

# Structural Insights into the Mechanism of Intramolecular Proteolysis

Qian Xu,\* Deirdre Buckley,\*<sup>§</sup> Chudi Guan,<sup>†</sup> and Hwai-Chen Guo\*<sup>†</sup>

\*Department of Biophysics  
Boston University School of Medicine  
Boston, Massachusetts 02118-2526

<sup>†</sup>New England Biolabs  
Beverly, Massachusetts 01915-5599

## Summary

A variety of proteins, including glycosylasparaginase, have recently been found to activate functions by self-catalyzed peptide bond rearrangements from single-chain precursors. Here we present the 1.9 Å crystal structures of glycosylasparaginase precursors that are able to autoproteolyze via an N → O acyl shift. Several conserved residues are aligned around the scissile peptide bond that is in a highly strained *trans* peptide bond configuration. The structure illustrates how a nucleophilic side chain may attack the scissile peptide bond at the immediate upstream backbone carbonyl and provides an understanding of the structural basis for peptide bond cleavage via an N → O or N → S acyl shift that is used by various groups of intramolecular autoprocessing proteins.

## Introduction

The posttranslational cleavage of peptide bonds is necessary for diverse biological processes such as zymogen activation (Khan and James, 1998), fibrinolysis and blood coagulation (Davie, 1995), protein degradation and antigen presentation (Baumeister et al., 1998), regulation of sodium balance and blood pressure (Vallet et al., 1997), regulation of the cell cycle (King et al., 1996), development (Chan and Jan, 1998), apoptosis (Salvesen and Dixit, 1997), and virus infection and membrane fusion (Klenk and Garten, 1994), as well as cytoprotection in the intestine or bronchial airways (Cocks et al., 1999). Many of these processes are mediated by regulatory factors or induced by a pH change. Recently, proteolytic cleavages to activate inactive precursors have been found to be self-catalyzed intramolecular processes in a number of important biological processes. These include autocleavage of Hedgehog proteins (Lee et al., 1994), protein splicing (Paulus, 1998; Perler, 1998a), maturation of pyruvoyl-dependent enzymes (Recsei et al., 1983), and autoproteolysis of N-terminal nucleophile (Ntn) hydrolase precursors (Zwickl et al., 1994; Brannigan et al., 1995; Aronson, 1996; Guan et al., 1996; Schmidtke et al., 1996).

Glycosylasparaginase (GA) is the prototype of proteins that self-catalyze peptide bond rearrangement

through an N → O or N → S acyl shift (Paulus, 1998; Perler, 1998b). In humans, a deficiency of GA leads to the disease aspartylglycosaminuria (Mononen et al., 1993). Central to the activities of GA is the conserved Thr-152 that provides the nucleophile in both autoproteolysis (Guan et al., 1996) and hydrolase activity (Kartinen et al., 1991; Tarentino et al., 1995). Facilitated by various functional groups, the side chain of Cys, Ser, or Thr is also utilized as a nucleophile in other autoprocessing proteins to activate the immediate upstream peptide bond and to substitute the peptide amide bond with a more reactive (thio)ester bond. These N → O or N → S acyl shifts initiate diverse autoprocessing pathways. Attack by a second nucleophile of various types results in cleavage of the linkage between the Cys, Ser, or Thr and the preceding amino acid. In a simple autoproteolytic reaction catalyzed by Ntn hydrolases, an activated water molecule serves as the second nucleophile (Guan et al., 1996). In the autoprocessing of the *hedgehog* family of eukaryotic developmental regulatory proteins, the hydroxyl group of cholesterol displaces the thioester and results in a covalent attachment of cholesterol to the signaling domain (Porter et al., 1996). In the case of protein splicing, the thiol or hydroxyl group of the Cys, Ser, or Thr at the junction of intein and C extein attacks the initial (thio)ester linkage, forming a branched protein intermediate and linking the N and C exteins with a second (thio)ester bond (Xu et al., 1994). Unlike the above autoprocessing pathways, breakdown of the ester intermediate in pyruvoyl enzymes is accomplished by β elimination, resulting in an N-terminal dehydroalanine that further hydrolyzes to form the catalytic pyruvoyl group at the N terminus of the α subunit (Recsei et al., 1983). Most of the evidence for the acyl intermediate has been derived indirectly from chemical and mutagenesis studies. Now, direct evidence of an ester intermediate has been found in the crystal structure of L-aspartate-α-decarboxylase (Albert et al., 1998).

The structures of precursors (proteins) are essential to understand the detailed mechanisms of autoprocessing reactions that proceed via an N → O or N → S acyl shift. Although structures of the processing products such as free inteins (Duan et al., 1997), autocleaved Ntn hydrolases (Smith et al., 1994; Brannigan et al., 1995; Duggleby et al., 1995; Löwe et al., 1995; Oinonen et al., 1995; Aronson, 1996; Isupov et al., 1996; Groll et al., 1997; Guo et al., 1998; Xuan et al., 1998), autoprocessed Hedgehog proteins (Hall et al., 1997), and pyruvoyl-containing proteins (Gallagher et al., 1993; Albert et al., 1998), as well as an inactive precursor (Ditzel et al., 1998), have revealed potential candidates that may be involved in facilitating (thio)ester formation and cleavage, there is always the possibility that slight conformational adjustments may have occurred after the formation and cleavage of the (thio)ester bond. Recently, a modified intein with a single alanine extension at its N terminus revealed a *cis* conformation at the peptide bond that would have been cleaved in an active precursor (Klabunde et al., 1998). However, it is unclear

<sup>†</sup>To whom correspondence should be addressed (e-mail: hcguo@bu.edu).

<sup>§</sup>Present address: Department of Biochemistry, University College Cork, Lee Maltings, Prospect Row, Cork, Ireland.

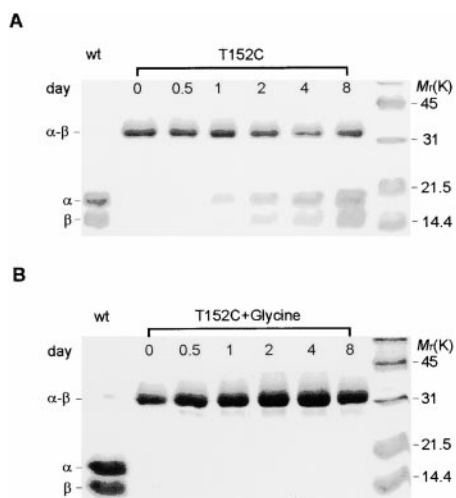


Figure 1. Activity and Inhibition of Autoproteolysis of the T152C Precursor

(A) One milliliter of T152C precursor (stabilized by glycine) was dialyzed at 4°C for 1 hr against 1 l of reaction buffer without glycine (20 mM Tris [pH 7.4], 50 mM NaCl, 1 mM EDTA). Autoproteolysis was initiated by incubating the precursor at 37°C. Samples were taken out at various time points, as indicated, and analyzed by SDS-PAGE. The precursor ( $\alpha$ - $\beta$ ) and autocleaved subunits ( $\alpha$  and  $\beta$ , wild-type in left lane) are marked.

(B) The same as in (A), except the 37°C incubation is in the presence of 10 mM glycine. The proteins were overloaded in (B) to show the absence of autoproteolytic products at time points up to 8 days.

whether a single amino acid could reflect the native folding and packing of the intact N-extein domain (130 amino acids) or have the same energy barrier of *cis-trans* isomerization at the putative scissile peptide bond. In this work, we have carried out the structural determination of intact precursors that are active in autoproteolysis via an  $N \rightarrow O$  acyl shift. To this end, we have generated several mutants that autoproteolyze but can be reversibly and indefinitely inhibited by glycine. Biochemical evidence is presented to show that the mutant GA precursors undergo the same autoproteolysis reaction as the wild-type precursor. The GA precursor structures reported here define a self-catalytic center for peptide bond rearrangement via an  $N \rightarrow O$  or  $N \rightarrow S$  acyl shift.

