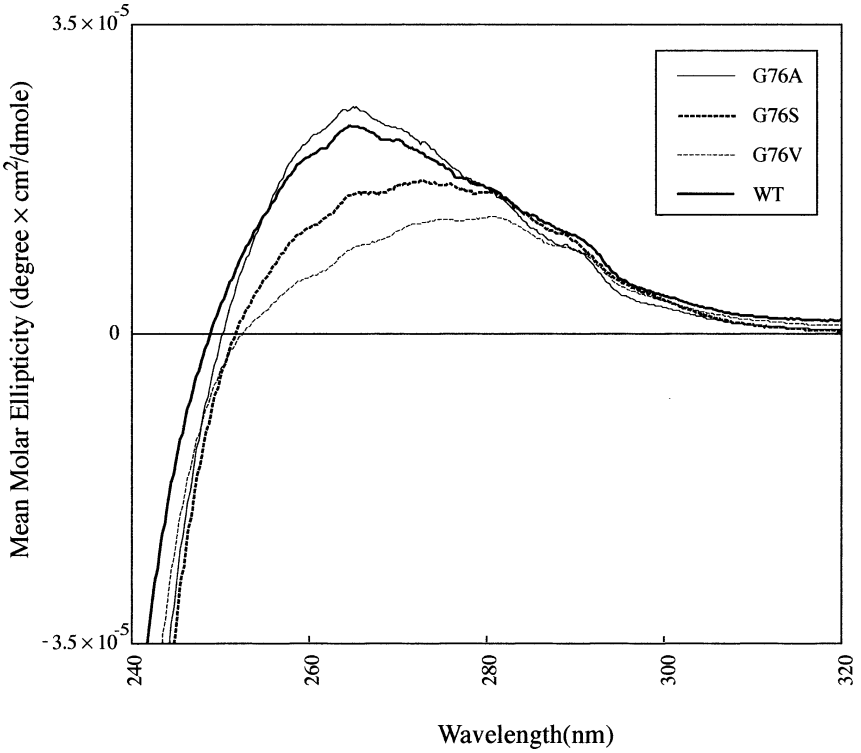

Figure 4 Near-UV spectra of the wild-type and mutant pepsins

Experiments were carried in 10 mM sodium acetate buffer, pH 5.3. Spectra were generated using 1 ml of 1 mg/ml purified enzyme solutions in a 1-cm pathlength cuvette after eight scans from 320 to 240 nm at room temperature under continuous nitrogen flush. Spectra were obtained from 1 mg/ml pepsin solution in 10 mM sodium acetate buffer, pH 5.3.

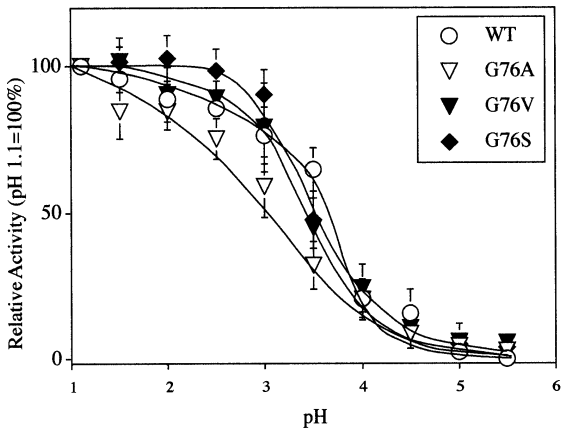

Figure 5 Relative amounts of pepsin formed during activation under different pH conditions for fusion-protein G76 mutants as determined by milk-clotting assay

Each point represents the mean ± S.D. of three measurements.

Table 3 First-order rate constants for activation of Trx-PG

Enzyme	k_{act} (min^{-1}) pH 1.1	pH 2.0	pH 3.0
Wild-type	4.449 ± 0.262	2.819 ± 0.152	1.043 ± 0.089
G76A	0.311 ± 0.076	0.332 ± 0.014	0.104 ± 0.029
G76V	0.132 ± 0.001	0.152 ± 0.013	0.054 ± 0.001
G76S	0.13 ± 0.014	0.166 ± 0.014	0.059 ± 0.003

The trend in the activation rate constants was similar to that for the catalytic rate constants of the mutants and wild-type enzyme. All the mutants were less efficient in catalysis and activation processes than the wild-type, and the alanine mutant had higher catalytic and activation efficiency than the two other mutants, valine and serine. These results indicate that G76 plays the same role in the activation and catalytic processes. During the activation process, a prosegment represents a substrate. Therefore, the results suggest that proteolytic cleavage is a limiting step of the activation process. This is in agreement with earlier results that any changes in pepsin flap sequence in position 77 had the same effect on the activation and catalytic processes and suggested that proteolytic cleavage limits the activation process of the fusion proteins [15].

