The Mechanism of Autocatalytic Activation of Plant-type L-Asparaginases*

Karolina Michalska, Alejandra Hernandez-Santoyo, and Mariusz Jaskolski

From the Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Poznan 60-780, Poland and the Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan 61-704, Poland

Plant L-asparaginases and their bacterial homologs, such as EcAIII found in Escherichia coli, form a subgroup of the N-terminal nucleophile (Ntn)-hydrolase family. In common with all Ntn-hydrolases, they are expressed as inactive precursors that undergo activation in an autocatalytic manner. The maturation process involves intramolecular hydrolysis of a single peptide bond, leading to the formation of two subunits (α and β) folded as one structural domain, with the nucleophilic Thr residue located at the freed N terminus of subunit β. The mechanism of the autocleavage reaction remains obscure. We have determined the crystal structure of an active site mutant of EcAIII, with the catalytic Thr residue substituted by Ala (T179A). The modification has led to a correctly folded but unprocessed molecule. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact. The atomic coordinates and structure factors (code 3C17) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

Posttranslational modifications of proteins can be divided into reactions leading to covalent attachment, usually at a side chain, of a specific chemical group, such as phosphate or carboxylate, and into reactions that lead to cleavage or cleavage/rearrangement of the polypeptide backbone. The latter processes can be enzyme-catalyzed or occur without an additional biocatalyst. An important class of backbone rearrangement is connected with autocatalytic processes in maturating proteins. Notable examples in this category include intein splicing and simple backbone cleavage. Protein splicing requires cleavage of two peptide bonds that surround the so-called intein, followed by ligation of the flanking polypeptides, the N- and C-exteins. The process consists of four steps: acyl rearrangement, transesterification, cyclization of an asparagine residue, and a second acyl rearrangement (1). In maturation following the autocleavage pathway, only a single peptide cleavage occurs, and the mechanism includes only acylation and water-dependent deacylation. Although the mechanism of protein splicing has been extensively studied and seems to be well understood, the autoproteolysis reactions are rather obscure. The known examples of autoproteolytic proteins include a large group of N-terminal nucleophile (Ntn)2-hydrolases (2). In Ntn enzymes, a cleavage of a precursor molecule is required to generate a catalytic residue at the newly formed N terminus, which can be a threonine, serine, or cysteine. The family of Ntn-hydrolases includes such enzymes as aspartylglycosaminidases (3, 4), penicillin acylases (5–7), taspase1 (8), and plant-type L-asparaginases (9–12). Sequence similarity within the Ntn-hydrolase family is very limited, but despite the variation of primary structure, the proteins share a common sandwich-like αβαβα fold created by two β-sheets surrounded by two layers of α-helices (13). The autocatalyzed maturation of Ntn-hydrolases involves either the removal of a propeptide (14) or the generation of two subunits that are then utilized for the formation of the active enzyme.

The autocatalytic apparatus of maturating proteins can be described by analogy to classic serine proteases. Their reactive machinery usually consists of a catalytic triad formed by a nucleophilic residue (Ser), a general base (His), and an additional acidic residue (Asp). The nucleophile carrying out the acylation step is activated by the general base that abstracts the proton from the serine OH group. The role of the acidic residue is to properly position the basic residue and enhance its basicity (15). Various deviations from the canonical triad have been noted, including enzymes that utilize only a dyad or even a single-residue active site.

Protein splicing and autoproteolytic maturation of Ntn-hydrolases have a common first step, consisting in an N→O (or N→S) acyl shift with the formation of a tetrahedral (thio)ester intermediate (1) (Fig. 1). This intermediate exists as a five-membered oxazolidine or oxothiazolidine ring. The rearrangement requires a nucleophilic attack of a hydroxyl (or thiol) group on the peptide carbonyl C atom. There are several controversial points about this process, the main one concerning the nucleophilic activation of the attacking group. In the case of splicing proteins, it has been proposed that a conserved His residue participates in the deprotonation of the -O(S)H nucleophile, but in the available crystal structures no suitable interaction has been found. However, a hydrogen bond between a His side chain and the scissile bond N atom was observed (16). In consequence, it has been proposed that the

* The abbreviations used are: Ntn, N-terminal nucleophile; hT179A, hexagonal form of EcAIII T179A mutant; oT179A, orthorhombic form of EcAIII T179A mutant.