## Results and Discussion

### Biochemical and Kinetic Characterization of Mutant GA Precursors

The GA precursor autoproteolyzes spontaneously at the scissile peptide bond between conserved residues Asp-151 and Thr-152. For crystallographic studies of the autoproteolytic mechanism, active site mutants that have a much slower rate of autoproteolysis were isolated (Guan et al., 1998). Two active mutants were chosen for this study: W11F, in which Trp-11 was replaced by a Phe, and T152C, in which Thr-152 was replaced by a Cys. As a control, we also studied an inactive mutant precursor, T152A, in which Thr-152 was replaced by an alanine.

Biochemical and kinetic data indicate that W11F and

Table 1. Kinetic Parameters of Autoproteolysis for Variants of GA Precursor

	$k_{cat}$ ( $s^{-1}$ )
WT	$1.2 \times 10^{-2}$
W11F	$2.0 \times 10^{-5}$
T152S	$5.8 \times 10^{-4}$
T152C	$3.8 \times 10^{-6}$
T152A	Not detectable
T170S	$8.3 \times 10^{-4}$
T170C	$2.0 \times 10^{-5}$
T170A	$3.8 \times 10^{-6}$

In vitro autoproteolysis was analyzed by SDS-PAGE at various time points. The overall rate constants in the activation buffer (20 mM Tris [pH 7.4]; 50 mM NaCl; 1 mM EDTA) were determined as described previously (Guan et al., 1998).

T152C mutants retained autoproteolysis activity. Like the wild-type precursor, both mutants underwent the same autocleavage into  $\alpha$  and  $\beta$  subunits (Figure 1A), and the reaction followed first-order kinetics (Guan et al., 1996). The W11F mutant had a 600-fold reduction in the reaction rate, whereas the T152C mutant had its rate reduced by three orders of magnitude (Table 1). Autoproteolysis of GA mutants can be accelerated by hydroxylamine, suggesting the intermediate is an ester linkage (Guan et al., 1996). The sites of autocleavage have been defined by N-terminal protein sequencing and mass spectrometry to be the same as the wild-type precursor (data not shown). By contrast, the T152A precursor had completely lost its autoproteolytic activity, even in the presence of hydroxylamine. CD spectra of the T152A precursor are identical to that of other active precursors (data not shown), suggesting that the lack of activity is not the result of misfolding. Instead, the lack of a nucleophilic group in the T152A precursor may prevent an  $N \rightarrow O$  or  $N \rightarrow S$  acyl shift.

In an effort to further slow down autoproteolysis to prepare precursors for crystallization, we found that small amino acids, such as glycine, reversibly inhibit the reaction (Guan et al., 1998). In the presence of 10 mM glycine, the T152C precursor was stable up to 8 days (Figure 1B). After removal of glycine by dialysis, the T152C precursor underwent autoproteolysis at position 152 (Figure 1A). Similarly, autocleavage of the W11F precursor was reversibly inhibited by glycine. The recombinant precursor variants were thus produced with glycine in the cell media and purification solutions. No autoproteolysis was detected for the T152A precursor in the absence or presence of glycine.

### Precursor Structures

The structures of three GA precursors were determined by X-ray cryocrystallography and refined to 1.9 Å resolution (Table 2). Apart from differences in the autoproteolytic centers where a single residue was changed (see below), the three precursor structures are essentially identical (root-mean-square deviation [rmsd] = 0.22–0.43 Å for 278 residues). In the single-chain precursor, a surface loop (P loop) of 15 residues makes the connection between domains (Figure 2); this loop is released from the active site and becomes disordered in the autocleaved enzyme (Guo et al., 1998). With the exception

Table 2. Statistics from the Crystallographic Analysis

	T152A	Hg(OAc) <sub>2</sub>	T152C	W11F
<b>Data collection</b>				
Resolution (Å)	2.1	2.5	1.9	1.9
Beamline	Lab Rigaku	Lab Rigaku	NSLS-X4a	NSLS-X4a
Observations	52,951	31,308	89,551	88,282
Unique reflections	29,429	18,452	41,303	43,920
$I/\sigma_I$	13.8	12.0	9.4	27.2
Data coverage (%)	90.9	95.8	94.3	97.5
$R_{\text{sym}}$ (%) <sup>a</sup>	7.6	10.2	5.6	2.6
$R_{\text{iso}}$ (%) <sup>b</sup>		25.9		
<b>SIRAS analysis</b>				
Heavy atom sites		4		
Phasing power <sup>c</sup>		1.55		
$R_{\text{cullis}}$ <sup>d</sup>		0.74		
SIGMAA combined phases <sup>e</sup> : FOM <sup>f</sup>		0.767		
DM modified phases: FOM <sup>f</sup>		0.867		
<b>Refinement</b>				
<b>Nonhydrogen atoms</b>				
Protein <sup>g</sup>	4,238		4,269	4,440
Water	124		151	226
$R_{\text{cryst}}$ (%) <sup>h</sup> : all F	23.18		23.15	19.60
$R_{\text{free}}$ (%) <sup>i</sup> : all F	29.73		27.82	23.96
<b>Rms deviations<sup>j</sup></b>				
Bond length (Å)	0.019		0.018	0.017
Bond angle (°)	3.3		3.6	2.9
B value (Å <sup>2</sup> )	4.8		4.7	5.5
NCS related molecules (Å)	0.38		0.30	0.17
<b>Average B value (Å<sup>2</sup>)</b>				
Main chain	10.2		9.9	7.8
Side chain	11.3		11.4	9.8
Water	16.4		16.4	15.3
<b>Residues in Ramachandran plot (%)</b>				
Most favored	91.3		91.3	90.3
Additional	8.7		8.7	9.7
Disallowed	0.0		0.0	0.0

<sup>a</sup>  $R_{\text{sym}} = \sum_n \sum_i |I_n - I_i| / \sum_n \sum_i I_i$  for the intensity ( $I$ ) of  $i$  observations of reflection  $h$ .

<sup>b</sup>  $R_{\text{iso}} = \sum_n |I_d - I_n| / \sum_n I_n$ , where  $I_d$  and  $I_n$  are the observed derivative and native (T152A) intensities, respectively.

<sup>c</sup> Phasing power =  $\langle F_H \rangle / E$ , where  $\langle F_H \rangle$  is the mean heavy atom structure factor and  $E$  is the mean lack of closure error.

<sup>d</sup>  $R_{\text{cullis}}$  is the lack of closure error divided by the isomorphous difference.

<sup>e</sup> Initial phases were obtained through a combination of SIRAS and MR using the models of mature enzyme (Guo et al., 1998).

<sup>f</sup> Figure of merit =  $|F(\text{hkl})_{\text{best}}| / F(\text{hkl})$ .

<sup>g</sup> The number of atoms includes both molecules in the asymmetric unit.

<sup>h</sup>  $R_{\text{cryst}}$  and  $R_{\text{free}} = \sum |F_{\text{obs}} - F_{\text{calc}}| / \sum |F_{\text{obs}}|$ , where  $F_{\text{obs}}$  and  $F_{\text{calc}}$  are the observed and calculated structure factor amplitudes, respectively.

<sup>i</sup>  $R_{\text{free}}$  was calculated with 8% of the amplitudes chosen randomly and omitted from the start of refinement.

<sup>j</sup> Rms deviations are deviations from ideal geometry, root-mean-square variation in the B factor of bonded atoms, and deviations for all main chain atoms between the two molecules in the asymmetric unit.

of the P loop, the main chain of the precursor structure is essentially superimposable with that of the autocleaved enzyme (rmsd = 0.53 Å for 273 residues). Both the precursor and the autocleaved enzyme form a four-layer  $\alpha$ - $\beta$ - $\beta$ - $\alpha$  structure (Figure 2), with two  $\beta$  sheets packed against each other to form a core that is "sandwiched" by two layers of  $\alpha$  helices. The interlayer loops cluster at one side of the structure to provide functional groups for the catalytic centers of autoproteolysis and hydrolase activity, whereas the intralayer loops are located on the other side of the structure. The most significant differences in the active sites between structures of the precursor and the autocleaved enzyme are in Thr-152 and Thr-203, where main chain atoms have shifted by more than 1 Å, with the nucleophile oxygen moved 1.7 Å. In the crystals, the GA precursor forms a dimeric structure

that is similar to the structure of the autocleaved enzyme crystallized in different crystal forms (Guo et al., 1998; Xuan et al., 1998). Such a dimer structure has also been observed in GA enzymes of other species either in solution (Kartinen et al., 1991; Tollersrud and Aronson, 1992) or in crystal (Oinonen et al., 1995).