All the mutants had altered kinetic constants compared with the wild-type; however, the degree of change was different for individual enzymes. The alanine mutant was affected less than either the valine or serine mutants, which had comparable catalytic constants. Alanine has the smallest van-der-Waals volume and accessible surface area compared with valine and serine. Therefore, it places a relatively smaller restriction on a polypeptide's conformational flexibility and has relatively less contacts with surrounding atoms. Valine and serine mutants have larger volumes and their presence in a polypeptide chain results in greater conformational restrictions due to the atom-

Table 4 Kinetic parameters for wild-type pepsin and the G76 mutants

Kinetic constants were calculated based on a minimum of two measurements for each enzyme.

Substrate	Enzyme	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)
(I) KPAEFF(NO_2)AL	Wild-type	0.043 ± 0.005	180.5 ± 10.3	4198 ± 544
	G76A	0.098 ± 0.008	47.0 ± 9.9	395 ± 96
	G76V	0.071 ± 0.011	11.1 ± 1.1	160 ± 30
	G76S	0.085 ± 0.005	2.5 ± 0.6	29 ± 7
(II) LSF(NO_2)NleAL-OCH_3	Wild-type	0.011 ± 0.002	88.7 ± 4.6	8064 ± 1525
	G76A	0.040 ± 0.009	43.4 ± 3.6	1085 ± 260
	G76V	0.044 ± 0.005	1.2 ± 0.1	28 ± 3
	G76S	0.041 ± 0.008	2.9 ± 0.3	70 ± 16

Table 5 Distances between atoms in the tip of the pepsin flap in pepsinogen, pepsin and inhibited pepsin (1.8–2.9 Å resolution)

I represents an inhibitor atom.

			Distance (Å)		
Atoms		Enzyme form ...	Pepsinogen	Pepsin	Inhibited pepsin
76:NH	75:O		2.25	2.28	2.28
	76:O		3.65	3.63	3.37
	I1:OH		–	–	3.3
	I3:O		–	–	2.9
76:O	77:NH		2.20	2.23	2.28
	77:OH		4.61	4.92	2.82
77:NH	77:OH		3.1	2.87	2.82

crowding effect. Consequently, the results suggested that all the mutations reduced polypeptide conformational freedom in the flap region. This is in agreement with molecular-dynamics calculations and structural analyses of the mutants, where the greatest local perturbations were observed for the G76V mutant. Similar observations were made in catalytic loops of other enzymes. The conversion of glycine in a flexible loop of F_1-ATPase into any residue bulkier than proline diminished the conformational mobility of the segment or destabilized β-turn [34]. Mutations of glycine 121 in the flexible loop of dihydrofolate reductase to valine and leucine modified loop flexibility and impaired its active-site lid function [35]. In glutathione synthetase, replacement of glycines and proline in the active-site loop with alanine and valine reduced flexibility of the loop [19]. Our results, combined with observations made in other loop structures, demonstrate that glycine in pepsin position 76 contributes to the flap flexibility and that this flexibility is essential for efficient catalysis.

In addition to reduced flap flexibility, the mutants could cause different distribution of hydrogen bonds in pepsin's binding cleft, especially the G76S mutant. The polar mutant had catalytic parameters lower than those of the wild-type and of a similar magnitude to the G76V mutant. However, serine's volume and accessible surface area were closer to alanine than valine. Therefore, if only reduced flap flexibility contributed to the lower catalytic parameters of G76S they should have resembled the catalytic efficiency of G76A rather than of G76V. However, the G76A mutant had catalytic parameters that were affected less than both the G76S and G76V mutants. At the same time, G76S introduced an additional polar group and, therefore, a hydrogen-bonding potential, as compared with the wild-type and the other mutants. Consequently, our results suggested that the additional hydrogen bond in the G76S mutant distorted the hydrogen-bond network in the active site in addition to the effect of reduced flap flexibility. In addition, molecular-dynamics calculation indicated that the G76V mutant would result in increased flexibility of T77 and, therefore, reduced probability of hydrogen-bond formation by the hydroxy group in position 77. Therefore, both G76S and G76V mutations would have resulted in altered hydrogen-bond distribution in comparison with the wild-type. The results support previous observations that the catalytic process in pepsin is sensitive to any changes in a delicate network of hydrogen bonds that participate in substrate alignment [10,15]. Differences in the hydrogen-bond patterns in aspartic proteinase binding clefts may contribute to substrate-specificity differences among the group members. For example, renin is the only aspartic proteinase that has serine in position 76 (Table 1) and, at the same time, it is highly specific towards