1 To whom correspondence should be addressed: Dept. of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Grunwaldzka 6, 60-780 Poznan, Poland. Tel.: 48-61-8291274; Fax: 48-61-8291505; E-mail: mariuszj@amu.edu.pl.

2 The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
His residue facilitates the breakdown of the tetrahedral intermediate by protonating the N atom. In inteins that utilize a Cys residue as the nucleophile, the thiol group does not require assistance of other groups (17). It has also been suggested that it is the stabilization of the tetrahedral intermediate by an oxyanion hole that drives the reaction. An additional factor is the unusual geometry of the scissile peptide bond, which is either cis (16) or distorted trans (18).

Similar problems plague the interpretation of the mechanism of Ntn-hydrolase maturation. One of the difficulties is the lack of a consensus sequence motif flanking the autocleavage point. Data obtained for the most studied Ntn-hydrolases, namely aspartylglucosaminidases, seemed to indicate that either a His (19) or an Asp residue N-terminal to the scissile bond is responsible for the activation of the nucleophilic threonine (20, 21). However, the His-Asp-Thr sequence is not conserved among other Ntn-hydrolases.

Our studies are focused on another group of the Ntn family, namely on plant-type L-asparaginases, which hydrolyze the side-chain amide bond of asparagine or of isoaspartyl dipeptides (9). Close homologs have been found not only in plants but also in bacteria, with the Escherichia coli protein, EcAIII (E. coli iaaA (ybiK) gene product), being particularly well studied (10, 22, 23). Enzymes from both sources have a catalytic Thr nucleophile located at the free N terminus of subunit β, liberated during the autoproteolytic maturation event that generates two separate subunits (α and β) from a single-chain precursor (10, 11). It is believed that the maturation mechanism involves the very same Thr nucleophile that is liberated in its consequence. However, the details of the maturation reaction are obscure, since the free α-amino group, supposed to activate the nucleophilic hydroxyl, is not present in the immature protein. A plausible candidate for the general base in the maturation of plant-type asparaginases has not been proposed so far; nor has the oxyanion hole been defined, which is necessary for the stabilization of the oxyoxazolidine intermediate.

To investigate the structural basis of EcAIII maturation, we have generated and crystallized an active site mutant with the catalytic Thr residue substituted by alanine (T179A). The structure of this EcAIII variant has been determined in two crystal forms, hexagonal (hT179A) and orthorhombic (oT179A). The latter one (Protein Data Bank code 2ZAK) has been reported earlier (24) in the context of EcAIII polymorphism. The analysis in this paper is mainly based on the hexagonal form and on comparison of the two structures.

**EXPERIMENTAL PROCEDURES**

**Cloning, Expression, and Purification**—The plasmid carrying the mutated gene of EcAIII was constructed using a pET-11d vector (Novagene) containing the wild type gene sequence (25) as a template. The plasmid DNA was PCR-amplified using the following primers: 5'-AATGGGCACCGGT-
GGGGG-3’ and 5’-CCCCC.ACGCGGCCCATT-3’. The codon change ACC→GCC (underlined) introduced the T179A mutation in the protein sequence. The recombinant plasmid was sequenced to confirm the correctness of the inserted sequence and subsequently transformed into E. coli strain BL21(DE3)pLysS (Novagene).