#### A Partially Formed Substrate Site and Glycine Binding

There is a deep cavity in the precursor located at one side of the  $\alpha$ - $\beta$ - $\beta$ - $\alpha$  sandwich (as viewed in the direction of dotted arrows in Figure 2). This cavity is adjacent to the autocleavage site and is a partial substrate site of the autocleaved enzyme. This cavity is formed from the conserved residues Arg-180, Asp-183, and Gly-204 that interact with the aspartate moiety of the GA hydrolase

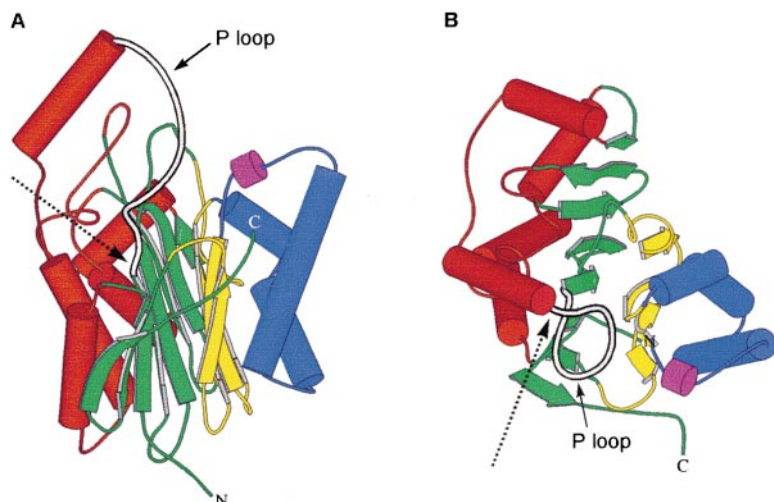


Figure 2. Overall Structure of the GA Precursor

The precursor consists of four layers of secondary structure elements: from the left, a five-helix outer layer (red), an eight-stranded  $\beta$  sheet (green), a four-stranded antiparallel  $\beta$  sheet (yellow), and an outer layer of three  $\alpha$  helices (blue) plus a  $3_{10}$  helix (purple). The solid arrows mark the precursor loop (P loop) that contains the scissile peptide bond, whereas the dotted arrows point toward the autoproteolytic center. The view in (B) is related to that in (A) by a rotation of  $90^\circ$  along the horizontal axis. Figures were made with the program MOLSCRIPT (Kraulis, 1991).

substrates (Oinonen et al., 1995; Guo et al., 1998). This cavity is solvent accessible through a small port (arrow in Figure 3A). In the precursor/inhibitor complexes, a glycine molecule binds in this cavity through the same hydrogen bond network as in the enzyme/product interactions (Oinonen et al., 1995): the  $\alpha$ -amino group of glycine makes hydrogen bonds with Asp-183 and Gly-204, and the  $\alpha$ -carboxylate of glycine forms a salt bridge with Arg-180. In the structure of the T152A precursor, where no glycine is present, this cavity is filled with solvent (possibly a cryoprotectant glycerol molecule).

#### The Autoproteolytic Center of a T152C Precursor

The autoproteolytic site is shielded by the P loop, next to the partially formed substrate cavity. Autoproteolysis results in structural rearrangements, that is, release/disorder of the P loop such that the substrate site is opened (Figure 3B) and the newly generated amino terminus is exposed for hydrolase activity. The autoproteolytic site is only accessible to solvent through a narrow port (arrow in Figure 3A). In this site, the scissile peptide bond is located at the junction of the P loop and a  $\beta$  strand, next to the partial substrate site occupied by a glycine molecule in the precursor/inhibitor complexes. Figure 4A shows the electron density map at the cleavage site of the T152C precursor. The backbone at the

scissile peptide bond makes a tight turn with the  $\alpha$  carbons of residues 150 and 153 less than 8 Å apart. One carboxylate oxygen of Asp-151 loops back via a six-member ring to form a hydrogen bond with the amide nitrogen of the scissile peptide bond, and the second carboxylate oxygen of Asp-151 forms a hydrogen bond with the hydroxyl oxygen of Thr-203. Unlike most major types of tight turns, there is no backbone hydrogen bond between the carbonyl oxygen of residue 150 and the amide nitrogen of residue 153 to stabilize the turn, and no Gly or Pro is involved. This unusual turn apparently creates conformational strain on the backbone, in that the planarity of the peptide bond ( $\omega$  value) near the scissile peptide bond consistently refined to a value that deviates significantly from ideality ( $180^\circ$  for *trans*). Residue 152 is the only residue with  $\omega$  values that deviated more than 3 standard deviations (more than  $20^\circ$ ) from ideality in both molecules of the asymmetric unit, in all three structures reported here (Figures 5A and 5B). In addition, the backbone angles N-C $\alpha$ -C' ( $\tau$ ) near the scissile peptide bond (e.g., Asp-151) deviate from ideality ( $110^\circ$ ) by more than 2 standard deviations ( $\sim 9^\circ$  for each residue). Together, these main chain distortions can raise the energy by more than 5 kcal/mol for each distorted residue (Balasubramanian and Ramakrishnan, 1972). Refinement of structure with much more tightly

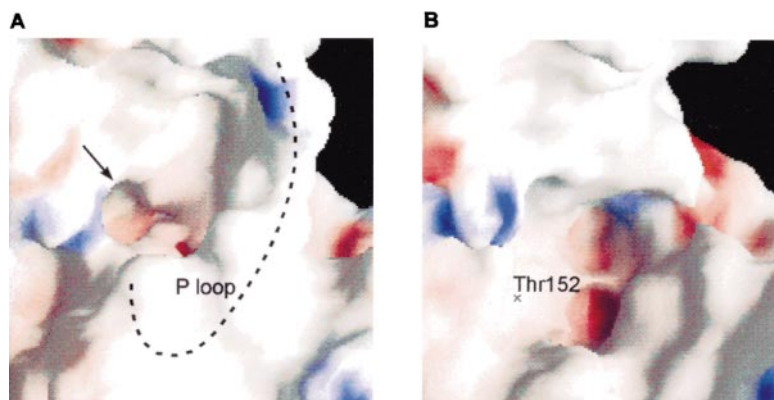


Figure 3. A Surface Cavity in the Precursor Adjacent to the Autoproteolysis Site Is Opened to Form the Substrate Site of the Autocleaved Hydrolase

The solvent-accessible surfaces (Nicholls et al., 1991) of the precursor (A) and the autocleaved hydrolase (B), colored according to electrostatic potential (blue, positively charged; red, negatively charged), reveal an enlargement of the surface cavity after autocleavage. The view toward the autoproteolytic center of the precursor is along the dotted arrow in Figure 2. The small port (black arrow in [A]) is approximately 4 Å wide. Dotted line traces the C $\alpha$  carbons of the P loop.