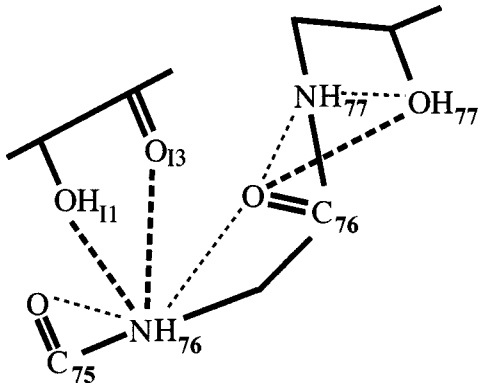

Figure 6 Schematic diagram of hydrogen bonds made by G76 and T77

Inhibitor atoms are denoted by the letter I. Hydrogen bonds present in pepsinogen, pepsin and inhibited pepsin are shown as thin broken lines. Distances equivalent to newly formed hydrogen bonds in inhibited pepsin are represented as thick broken lines.

angiotensinogen [36]. Consequently, S76 in renin (pepsin numbering) is likely to contribute to the lower flap mobility and different hydrogen-bond pattern and, therefore, to the high specificity of this enzyme.

The reduced flexibility of the flap could indirectly affect hydrogen-bond distribution near the active site. In fungal and mammalian pepsins it has been shown that hydrogen-bonding potential in position 77 is essential for enzyme specificity and efficient catalysis [15,37]. Lower conformational freedom of the mutated flap could prohibit the residue in position 77 from forming a hydrogen bond with a substrate that was supported by energy minimization. In addition, the steric hindrance caused by overcrowding of atoms in G76 mutants could change a pattern of hydrogen bonds made by the G76 backbone and, therefore, the stability of reaction intermediates.

The distribution of hydrogen bonds in loops was thought to be responsible for their conformations [4]. Therefore, we also investigated hydrogen bonds between the flap backbone and side-chain atoms in pepsinogen, pepsin and inhibited pepsin. Table 5 lists possible distances of hydrogen bonds that were identified at the tip of the flap from crystal structures obtained at resolutions of 1.8–2.9 Å. Therefore, bond distances should be considered with a 1.8 Å confidence region. Figure 6 shows a schematic diagram of the hydrogen-bond distances at the tip of the flap in the three enzyme forms. Comparison of the three

enzyme forms, i.e. zymogen, active and inhibited pepsin, revealed that during catalysis changing distances in the flap region corresponded to formation of hydrogen bonds between G76-backbone atoms and an inhibitor and T77. In the inhibited enzyme, the amide nitrogen of G76 came into hydrogen-bonding distance of the hydroxy and carbonyl oxygens of an inhibitor, and a carbonyl oxygen in position 76 came into hydrogen-bonding distance of the hydroxy group of T77. These results indicate that G76-backbone atoms participate in a hydrogen bonding to a substrate and aid in positioning the T77 hydroxy group in the most favourable conformation for substrate alignment. Similar observations were made in the flexible loop of triosephosphate isomerase, where G171 hydrogen-bonded to a ligand and in this way was essential for substrate alignment [38]. Likewise, the backbone nitrogen of loop glycine in lectin formed an essential part of the monosaccharide-binding site. In dihydrofolate reductase, glycine located in the catalytically important loop indirectly influenced ligand conformation by positioning the active-site loop [39–41].

The four remaining hydrogen bonds between atoms of residues 75, 76 and 77 are present in all pepsin forms, which indicates that they are responsible for the conformation of the flap tip and preservation of a turn structure. The same observations were made in fructose-1,6-bisphosphatase, dihydrofolate reductase and thymidylate synthase [41–43], where networks of hydrogen bonds have been found to stabilize conformations of catalytically important loops. This is in agreement with findings that internal-loop hydrogen bonds are essential for maintaining the conformation of these structures [4].