30 ml of LB medium with 100 µg/ml−1 ampicillin were inoculated with 20 µl of BL21(DE3)pLysS cells containing the recombinant plasmid and grown overnight at 37 °C. The culture was used to inoculate 1 liter of LB medium with appropriate antibiotics and grown for 4 h, when the expression of the recombinant protein was induced by the addition of isopropyl thio-β-D-galactoside to a final concentration of 1 mM. The bacteria were incubated for 4 h at 37 °C and then centrifuged at 6000 rpm, 4 °C, for 25 min and resuspended in 50 mM Tris/HCl buffer, pH 8.5. After sonication, the lysate was centrifuged at 9000 rpm, 4 °C, for 25 min, to remove cell debris. All purification steps were carried out at room temperature. The supernatant was fractionated with ammonium sulfate. The precipitate formed at 35–65% saturation, containing the recombinant protein, was collected, resuspended in water, and dialyzed against water and subsequently against 50 mM Tris/HCl buffer, pH 8.5. The solution was loaded onto a DEAE-cellulose column equilibrated with 50 mM Tris/HCl buffer, pH 8.5. The column was washed with the same buffer, and then a 0–1 M linear NaCl gradient in the same buffer was applied. The fractions containing T179A EcAIII were collected and rechromatographed under the same conditions but in a 0–0.3 M NaCl gradient. Fractions with EcAIII were collected, concentrated using a YM-10 membrane Amicon 8010 system (Millipore), and loaded onto a Sephadex G75 gel filtration column (Amersham Biosciences) equilibrated with 50 mM Tris/HCl buffer, pH 8.5, and 0.2 M NaCl. The pure protein was pooled, concentrated, and kept frozen at −80 °C until use. The expression product migrates in SDS-PAGE as a single band with a molecular mass of 33 kDa, corresponding to intact precursor molecule.

**Crystallization**—Crystallization was carried out by the hanging drop vapor diffusion method at room temperature. Drops (4 µl) contained a 1:1 volume ratio of protein at a 7 mg/ml concentration in a solution composed of 50 mM Tris/HCl buffer, pH 8.5, and 0.1 M NaCl and a range of precipitants from HMP (31). The reservoir solution contained 0.2M MgCl2, 30% (w/v) polyethylene glycol 4000, and 0.1 M Tris/HCl, pH 8.5.

**Data Collection and Processing**—The diffraction experiment for the hexagonal crystals was performed using the X12 European Molecular Biology Laboratory beamline of the Deutsches Elektronen Synchrotron and a MarResearch CCD225 detector. The crystals did not require additional cryoprotection and were directly flash-cooled in a nitrogen gas stream at 100 K. The images were indexed and integrated using Mosflm (26) and merged using Scala (27) in the 6/mmm Laue symmetry. Since the 00l reflections were missing in the data set, the space group was determined at the molecular replacement stage. Data collection and processing statistics are given in Table 1.

**Structure Determination and Refinement**—The structure of the hexagonal crystals was solved by molecular replacement in Phaser (28), using as a probe the native EcAIII (αβ2)2 dimer of heterodimers deposited in the PDB with the accession code 1K2X. Due to space group ambiguity, all possibilities in the 622 class were tested. The correct solution, with the rotation and translation Z-scores of 15.3 and 78.7, respectively, indicated the space group symmetry (P6_1,22) and two monomeric precursor molecules in the asymmetric unit, corresponding to 76% solvent content. Maximum likelihood refinement (including TLS optimization) (29) in Refmac5 (30) alternated with manual model rebuilding done in Coot (31). The refinement statistics are given in Table 1. Structure determination and refinement of the orthorhombic form has been described elsewhere (24).

| Parameter | Value |
|-----------|-------|
| Data collection and processing | X12, EMBL Hamburg |
| Wavelength (Å) | 1.000 |
| Temperature of measurements (K) | 100 |
| Space group | P6_1,22 |
| Cell dimensions (Å) | a = 149.5, c = 214.3 |
| Resolution range (Å) | 47.7–1.95 (2.06–1.95)* |
| Reflections collected | 579,986 |
| Unique reflections | 98,405 (14,515) |
| Completeness (%) | 96.6 (98.5) |
| Redundancy | 5.9 (5.2) |
| Favored | 14.3 (2.2) |
| Outliers | 0.107 (0.760) |

*Values in parentheses correspond to the last resolution shell.