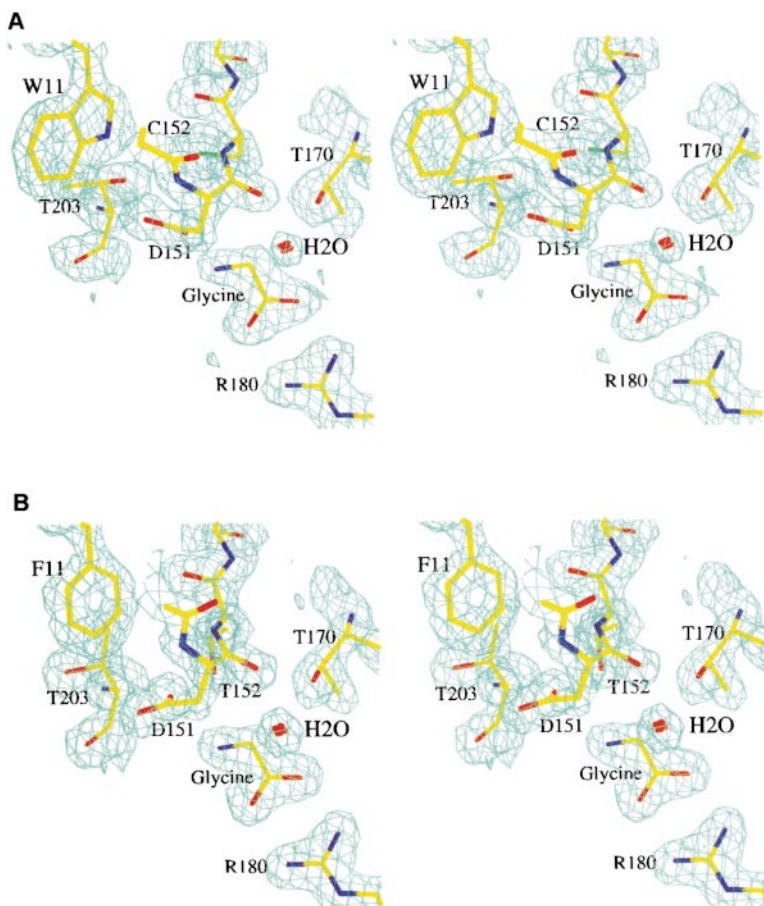


Figure 4. Stereo View of the Electron Density Maps at the Autoproteolytic Sites

The maps are of the  $2F_o - F_c$  type at 1.9 Å resolution and contoured at  $1.3 \sigma$  for the T152C (A) and W11F (B) precursors. The final model of the active site residues are shown by atom type: yellow for carbons, blue for nitrogens, red for oxygens, and green for sulfur. Also shown is a glycine inhibitor.

restrained stereochemical geometry leads to a poorer model (assessed by R factors and the difference maps), indicating that the backbone distortion is well supported by the experimental data. Furthermore, when the restraints in this region are relaxed, a model with a more distorted backbone gives similar R factors, suggesting that the main chain distortion could be underestimated when using the “ideal” weight between experimental data and empirical energy function in the refinement (Brunger, 1992a). This unusual peptide conformation does not result from the binding of the glycine inhibitor, since in the absence of glycine the structure of the T152A precursor also reveals a strained tight turn. In the T152C precursor, the binding of glycine locks the thiol group in an inactive side chain rotamer, pointing away from the carbonyl carbon of the scissile peptide bond. This appears to be due to a favorable packing of the thiol group into a small pocket formed between side chains of Cys-168 and Thr-203 and main chain atoms of a  $\beta$  strand. A similar inactive conformation has also been observed in the T152C autocleaved enzyme (Guo et al., 1998) and in the glutaminase domain of glucosamine 6-phosphate synthase where Cys-1 is the wild-type N-terminal nucleophile (Isupov et al., 1996). Along with the nucleophile Cys-152, several conserved residues are aligned at the autoproteolytic center (orange bonds in Figure 5C). The  $N_\epsilon$  atom of Trp-11 is hydrogen bonded to the hydroxyl oxygen of Thr-203, which in turn forms a hydrogen bond with a carboxylate oxygen of Asp-151

(Figure 6). On the other side of the autocleavage site, the carbonyl of Asp-151 forms hydrogen bonds with the hydroxyl oxygen of Thr-170 and a water molecule.

#### The Autoproteolytic Center of a W11F Precursor

The structure of a second autoproteolysis active precursor, W11F, also reveals a highly strained tight turn at the scissile peptide bond. An electron density map from the cleavage site of the W11F precursor is shown in Figure 4B. As a result of the tight turn, the side chains of residues Asp-151 and Thr-152 are in close contact, such that one side chain carboxylate oxygen of the Asp-151 forms a hydrogen bond with the hydroxyl oxygen of Thr-152 to enhance its nucleophilicity. Furthermore, the hydroxyl group of Thr-152 is oriented toward the scissile peptide bond in an orientation suitable for a nucleophilic attack at the carbonyl carbon of Asp-151. The hydroxyl oxygen of Thr-152 is 3.1 Å from the carbonyl carbon of Asp-151. This active conformation of the nucleophile is apparently held by a hydrogen bond with the  $O_\gamma$  of Thr-170. On the other hand, substitution of Trp-11 with a Phe results in the loss of a hydrogen bond with Thr-203, which in turn changes its side chain conformation to the most popular rotamer (+ rotamer; Ponder and Richards, 1987), with its hydroxyl group steering away from the active center. As a result, the hydrogen bond between Thr-203 and Asp-151 is broken, and both Asp-151 and Thr-152 are displaced slightly

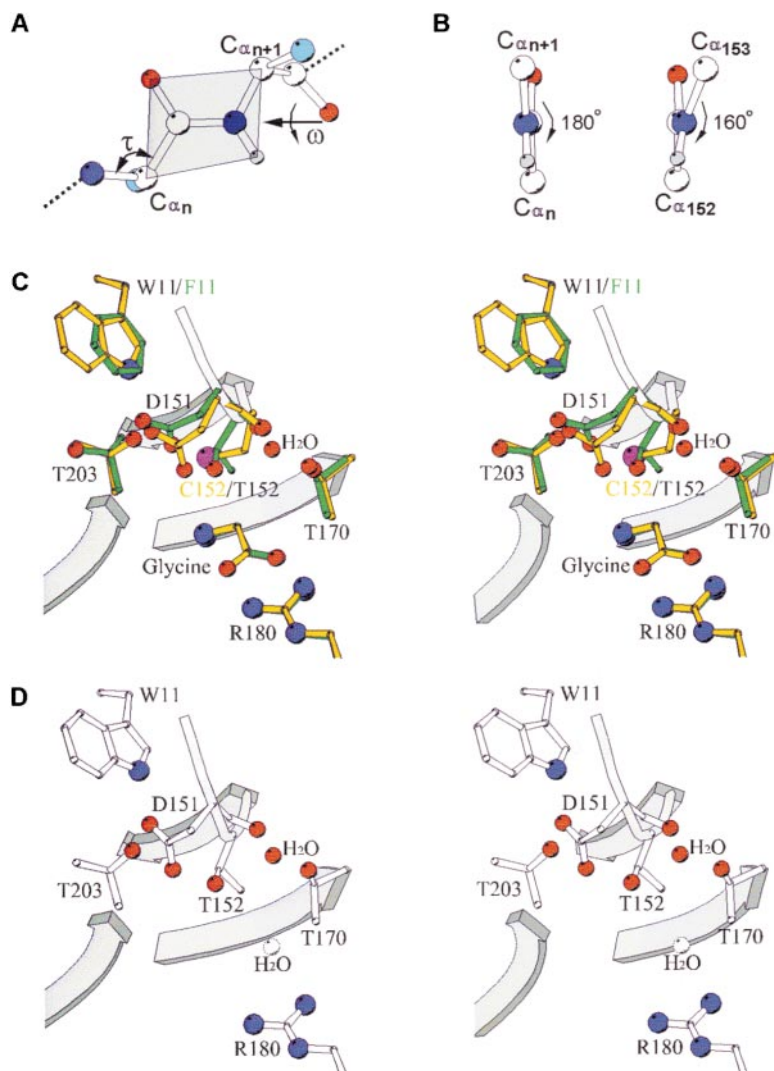


Figure 5. Atomic Structures of the Autoproteolytic Sites

(A) Part of a polypeptide chain that is linked by a peptide bond (planar amide group), represented as a shaded block. The angle of rotation around the  $C'_n-N_{n+1}$  bond is called  $\omega$ , and the backbone angle  $N-C\alpha-C'$  is called  $\tau$ . Atoms are shown by atom type: gray for hydrogens, white for carbons, blue for nitrogens, red for oxygens, and cyan for side chains.

(B) Comparison of  $\omega$  angles between an ideal *trans* peptide bond (left) and the distorted *trans* peptide bond in the GA precursors (right). The view is looking down on the  $C'_n$  from the direction of the  $N_{n+1}$  atom, as depicted by an arrow in (A).

(C) Stereo view of superimposition of the active site residues in T152C (orange bonds) and W11F (green bonds) precursors. Also shown are the inhibitor glycine molecules. Only side chain nonhydrogen atoms are shown, with the exception of the carbonyl group of Asp-151. Atoms are shown by atom type: orange and green for carbons of T152C and W11F, respectively, blue for nitrogens, red for oxygens, and purple for sulfur.