All the observations made of the flap hydrogen-bond patterns suggest that the carbonyl oxygen of G76 helps to position the 77 hydroxy group, and that the 76 backbone nitrogen stabilizes the transition state. All the mutants in position 76 had altered catalytic properties since T77 was not able to adopt the proper conformation and G76-backbone hydrogen bonds were impaired. This is in agreement with previous observations that preservation of hydrogen-bond potential, as in a T77S pepsin mutant, did not change enzyme catalytic properties [15]. A different distribution of hydrogen bonds in all G76 mutants and T77 apolar mutants could distort flap conformation by changing interactions of atoms that maintain the flap structure. Consequently, different distribution of the hydrogen bonds at the flap tip could affect substrate alignment, geometry of the transition state and change conformation of the flap.

In conclusion, the results indicated that G76 is involved in the catalytic process. Introduction of any additional side chain to this position, as in the G76A, G76V and G76S mutants, resulted in less efficient catalysis and a slower activation process. Alanine, which has the smallest side chain of the three substituted amino acids, had the smallest effect on enzyme catalysis and local conformation. Introduction of bulkier amino acids, serine and valine, resulted in greater decreases in catalytic parameters and activation rate constants. Therefore, it was concluded that G76 is essential for pepsin-catalysed proteolysis by contributing to the flap flexibility. We suggest that flexibility is essential for hydrogen bonding of the T77 hydroxy group to a substrate. Detailed examination of possible hydrogen-bond distances at the tip of the flap revealed that G76-backbone atoms participated in maintenance of the flap conformation and in the geometry of reaction intermediates. Based on our results and observations made of structures of other enzymes, it is postulated that terminal glycine residues in catalytically important loops and turns serve to enable movement of the residues essential in enzyme action.

All the above results and our previous studies [15] combined indicate that the tip of the flap, i.e. residues 76 and 77, participates in the catalytic and activation processes by hydrogen bonding to a substrate, or the prosegment, respectively. Binding is possible by changing the position of the flap, which is facilitated by G76 contributing to flap flexibility. It is suggested that G76 acts as a hinge at the tip of the flap and enables formation of the hydrogen bond by T77, and that G78, which is preserved in all the compared aspartic proteinases (Table 1), serves as a second hinge of the flap tip. This is in agreement with observations made in the crystal structures of aspartic proteinases [14] that the tip of the flap, residues 76–78 (pepsin numbering), is the most flexible fragment of this structure. Our hypothesis is supported further by the results of Tanaka et al. [19] and Schneider et al. [20], who found that glycines act as catalytic loop hinges in glutathione synthetase and tryptophan synthase, respectively. A change of the flap position enables interactions of position-76 backbone nitrogen with reaction intermediates and possibly contributes to stabilization of the transition state, which is in agreement with previous suggestions [16,17]. Additionally, we have shown that G76 and T77 contribute to the hydrogen-bond network that is responsible for substrate alignment and efficient catalysis. The flap's internal hydrogen bonds are responsible for the conformation of this structure.