**RESULTS AND DISCUSSION**

**Expression and Purification**—The T179A mutant of plant-type asparaginase from E. coli has been expressed and purified to homogeneity. The expression was carried out in an E. coli strain that constitutively produces small amounts of unmutated EcAIII. However, from a visual inspection of the SDS-polyacrylamide gels after purification, there was no indication of the presence of the native protein, which should appear as
two bands corresponding to the two autolytic products with molecular masses of 19 and 14 kDa. The single 33 kDa band corresponds to the intact T179A precursor protein, which does not undergo autoprocessing, and not to the wild-type EcAIII, since at the end of the purification procedure, the latter protein exists only in the activated form (9). The maturation can theoretically take place either in cis or in trans mode. In the cis mechanism, the reaction occurs within one precursor molecule, whereas a trans reaction would involve two molecules, one acting as the catalyst and the other as a substrate. Since in the applied expression system there should be at least catalytic amounts of wild-type EcAIII, our failure to detect cleaved EcAIII suggests that the autoprocessing reaction occurs in the cis mode. In addition, this result confirms that the EcAIII precursor with Thr\textsuperscript{179} replaced by alanine is incapable of autoprocessing.

**Structure Solution**—The hT179A structure was solved by molecular replacement using the Protein Data Bank 1K2X model of mature EcAIII as the molecular probe. The asymmetric unit of the P\textsubscript{6}\textsubscript{1}2\textsubscript{2} cell contains two intact polypeptides with T179A EcAIII sequence, organized into an AB homodimer. The overall fold of this EcAIII precursor is essentially identical to that of the mature enzyme. The protein core is established by two β-sheets facing each other in a parallel manner (Fig. 2a). The smaller β-sheet comprises four antiparallel strands (S4\textsubscript{β}, S5\textsubscript{β}, S6\textsubscript{β}, and S7\textsubscript{β}), whereas the bigger one contains nine strands (S9\textsubscript{β}, S8\textsubscript{β}, S1\textsubscript{α}, S5\textsubscript{α}, only in hT179A; see below), S1\textsubscript{β}, S2\textsubscript{β}, S2\textsubscript{α}, S3\textsubscript{α}, and S4\textsubscript{α}), which, except for the last interaction, are organized in an antiparallel fashion. This β-sheet is extended by an additional β-strand (S3\textsubscript{β}) belonging to the second molecule. The envelope of the protein core is formed by two layers of helices. The layer flanking the bigger β-sheet is made of seven helices (H1\textsubscript{α}, H2\textsubscript{α}, H3\textsubscript{α}, H4\textsubscript{α}, H5\textsubscript{α}, H6\textsubscript{α}, and H7\textsubscript{α}), two of which have the 3\textsubscript{10} classification (H1\textsubscript{α} and H6\textsubscript{α}), whereas the smaller β-sheet is covered by four α-helices (H1\textsubscript{β}, H2\textsubscript{β}, H3\textsubscript{β}, and H4\textsubscript{β}). This overall fold, usually described as the αββα sandwich, is found in all Ntn-hydrolases.

**The Linker Region**—The fragment of the polypeptide chain that acts as a linker (residues Ala\textsuperscript{161}–Gly\textsuperscript{178}) between the subunits was not visible in the structures of mature EcAIII, either due to disorder or because of C-terminal degradation occurring during the purification procedure (9). The highest degree of completeness of subunit α in the mature protein (up to Ala\textsuperscript{161}) was observed in the 2ZAK structure. In all other structures, the visible C-terminal chain of subunit α is several residues shorter. The covalent anchoring of this fragment in the present mutant protein has stabilized its conformation, making it possible to model as many as 11 residues upstream of the 178–179 cleavage point. Nevertheless, a small portion of the linker region, about seven residues long, is still not visible in the electron density. In the orthorhombic structure, the disordered fragment is even longer, 14 residues (for chain A). This result indicates that there is intrinsic flexibility in the linker region, regardless of the maturation stage of the protein. The fragments of the linker region that have well defined electron density mostly do not adopt any regular secondary structure, with the exception of residues Leu\textsuperscript{171}–

---

**FIGURE 2. Overall structure of the T179A mutant of EcAIII (stereo).** a, schematic representation of the protein fold with bound ions. The polypeptide chains A and B are shown in sand and gold color, respectively. Sodium ions are shown as purple spheres, and chloride ions are shown as green spheres. The broken line indicates a gap in the model in the linker region. b, superposition of EcAIII molecules. Pink, T179A mutant, hexagonal form (this work); gray, T179A mutant, orthorhombic form (Protein Data Bank code 2ZAK); blue (subunit α) and yellow (subunit β), mature EcAIII (Protein Data Bank code 1K2X). Selected secondary structure elements discussed under “Results and Discussion” are indicated by their labels.