(D) The same stereo view of the active site model of the wild-type precursor. The color scheme for atoms is the same as in (C), except carbons are white. The inhibitor molecule has been replaced with a water molecule (open circle). Figures were made with the program MOLSCRIPT (Kraulis, 1991).

away from the glycine-binding site (green bonds in Figure 5C) relative to those in the T152C precursor (displaced by  $\sim 1$  Å; the rmsd of all other active site residues between the two precursor structures is 0.20 Å). It is of interest to note that autoproteolysis of the wild-type precursor is not inhibited by glycine, suggesting that in the wild-type precursor, the partially formed substrate cavity is smaller or less flexible than that of the W11F precursor. Thus, in the wild-type precursor, the two residues bracketing the scissile peptide bond would have conformations closer to those in the T152C precursor than in the W11F precursor. The rest of the structure is essentially identical to that of the T152C precursor (rmsd = 0.22 Å for 278 residues; Figure 5C).

#### Reversible Inhibition by Glycine

As discussed above, glycine inhibits autoproteolysis of the T152C precursor by locking its nucleophile in an inactive conformation. Furthermore, in the T152C precursor, the  $\alpha$ -amino group of glycine also makes a hydrogen bond with the side chain of Asp-151 (Figure 5C); a protonated amino group would make Asp-151 less

basic. This mode of inhibition by glycine could also occur in the T152S (a Thr to Ser mutation) precursor. However, as pointed out above, glycine does not inhibit autoproteolysis of the wild-type enzyme. Nature has thus selected a Thr as the nucleophile for GA to prevent inhibition by small amino acids that are abundant in cells. It is of interest to note that proteasomes also use a Thr as the nucleophile even though replacement by a Ser in the archaeobacterial proteasome still allows full proteolytic activity (Seemüller et al., 1995). It is possible that Thr has also been conserved in the active sites of proteasomes, from bacteria to human, to ensure enzymatic activation by autoproteolysis. In the case of the W11F precursor, Thr-152 is in an appropriate orientation to attack the scissile peptide bond even in the presence of the inhibitor glycine. Small but significant reorganizations were observed on residues 151 and 152 in the W11F structure (Figure 5C). These geometric perturbations could account for the 600-fold reduction in the autoproteolysis rate of the W11F precursor. In an analogous case with isocitrate dehydrogenase (Mesecar et al., 1997), small changes in distance ( $< 1.55$  Å) and orientation of reacting groups results in a large reduction

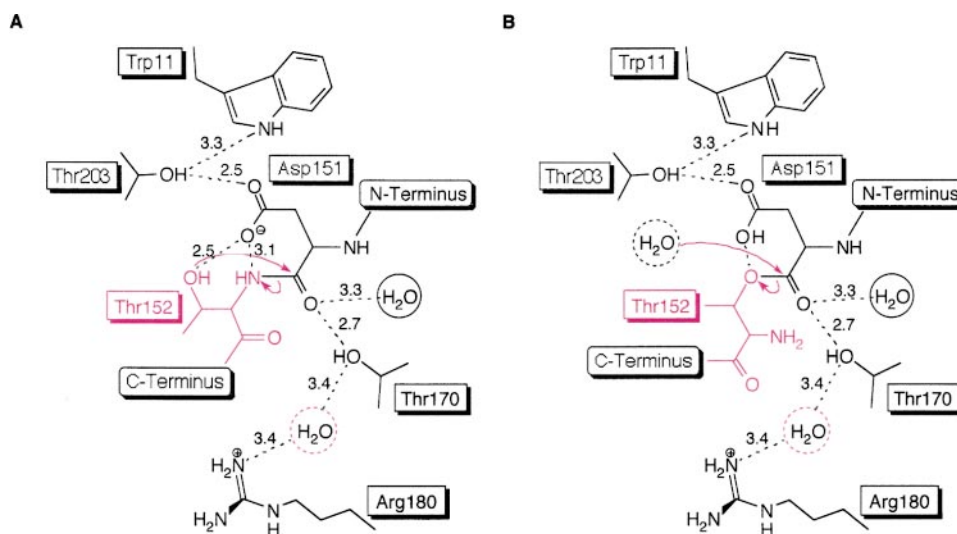


Figure 6. A Proposed Two-Step Autoproteolysis Mechanism for the Wild-Type Precursor

Dashed lines indicate hydrogen-bonding interactions with distances in angstroms. The model is based on the refined structure of the W11F precursor, with some modification (see the text). A water molecule (dotted red circle) has been modeled after removal of the glycine inhibitor. Red arrows indicate nucleophilic attacks and the routes of leaving groups.

(A) The first step of an N  $\rightarrow$  O acyl shift, resulting in an ester linkage.

(B) The second step of hydrolysis of the ester linkage, finishing the cleavage of the precursor protein with a neighboring water molecule (dotted black circle).

( $10^{-3}$  to  $10^{-5}$ ) in the reaction rate. But questions remain concerning the mechanism of glycine inhibition of autoproteolysis of the W11F precursor (Guan et al., 1998). In addition to locking residues 151 and 152 to a nonproductive geometry, we propose that glycine could also replace solvent molecules that bridge and relay the positive charge from Arg-180 to the proposed oxyanion hole of Thr-170 (see below). In line with this proposal, mutation of Arg-180, which is more than 6 Å away from other residues of the autoproteolytic center, dramatically reduces autoproteolysis (Guan et al., 1998; Liu et al., 1998).

#### An Atomic Model of the Wild-Type Precursor

Based on the structures of the two active precursors W11F and T152C, an atomic model for the wild-type precursor has been constructed (Figure 5D). This model was mainly derived from the W11F structure because of the higher quality of its diffraction data and refined structure (see Experimental Procedures and Table 2), with some modifications. As discussed above, the conformations of Asp-151 and Thr-152 in the wild-type precursor are likely to be closer to those in the T152C structure. We thus rebuilt these two residues based on the T152C structure, but the side chain orientation of Thr-152 is modeled to reflect the active conformation in the W11F structure. In addition, side chain conformations of Trp-11 and Thr-203 were rebuilt to restore the hydrogen bond network in the T152C structure. The rebuilt model was further subjected to energy minimization (see Experimental Procedures).

#### Autoproteolytic Mechanism

With the atomic model of the wild-type GA precursor, we propose a molecular mechanism of autoproteolysis that utilizes an N  $\rightarrow$  O acyl shift (Figure 6). First, Trp-11

holds the side chain of Thr-203 into a less popular rotamer conformation, such that it forms a hydrogen bond with a carboxylate oxygen of the Asp-151 side chain. As a result, the second carboxylate oxygen of Asp-151 is positioned to act as a general base and deprotonate the hydroxyl oxygen of Thr-152 to enhance its nucleophilicity. In vivo, human GA precursor autoproteolyzes in the endoplasmic reticulum, where the equivalent to Asp-151 is likely to be mainly in its basic form (Ikonen et al., 1993). Determination of the  $pK_a$  value of Asp-151 in the precursor is underway. Second, a nucleophilic attack on the carbonyl carbon of Asp-151 by the hydroxyl oxygen of Thr-152 would result in a tetrahedral intermediate with a negatively charged carbonyl oxygen stabilized by an oxyanion hole. The oxyanion hole is formed by a bound water molecule and the side chain hydroxyl group of Thr-170 that is positively polarized by another solvent molecule interacting with Arg-180. This oxyanion hole is different from the one that functions in GA hydrolase activity and for which Thr-203 and Gly-204 form the putative oxyanion hole (Guo et al., 1998). Third, collapse of the tetrahedral intermediate shifts the linkage from an amide bond to an ester bond (N  $\rightarrow$  O acyl shift). In this step, the protonated side chain of Asp-151 could aid the leaving amino group. The strained tight turn observed in the peptide backbone near the scissile peptide bond would drive the equilibrium toward the N  $\rightarrow$  O shift. Fourth, a neighboring water molecule nucleophilically attacks the same carbonyl carbon to form another tetrahedral transition state that collapses to resolve into the  $\alpha$  and  $\beta$  subunits, which constitute an active GA hydrolase. Mutagenesis studies have confirmed the critical roles of Trp-11, Asp-151, Thr-152, Thr-170, Arg-180, and Thr-203 in autoproteolysis (Guan et al., 1998; Liu et al., 1998). His-150 is important but not

absolutely required for autoproteolysis (Guan et al., 1998; Saarela et al., 1998). This could be due to an optimal packing of the imidazole ring of histidine against a trough between Trp-11 and Phe-13 that helps in maintaining the precise geometry of the scissile peptide bond. This represents a subtle structural factor that has also been implicated in driving protein splicing of *Saccharomyces cerevisiae* VMA intein (Chong et al., 1998).