REFERENCES

1 Leszczynski, J. F. and Rose, G. D. (1986) Loops in globular proteins: a novel category of secondary structures. Science **234**, 849–855
2 Fetrow, J. S. (1995) Omega loops: nonregular secondary structures significant in protein function and stability. FASEB J. **9**, 708–717
3 Rose, G. D., Gierasch, L. M. and Smith, J. A. (1985) Turns in peptides and proteins. Adv. Protein Chem. **37**, 1–109
4 Tramontano, A., Chothia, C. and Lesk, A. M. (1989) Structural determinants of the conformations of medium-sized loops in proteins. Proteins **6**, 382–394
5 Gerstein, M. and Chothia, C. (1991) Analysis of protein loop closure. Two types of hinges produce one motion in lactate dehydrogenase. J. Mol. Biol. **220**, 133–149
6 Sampson, N. S. and Knowles, J. R. (1992) Segmental motion in catalysis: investigation of a hydrogen bond critical for loop closure in the reaction of triosephosphate isomerase. Biochemistry **31**, 8488–8494
7 Kempner, E. S. (1993) Movable lobes and flexible loops in proteins. FEBS Lett. **326**, 4–10
8 Sielecki, A. R., Fedorov, A. A., Boodhoo, A., Andreeva, N. S. and James, M. N. G. (1990) Molecular and crystal structures of monoclinic porcine pepsin refined at 1.8 Å resolution. J. Mol. Biol. **214**, 143–170
9 Tang, J., James, M. N. G., Hsu, I. N., Jenkins, J. A. and Blundell, T. L. (1978) Structural evidence for gene duplication in the evolution of the acid proteases. Nature (London) **271**, 618–621
10 Hartsuck, J. A., Koelsch, G. and Remington, J. S. (1992) The high-resolution crystal structure of porcine pepsinogen. Proteins **13**, 1–25
11 Chen, L., Erickson, J. W., Rydel, T. J., Park, C. H., Neidhart, D., Luly, J. and Abad-Zapatero, C. (1992) Structure of pepsin/renin inhibitor complex reveals a novel crystal packing induced by minor chemical alterations in the inhibitor. Acta Cryst. **B48**, 476–488
12 Glick, M. D., Hilt, C. R. and Mende-Mueller, L. (1991) Conformational change that accompanies pepsinogen activation observed in real time by fluorescence energy transfer. Int. J. Peptide Protein Res. **37**, 230–235
13 Tanaka, T., Teo, K. S., Lamb, K. M., Harris, L. and Yada, R. Y. (1998) Effect of replacement of the conserved Tyr75 on the catalytic properties of porcine pepsin A. Protein Peptide Lett. **5**, 19–26
14 Andreeva, N., Bochkarev, A. and Pechik, I. (1995) A new way of looking at aspartic proteinases: a comparison of pepsin structure to other aspartic proteinases in the near active site region, in Aspartic Proteinases (Takahashi, K., ed.), pp. 19–32, Plenum Press, New York
15 Okoniewska, M., Tanaka, T. and Yada, R. (1999) The role of flap residue, threonine 77, in the activation and catalytic activity of pepsin A. Protein Eng. **12**, 55–61
16 James, M. N. G. and Sielecki, A. R. (1985) Stereochemical analysis of peptide bond hydrolysis catalyzed by the aspartic proteinase penicillopepsin. Biochemistry **24**, 3701–3713
17 Pearl, L. H. (1987) The catalytic mechanism of aspartic proteinases. FEBS Lett. **214**, 8–12
18 Creighton, T. E. (1994) Proteins. Structures and Molecular Properties, 2nd edn., W. H. Freeman and Company, New York

19 Tanaka, T., Yamaguchi, H., Kato, H., Nishioka, T., Katsube, Y. and Oda, J. (1993) Flexibility impaired by mutations revealed the multifunctional roles of the loop in glutathione synthetase. Biochemstry **32**, 12398–12404

20 Schneider, T. R., Gerhardt, E., Lee, M., Liang, P.-H., Anderson, K. and Schlichting, I. (1998) Loop closure and intersubunit communication in tryptophan synthase. Biochemistry **37**, 5394–5406

21 Fry, D. C., Kuby, S. A. and Mildvan, A. S. (1986) ATP-binding site of adenylate kinase: mechanistic implications of its homology with *Ras*-encoded p21, F_1-ATPase, and other nucleotide-binding proteins. Proc. Natl. Acad. Sci. U.S.A. **83**, 907–911

22 Wilks, H. M., Moreton, K. M., Halsall, D. J., Hart, K. W., Sessions, R., Clarke, A. R. and Holbrook, J. J. (1992) Design of a specific phenyllactate dehydrogenase by peptide loop exchange on the *Bacillus stearothermophilus* lactate dehydrogenase framework. Biochemistry **31**, 7802–7806

23 Larson, E. M., Larimer, F. W. and Hartman, F. C. (1995) Mechanistic insights provided by deletion of a flexible loop at the active site of ribulose-1,5-bisphosphate carboxylase/oxygenase. Biochemistry **34**, 4531–4537

24 Kunkel, T. A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. U.S.A. **82**, 488–492

25 Tanaka, T. and Yada, R. Y. (1996) Expression of soluble cloned porcine pepsinogen A in *Escherichia coli.* Biochem. J. **315**, 443–446

26 Lin, X.-l., Wong, R. N. S. and Tang, J. (1989) Synthesis, purification, and active site mutagenesis of recombinant porcine pepsinogen. J. Biol. Chem. **264**, 4482–4489