---

3 The topology of the protein is described as in mature EcAIII (10) (i.e. a Greek letter following the symbol S for strand, H for helix); the consecutive number of a secondary structure element indicates its location in the α or β portion of the sequence.
Mechanism of Autocleavage of EcAll

Asp^{172}, which have a β-strand conformation, establishing the S5α element involved in antiparallel β-sheet interactions with the C terminus of strand S1α.

Comparison with Other EcAlll Structures—Superposition of the hexagonal T179A EcAlll mutant structure on the mature EcAlll model deposited in the Protein Data Bank as 1K2X reveals only minor changes in the overall conformation of the protein. The most pronounced differences are found at H1α, H2α, and the region directly preceding these helices, namely the C-terminal fragment of S1α and the S1α/H1α loop (Fig. 2b). In particular, according to DSSP secondary structure assignment (34), the β-strand S1α in the precursor is slightly longer than in mature EcAlll. This extension is stabilized by hydrogen bonds with the two-residue S5α element that is not present in the mature molecule. The 3_10 helix H1α and the N terminus of H2α are in the mutant structure shifted toward solvent. Additional small changes are observed in the positions of helices H6α and H7α.

When the hexagonal and orthorhombic models of the mutant protein are compared, the only visible discrepancies are located in the same areas as in the comparison with 1K2X (Fig. 2b). Additionally, in the oT179A form, the intersubunit linker is fully disordered, mimicking the situation in all mature EcAlll structures. The only exception is the presence of the Gly^{178} residue, which together with Ala^{179} defines the scissile peptide bond. Interestingly, the lack of the S5α strand in oT179A influences the length of S1α as well, which without the stabilization within the S1α-S5α β-sheet contains the same number of residues as in the mature protein.

The root mean square deviations (Å) between equivalent Ca atoms in pairwise comparisons show the same level of similarity for the hT179A/oT179A (0.66) and hT179A/1K2X (0.65) pairs, whereas for the oT179A/1K2X pair (0.49), the superposition is slightly better. Such a result indicates that the mutation itself and the intact polypeptide chain that is its consequence are not responsible for the observed conformational changes. The root mean square deviation values also show that the hexagonal form of T179A EcAlll is slightly different from other EcAlll molecules, probably due to different crystal packing interactions. Analysis of superpositions of the separate α and β portions in mature and precursor EcAlll indicates that there is no subunit reorientation upon maturation.

Inorganic Ions—The high concentration of NaCl used for crystallization resulted in a number of inorganic ions bound to the protein molecule (Fig. 2a). The identity of sodium ions with complete octahedral coordination was confirmed by the bond-valence method (35). Interestingly, except for one ion, the distribution of the Na⁺ cations follows the molecular NC5 dyad. The binding of the Cl⁻ ions is less regular; only 12 of 17 obey the symmetry of the protein dimer.

In a manner identical to that observed in all mature EcAlll enzymes, one of the sodium cations occupies the cavity within a loop formed by residues Leu^{60}–Ile^{70}, where it is coordinated by six main-chain carbonyl groups. Previously, we suggested that this Na⁺ cation is essential for proper conformation of the Leu^{60}–Ile^{70} loop, which in turn supports the required orientation of the catalytic Thr^{179} residue for the enzymatic reaction (10, 11). In the active enzyme, this stabilization is achieved by a hydrogen bond between the free α-amino group of Thr^{179} and the main-chain carbonyl of Asn^{67}. In the T179A mutant, such an interaction is not possible, because the α-amino group of Ala^{179} is engaged in the Gly^{178}–Ala^{179} peptide bond, which forces its N-H donor to point in a different direction. Instead, the side-chain amide donor of Asn^{67} forms a hydrogen bond with the carbonyl group of the scissile peptide (Fig. 3, a and b).

It appears, therefore, that the “sodium loop” of EcAlll may be important not only for maintaining the catalytic apparatus of the mature enzyme in its active conformation but also for proper orientation of the “internal substrate” of the precursor protein for the autocatalytic reaction.