It is generally conceived that a *cis* peptide bond has a higher energy than the *trans* isomer. However, since peptide bonds are biochemically synthesized in a *trans* configuration, the existence of *cis* peptide bonds in crystal structures indicates that distorted *trans* isomers at these positions would be even more energetically unfavorable. Indeed, the autoproteolytic site of GA precursors reveals a highly strained and distorted *trans* configuration that may generate the potential for autocleavage. A spatial motif composed of the strained loop (GA residues 149 to 154) was used to search similar loop conformations in the Protein Data Bank database (Kleywegt, 1999). The best matches were residues 57 to 62 of endocellulase E1 (Sakon et al., 1996; PDB code 1ece; rmsd = 0.95 Å) and residues 35 to 40 of bacterial luciferase  $\beta$  subunit (Fisher et al., 1996; PDB code 1luc; rmsd = 0.95 Å). However, no backbone distortion was found in these loops. It is unclear why such an unfavorable conformation in the GA precursor does not collapse into a presumably more favorable *cis* isomer. One possibility is that this distorted configuration was created after the protein had folded, for example, through a dimerization of precursors at their extensive hydrophobic interfaces (Guo et al., 1998). Precursor dimerization has been implicated as an important factor for the autoproteolysis of GA (Guo et al., 1998; Saarela et al., 1998). Similarly, subunit interactions also appear to be a factor for the autoprocessing and  $\beta$  elimination of aspartate decarboxylase (Albert et al., 1998). Although prolyl isomerase offers some possibility for the *cis/trans* isomerization of proline imidic peptide bonds, this enzyme does not appear to act efficiently on nonprolyl amide bonds or on folded proteins (Lang et al., 1987). In the absence of the catalytic isomerization of nonprolyl amide bonds at physiological temperatures, the biochemically synthesized *trans* configuration may be determined within the folded protein. In other cases, the conformational strain generated by an unusual  $\beta$ -fold framework also appears to be a factor for the autocatalytic cleavage of the pyruvoyl-dependent histidine decarboxylase (Gallagher et al., 1993) or for protein splicing of the GyrA intein (Klabunde et al., 1998). Nonetheless, a single alanine extension within the structure of GyrA intein might have a lower energy barrier than the native precursor to allow the relaxation of a strained backbone into a *cis* peptide configuration.

However, more than a strained backbone is required for autoproteolysis. Many mutants of GA, including the T152A mutant, with a similarly strained backbone conformation do not autoproteolyze (Guan et al., 1998). In addition to the nucleophile, other functional groups must be precisely aligned to allow autoproteolysis to proceed via an  $N \rightarrow O$  shift with ester formation. Formation of an ester requires deprotonation of the nucleophilic hydroxyl group, stabilization of the transient tetrahedral intermediate by an oxyanion hole, and protonation of

the leaving amino group. In general, the equilibrium favors the amide bond over the ester form of Thr, Ser, or Cys. The equilibrium, however, can be shifted toward the ester when its formation may relieve some backbone conformational strain such as that observed in the GA precursor structures. The equilibrium can also be shifted toward ester formation by the presence of a general acid-base catalyst, such as Asp-151, which may rapidly protonate the leaving amino group.

Although a similar mechanism is likely to occur in other self-catalyzed peptide bond rearrangements via  $N \rightarrow O$  or  $N \rightarrow S$  acyl shifts, some differences are noted among these systems. Depending on their individual structures, either Thr, Ser, or Cys functions as the nucleophile. Nature has selected Thr as the nucleophile for GA to eliminate a nonproductive side chain rotamer conformation and to prevent inhibition of autoproteolysis by small amino acids in cells. In contrast, Cys is the preferred nucleophile in many other cases (Perler et al., 1997). This could be due to the lower  $pK_a$  of the thiol group of the Cys and presumably because of the lack of a general base residue in these structures to assist in the deprotonation of the nucleophile. Also unlike the GA precursors, the conserved residue immediately upstream of the autoproteolytic bond is Gly in proteasomes and some inteins, indicating a different mechanism for the activation of the nucleophile. In addition, Ntn hydrolases have an  $\alpha$ - $\beta$ - $\beta$ - $\alpha$  sandwich structural fold that is different from the all- $\beta$  structure of Hedgehog proteins (Hall et al., 1997) or intein-splicing domains (Duan et al., 1997; Klabunde et al., 1998). A detailed knowledge of the individual high-resolution structures around the autoprocessing site is necessary to understand these differences.

## Conclusion

This study of GA precursor provides a structure of precursor proteins that are active in self-catalyzed peptide bond rearrangements. The autoproteolytic potential originates at least in part from a highly strained, tight turn conformation at the scissile peptide bond. In addition, conserved residues are properly aligned as the general acid-base and oxyanion hole that are critical to facilitate the  $N \rightarrow O$  or  $N \rightarrow S$  acyl shift or to stabilize the reaction intermediate. These principles and mechanisms are likely to be found in other systems of protein processing through the  $N \rightarrow O$  or  $N \rightarrow S$  acyl shift. Thus, the structure of GA precursor will guide additional studies of all the self-catalyzed protein processing pathways.

## Experimental Procedures

### Kinetics of Autoproteolysis

In vitro autoproteolysis was analyzed by SDS-PAGE at various time points. The overall rate constants of autoproteolysis were determined as described previously (Guan et al., 1998).

### Crystal Preparation and Data Collection

Protein overexpression, purification, and crystallization of *Flavobacterium* GA precursors are described elsewhere (Cui et al., 1999). The crystals form in space group P1 with  $a = 46.3$  Å,  $b = 52.8$  Å,  $c = 62.4$  Å,  $\alpha = 80.8^\circ$ ,  $\beta = 90.5^\circ$ , and  $\gamma = 105.1^\circ$  and contain two molecules in the asymmetric unit. Heavy atom derivatives were obtained by soaking the native T152A crystals in the crystallization buffer supplemented with 10% glycerol and 1 mM mercuric acetate

for 17 hr, before transferring the crystals to a final glycerol concentration of 20% as cryoprotectant.

Diffraction data were collected with crystals flash cooled at 100 K using either a laboratory X-ray source (Rigaku generator with Raxis II detector) or synchrotron X-radiation (beamline X-4a, the National Synchrotron Light Source, Brookhaven National Laboratory). All intensity data were processed and scaled using the programs DENZO and SCALEPACK (Otwinowski and Minor, 1997) and converted to structure factors using TRUNCATE from the CCP4 software package (CCP4, 1994).

#### Structure Determination and Refinement

The structures were determined by a combination of single isomorphous replacement with anomalous scattering (SIRAS) and molecular replacement (MR). Heavy atom positions were initially obtained from isomorphous difference Patterson maps calculated in XTALVIEW (McRee, 1992). The heavy atom parameters were then refined by MLPHARE from the CCP4 package (CCP4, 1994). To determine an atomic model of the precursor free of MR model bias, starting models of the autocleaved GA (Guo et al., 1998) with one domain omitted at a time (in both noncrystallographic symmetry, NCS, related molecules, ~20% of the structure) were used in the MR phasing. These partial models were further subjected to simulated annealing before being used to calculate MR phases and then combined with the SIRAS phases. At this stage, the overall figure of merit was 0.767. The combined phases were further improved by NCS averaging, solvent flattening, and histogram matching using the program DM (CCP4, 1994) with an overall figure of merit of 0.867. This real-space averaging yielded high-quality electron density maps, and the model for the omitted domain was then built into the density. This procedure, with one domain omitted at a time, was repeated throughout the entire structure.