27 McPhie, P. (1976) The origin of the alkaline inactivation of pepsinogen. Anal. Biochem. **73**, 258–261

28 Chang, C. T., Wu, C.-S. C. and Yang, J. T. (1978) Circular dichroic analysis of protein conformation: inclusion of the β-turns. Anal. Biochem. **91**, 13–31

29 Lin, Y., Fusek, M., Lin, X., Hartsuck, J. A., Kezdy, F. J. and Tang, J. (1992) pH dependence of kinetic parameters of pepsin, rhizopuspepsin, and their active-site hydrogen bond mutants. J. Biol. Chem. **267**, 18413–18418

30 Sakoda, M. and Hiromi, K. (1976) Determination of the best-fit values of kinetic parameters of the Michaelis–Menten equation by the method of least squares with Taylor expansion. J. Biochem. (Tokyo) **80**, 547–555

31 Andreeva, N. S., Zdanov, A. S., Gustchina, A. E. and Fedorov, A. A. (1984) Structure of ethanol inhibited porcine pepsin at 2 Å resolution and binding of the methyl ester of phenylalanyl-diiodotyrosine to the enzyme. J. Biol. Chem. **259**, 11353–11365

32 Lee, A. Y., Sergei, V. G. and Erickson, J. W. (1998) Conformational switching in an aspartic proteinase. Nat. Struct. Biol. **5**, 866–871

33 Tanaka, T. and Yada, R. Y. (1997) Engineered porcine pepsinogen exhibits dominant unimolecular activation. Arch. Biochem. Biophys. **340**, 355–358

34 Seeburg, P. H., Colby, W. W., Capon, D. J., Goeddel, D. V. and Levinson, A. D. (1984) Biological properties of human C-Ha-ras1 genes mutated at codon 12. Nature (London) **312**, 71–75

35 Gekko, K., Yamagami, K., Kunori, Y., Ichihara, S., Kodama, M. and Iwakura, M. (1993) Effects of point mutation in a flexible loop on the stability and enzymatic function of *Escherichia coli* dihydrofolate reductase. J. Biochem. **113**, 74–80

36 Fujinaga, M., Chernaia, M. M., Tarasova, N. I., Mosimann, S. C. and James, M. N. G. (1995) Crystal structure of human epsin and its complex with pepstatin. Protein Sci. **4**, 960–972

37 Shintani, T., Nomura, K. and Ichishima, E. (1997) Alteration of S_1 substrate specificity of pepsin to those of fungal aspartic proteinases by site-directed mutagenesis. J. Biol. Chem. **272**, 18855–18861

38 Pompliano, D. L., Peyman, A. and Knowles, J. R. (1990) Stabilization of a reaction intermediate as a catalytic device: definition of the functional role of the flexible loop in triosephosphate isomerase. Biochemistry **29**, 3186–3194

39 Yuksel, U.K., Sun, A.-Q., Gracy, R. W. and Schnackerz, K. D. (1994) The hinged lid of yeast triose-phosphate isomerase. J. Biol. Chem. **269**, 5005–5008

40 Hoedemaeker, F. J., van Eijsden, R. R., Diaz, C. L., de Pater, B. S. and Kijne, J. W. (1993) Destabilization of pea lectin by substitution of a single amino acid in a surface loop. Plant Mol. Biol. **22**, 1039–1046

41 Miller, G. P. and Benkovic, S. J. (1998) Deletion of highly motional residue affects formation of the Michaelis complex for *Escherichia coli* dihydrofolate reductase. Biochemistry **37**, 6327–6335

42 Choe, J.-Y., Poland, B. W., Fromm, H. J. and Honzatko, R. B. (1998) Role of a dynamic loop in cation activation and allosteric regulation of recombinant porcine fructose-1,6,-bisphosphatase. Biochemistry **37**, 11441–11450

43 Schiffer, C. A., Clifton, I. J., Davisson, V. J., Santi, D. V. and Stroud, R. M. (1995) Crystal structure of human thymidylate synthase: a structural mechanism for guiding substrates into the active site. Biochemistry **34**, 16279–16287

44 Abad-Zapatero, C., Rydel, T. J. and Erickson, J. (1990) Revised 2.3 Å structure of porcine pepsin: evidence for a flexible subdomain. Proteins **8**, 62–81

Received 15 December 1999/31 March 2000; accepted 25 April 2000