The first maps were calculated at 2.5 Å resolution for model building in the program O (Jones et al., 1991). The first 272 out of 295 residues were built based on the DM-modified maps. Automated refinement included rigid body, overall temperature factor, and positional and restrained atomic temperature factor refinement, as well as simulated annealing using a slow-cooling protocol in X-PLOR (Brünger, 1992a). After the first round of manual rebuilding, the structure was used for independent model rebuilding and refinement of all three structures (T152A, T152C, and W11F). Initially, strict NCS constraints were applied, and in later stages of refinement, tight NCS restraints were applied, exclusive of residues that were involved in crystal contacts. After a few rounds of model rebuilding, stepwise resolution extension, and automated refinement, clear electron density could be seen for all residues in the final model. In the W11F precursor structure, which has the best quality of diffraction data, continuous electron density for the main chain and most of the side chains of the precursor loop (P loop) was evident, with one break in the chain between residues 137 and 140, despite the exposed nature of the loop and the relatively high B factors (main chain average from residue 141 to 151 is 20.5 Å<sup>2</sup>; overall main chain average is 7.8 Å<sup>2</sup>). Refinement protocols were aimed at decreasing the  $R_{\text{free}}$  (Brünger, 1992b) rather than the conventional  $R_{\text{cryst}}$  to avoid errors introduced by overfitting of the data. When the  $R_{\text{free}}$  appeared to have reached a minimum at the final resolution, water molecules were added and the structure subjected to another round of positional refinement and manual rebuilding. The statistics of the final structures are shown in Table 2, with an rmsd of 0.17–0.38 Å for main chain atoms between crystallographically independent molecules. The modeled structure of the wild-type precursor was refined with ten cycles of conjugate gradient energy minimization without the crystallographic residual term (Brünger, 1992a).

#### Structural Comparisons

All superimpositions of different structures were performed using LSQKAB (CCP4, 1994). For Figure 5C, common atoms of residues surrounding the scissile peptide bond are superimposed (Trp-11, Thr-170, Arg-180, Asp-183, Thr-203, and Gly-204). Spatial motif searches were performed using SPASM (Kleywegt, 1999).

#### Acknowledgments

The authors thank C. Ogata and D. Cook for aid at beamline X4a of the National Synchrotron Light Source (Brookhaven National Lab-

oratory); C. W. Akey, D. Atkinson, and G. G. Shipley for critical reading of the manuscript; and members of the lab for useful comments. H.-C. G. is grateful to D. M. Small for his continuous support of this study. This research was supported by a grant from the National Institutes of Health (DK53893).

Received May 28, 1999; revised August 3, 1999.

#### References

- Albert, A., Dhanaraj, V., Genschel, U., Khan, G., Ramjee, M.K., Pulido, R., Sibanda, B.L., von Delft, F., Witty, M., Blundell, T.L., et al. (1998). Crystal structure of aspartate decarboxylase at 2.2 Å resolution provides evidence for an ester in protein self-processing. *Nat. Struct. Biol.* 5, 289–293.
- Aronson, N.N., Jr. (1996). Lysosomal glycosylasparaginase: a member of a family of amidases that employ a processed N-terminal threonine, serine or cysteine as a combined base-nucleophile catalyst. *Glycobiology* 6, 669–675.
- Balasubramanian, R., and Ramakrishnan, C. (1972). Stereochemical criteria for polypeptide and protein chain conformations. VIII. Energy maps for a pair of non-planar peptide units having distortion of bond angle at the  $\alpha$ -carbon atom. *Int. J. Peptide Protein Res.* 4, 91–99.
- Baumeister, W., Walz, J., Zühl, F., and Seemüller, E. (1998). The proteasome: paradigm of a self-compartmentalizing protease. *Cell* 92, 367–380.
- Brannigan, J.A., Dodson, G., Duggleby, H.J., Moody, P.C.E., Smith, J.L., Tomchick, D.R., and Murzin, A.G. (1995). A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. *Nature* 378, 416–419.
- Brünger, A.T. (1992a). X-PLOR Version 3.1. A System for X-Ray Crystallography and NMR (New Haven, CT: Yale University Press).
- Brünger, A.T. (1992b). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355, 472–475.
- CCP4 (Collaborative Computational Project Number 4). (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* 50, 760–763.
- Chan, Y.-M., and Jan, Y.N. (1998). Roles for proteolysis and trafficking in Notch maturation and signal transduction. *Cell* 94, 423–426.
- Chong, S., Williams, K.S., Wotkowicz, C., and Xu, M.-Q. (1998). Modulation of protein splicing of the *Saccharomyces cerevisiae* vacuolar membrane ATPase intein. *J. Biol. Chem.* 273, 10567–10577.
- Cocks, T.M., Fong, B., Chow, J.M., Anderson, G.P., Frauman, A.G., Goldie, R.G., Henry, P.J., Carr, M.J., Hamilton, J.R., and Moffatt, J.D. (1999). A protective role for protease-activated receptors in the airways. *Nature* 398, 156–160.
- Cui, T., Liao, P.-H., Guan, C., and Guo, H.-C. (1999). Purification and crystallization of precursors and autoprocessed enzymes of *Flavobacterium* glycosylasparaginase: an N-terminal nucleophile hydrolase. *Acta Crystallogr. D*, in press.
- Davie, E.W. (1995). Biochemical and molecular aspects of the coagulation cascade. *Thromb. Haemost.* 74, 1–6.
- Ditzel, L., Huber, R., Mann, K., Heinemeyer, W., Wolf, D.H., and Groll, M. (1998). Conformational constraints for protein self-cleavage in the proteasome. *J. Mol. Biol.* 279, 1187–1191.
- Duan, X., Gimble, F.S., and Quioco, F.A. (1997). Crystal structure of PI-Scel, a homing endonuclease with protein splicing activity. *Cell* 89, 555–564.
- Duggleby, H.J., Tolley, S.P., Hill, C.P., Dodson, E.J., Dodson, G., and Moody, P.C.E. (1995). Penicillin acylase has a single-amino-acid catalytic centre. *Nature* 373, 264–268.
- Fisher, A.J., Thompson, T.B., Thoden, J.B., Baldwin, T.O., and Raymond, I. (1996). The 1.5-Å resolution crystal structure of bacterial luciferase in low salt conditions. *J. Biol. Chem.* 271, 21956–21968.

- Gallagher, T., Rozwarski, D.A., Ernst, S.R., and Hackert, M.L. (1993). Refined structure of the pyruvoyl-dependent histidine decarboxylase from *Lactobacillus* 30a. *J. Mol. Biol.* *230*, 516–528.
- Groll, M., Ditzel, L., Lowe, J., Stock, D., Bochtler, M., Bartunik, H.D., and Huber, R. (1997). Structure of 20S proteasome from yeast at 2.4 Å resolution. *Nature* *386*, 463–471.
- Guan, C., Cui, T., Rao, V., Liao, W., Benner, J., Lin, C.-L., and Comb, D. (1996). Activation of glycosylasparaginase: formation of active N-terminal threonine by intramolecular autoprolysis. *J. Biol. Chem.* *271*, 1732–1737.
- Guan, C., Liu, Y., Shao, Y., Cui, T., Liao, W., Ewel, A., Whitaker, R., and Paulus, H. (1998). Characterization and functional analysis of the cis-autoprolysis active center of glycosylasparaginase. *J. Biol. Chem.* *273*, 9695–9702.
- Guo, H.-C., Xu, Q., Buckley, D., and Guan, C. (1998). Crystal structures of *Flavobacterium* glycosylasparaginase: an N-terminal nucleophile hydrolase activated by intramolecular proteolysis. *J. Biol. Chem.* *273*, 20205–20212.
- Hall, T.M.T., Porter, J.A., Young, K.E., Koonin, E.V., Beachy, P.A., and Leahy, D.J. (1997). Crystal structure of a Hedgehog autoprocessing domain: homology between Hedgehog and self-splicing proteins. *Cell* *91*, 85–97.
- Ikonen, E., Julkunen, I., Tollersrud, O.-K., Kalkkinen, N., and Peltonen, L. (1993). Lysosomal aspartylglucosaminidase is processed to the active subunit complex in the endoplasmic reticulum. *EMBO J.* *12*, 295–302.
- Isupov, M.N., Obmolova, G., Butterworth, S., Badet-Denisot, M.-A., Badet, B., Polikarpov, I., Littlechild, J.A., and Teplyakov, A. (1996). Substrate binding is required for assembly of the active conformation of the catalytic site in Ntn amidotransferase: evidence from the 1.8 Å crystal structure of the glutaminase domain of glucosamine 6-phosphate synthase. *Structure* *4*, 801–810.
- Jones, T.A., Zou, J.-Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron-density maps and the location of errors in these models. *Acta Crystallogr. A* *47*, 110–119.
- Kaartinen, V., Williams, J.C., Tomich, J., Yates, J.R., III, Hood, L.E., and Mononen, I. (1991). Glycosylasparaginase from human leukocytes: inactivation and covalent modification with diazo-oxonorvaline. *J. Biol. Chem.* *266*, 5860–5869.
- Khan, A.R., and James, M.N.G. (1998). Molecular mechanisms for the conversion of zymogens to activate proteolytic enzymes. *Protein Sci.* *7*, 815–836.
- King, R.W., Deshaies, R.J., Peters, J.-M., and Kirschner, M.W. (1996). How proteolysis drives the cell cycle. *Science* *274*, 1652–1659.
- Klabunde, T., Sharma, S., Telenti, A., Jacobs, W.R., Jr., and Sacchettini, J.C. (1998). Crystal structure of GyrA intein from *Mycobacterium xenopi* reveals structural basis of protein splicing. *Nat. Struct. Biol.* *5*, 31–36.
- Klenk, H.-D., and Garten, W. (1994). Host cell proteases controlling virus pathogenicity. *Trends Microbiol.* *2*, 39–43.
- Kleywegt, G.J. (1999). Recognition of spatial motifs in protein structures. *J. Mol. Biol.* *285*, 1887–1897.
- Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* *24*, 946–950.
- Lang, K., Schmid, F.X., and Fischer, G. (1987). Catalysis of protein folding by prolyl isomerase. *Nature* *329*, 268–270.
- Lee, J.J., Ekker, S.C., von Kessler, D.P., Porter, J.A., Sun, B.I., and Beachy, P.A. (1994). Autoprolysis in *hedgehog* protein biogenesis. *Science* *266*, 1528–1537.
- Liu, Y., Guan, C., and Aronson, N.N., Jr. (1998). Site-directed mutagenesis of essential residues involved in the mechanism of bacterial glycosylasparaginase. *J. Biol. Chem.* *273*, 9688–9694.
- Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., and Huber, R. (1995). Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. *Science* *268*, 533–539.
- McRee, D.E. (1992). A visual protein crystallographic software system for X11/XView. *J. Mol. Graph.* *10*, 44–46.
- Mesecar, A.D., Stoddard, B.L., and Koshland, D.E., Jr. (1997). Orbital steering in the catalytic power of enzymes: small structural changes with large catalytic consequences. *Science* *277*, 202–206.
- Mononen, I., Fisher, K.J., Kaartinen, V., and Aronson, N.N., Jr. (1993). Aspartylglucosaminuria: protein chemistry and molecular biology of the most common lysosomal storage disorder of glycoprotein degradation. *FASEB J.* *7*, 1247–1256.
- Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* *11*, 281–293.
- Oinonen, C., Tikkanen, R., Rouvinen, J., and Peltonen, L. (1995). Three-dimensional structure of human lysosomal aspartylglucosaminidase. *Nat. Struct. Biol.* *2*, 1102–1108.
- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* *276*, 307–326.
- Paulus, H. (1998). The chemical basis of protein splicing. *Chem. Soc. Rev.* *27*, 375–386.
- Perler, F.B. (1998a). Protein splicing of inteins and Hedgehog autoproteolysis: structure, function, and evolution. *Cell* *92*, 1–4.
- Perler, F.B. (1998b). Breaking up is easy with esters. *Nat. Struct. Biol.* *5*, 249–252.
- Perler, F.B., Olsen, G.J., and Adam, E. (1997). Compilation and analysis of intein sequences. *Nucleic Acids Res.* *25*, 1087–1093.
- Ponder, J.W., and Richards, F.M. (1987). Tertiary templates for proteins: use of packing criteria in the enumeration of allowed sequences for different structural classes. *J. Mol. Biol.* *193*, 775–791.
- Porter, J.A., Young, K.E., and Beachy, P.A. (1996). Cholesterol modification of Hedgehog signaling proteins in animal development. *Science* *274*, 255–259.
- Recsei, P.A., Huynh, Q.K., and Snell, E.E. (1983). Conversion of prohistidine decarboxylase to histidine-decarboxylase peptide-chain cleavage by non-hydrolytic serinolysis. *Proc. Natl. Acad. Sci. USA* *80*, 973–977.
- Saarela, J., Laine, M., Tikkanen, R., Oinonen, C., Jalanko, A., Rouvinen, J., and Peltonen, L. (1998). Activation and oligomerization of aspartylglucosaminidase. *J. Biol. Chem.* *273*, 25320–25328.
- Sakon, J., Adney, W.S., Himmel, M.E., Thomas, S.R., and Karplus, P.A. (1996). Crystal structure of thermostable family 5 endocellulase E1 from *Acidothermus cellulolyticus* in complex with cellotetraose. *Biochemistry* *35*, 10648–10660.
- Salvesen, G.S., and Dixit, V.M. (1997). Caspases: intracellular signaling by proteolysis. *Cell* *91*, 443–446.
- Schmidtke, G., Kraft, R., Kostka, S., Henklein, P., Frömmel, C., Löwe, J., Huber, R., Kloetzel, P.M., and Schmidt, M. (1996). Analysis of mammalian 20S proteasome biogenesis: the maturation of  $\beta$ -subunits is an ordered two-step mechanism involving autocatalysis. *EMBO J.* *15*, 6887–6898.
- Seemüller, E., Lupas, A., Stock, D., Löwe, J., Huber, R., and Baumeister, W. (1995). Proteasome from *Thermoplasma acidophilum*: a threonine protease. *Science* *268*, 579–582.
- Smith, J.L., Zaluzec, E.J., Wery, J.-P., Niu, L., Switzer, R.L., Zalkin, H., and Satow, Y. (1994). Structure of the allosteric regulatory enzyme of purine biosynthesis. *Science* *264*, 1427–1433.
- Tarentino, A.L., Quinones, G., Hauer, C.R., Changchien, L.-M., and Plummer, T.H., Jr. (1995). Molecular cloning and sequence analysis of *Flavobacterium meningosepticum* glycosylasparaginase: a single gene encodes the  $\alpha$  and  $\beta$  subunits. *Arch. Biochem. Biophys.* *316*, 399–406.
- Tollersrud, O.K., and Aronson, N.N., Jr. (1992). Comparison of liver glycosylasparaginases from six vertebrates. *Biochem. J.* *282*, 891–897.
- Vallet, V., Ahmed, C., Gaeggler, H.-P., Horisberger, J.-D., and Rossier, B.C. (1997). An epithelial serine protease activates the amiloride-sensitive sodium channel. *Nature* *389*, 607–610.

Xu, M.-Q., Comb, D.G., Paulus, H., Noren, C.J., Shao, Y., and Perler, F.B. (1994). Protein splicing: an analysis of the branched intermediate and its resolution by succinimide formation. *EMBO J.* *13*, 5517–5522.

Xuan, J., Tarentino, A.L., Grimwood, B.G., Plummer, T.H., Jr., Cui, T., Guan, C., and Van Roey, P. (1998). Crystal structure of glycosyl-asparaginase from *Flavobacterium meningosepticum*. *Protein Sci.* *7*, 774–781.

Zwickl, P., Kleinz, J., and Baumeister, W. (1994). Critical elements in proteasome assembly. *Nat. Struct. Biol.* *1*, 765–770.

#### Protein Data Bank ID Codes

Coordinates for the T152A, T152C, and W11F mutants have been deposited in the Protein Data Bank under ID codes 9gaa, 9gac, and 9gaf, respectively.