The active site of the T179A mutant in both crystal structures is occupied by a chloride anion tightly anchored by the guanidinium group of Arg^{207} (Fig. 3a). In the EcAlll structure crystallized in complex with L-aspartate (presumably also in a substrate complex), this position is occupied by the α-carboxylate group of the ligand molecule, which forms an ideal partner for a strong salt bridge interaction with Arg^{207}. In a homologous protein from yellow lupine (LIA) (11), a chloride anion was found to occupy the same position. These observations clearly emphasize the affinity of the active site for a negatively charged species and the ability of the Arg^{207} anchor to capture anionic components of the buffer if no suitable substrate is available for docking in the active site.

At about 2.7 Å from the Cl⁻ anion, a sodium cation is present in the catalytic cavity (Fig. 3a). In addition to the Cl⁻ anion, its coordination sphere includes two water molecules, one of which is Wat1, and two main-chain carbonyl O atoms (Gly^{231} and Lys^{176}). The Na⁺ cation is located relatively close to Gly^{178}, at 4 Å from the O atom, but it is unlikely that it participates in the autoproteolytic mechanism. Theoretically, its location near the Gly^{178} oxygen atom, which develops a negative charge during the autocatalytic reaction, could suggest that it serves to stabilize the oxyanion. Alternatively, it could be required before the reaction takes place, to polarize the C=O bond and enhance the electrophilicity of the carbonyl C atom, making it more susceptible to nucleophilic attack. However, the carbonyl group of Gly^{178} does not participate directly in the sodium coordination sphere. Instead, it forms a hydrogen bond with one of the coordinated water molecules (Wat1; Fig. 3a). In principle, the polarizing effect of the sodium cation could influence the carbonyl group of the scissile bond indirectly, through mediation of a water molecule, but it is difficult to decide with certainty if this metal cation can play a role in the autocatalytic mechanism.

It is of note that a calcium cation was found to be important for the maturation of penicillin amidase (36), another member of the Ntn-hydrolase family. The metal has no direct involvement in the reaction but is required to stabilize the protein in a conformation that is necessary for autoproteolysis. This is supported by structural data showing that the Ca²⁺ cation is buried within the protein core in a binding site located away from the catalytic pocket (37). Although biochemical data clearly show that Ca²⁺ ions are necessary for the autoproteolytic reaction of penicillin amidase, this does not automatically prove that the
sodium cations in EcAIII play the same role, since their binding sites have quite different localizations.

The Cleavage Site—The Thr$^{179}$ → Ala substitution has abolished auto-proteolysis of EcAIII, thus allowing the determination of the structure of the proenzyme, with the scissile Gly$^{178}$-Ala$^{179}$ peptide bond intact. The total arrest of the maturation process confirms the previous prediction that the OH group of Thr$^{179}$ is crucial for this reaction. Although the structure of T179A EcAIII is not an exact representation of the wild-type precursor molecule, the native state can be modeled by changing the Ala$^{179}$ side chain to Thr. Such an approach is justified by the relatively small changes in the overall structure of the mutated variant in comparison with the mature protein. Moreover, also the architecture of the active site is well conserved (Fig. 3).

The conformation of the main chain near the scissile peptide bond in hT179A has standard β-strand geometry. Also, the ω torsion angle is normal, with no deviation from planarity. In the oT179A structure, the main-chain torsion angles of Ala$^{179}$ are ϕ = −83°−57° and ψ = 178°/177°, and the ω torsion angle is standard. Although this conformation is also acceptable in the Ramachandran diagram, it changes the arrangement within the active site, especially the predicted orientation of the nucleophilic hydroxyl group relative to the Gly$^{178}$ carbonyl C atom (Fig. 3).

Conformation of Thr$^{179}$—A theoretically modeled Thr$^{179}$ side chain can adopt three favored conformations: gauche+, corresponding to the χ torsion angle (N-Cα-Cβ-Oγ) of 60°; gauche− (−60°); and anti (180°). In the mature protein, the Thr residue exists in the gauche− conformation, dictated by adequate orientation of the hydroxyl group with respect to the scissile bond of the asparagine or isoaspartyl peptide substrate. In the proenzyme, the conformation of the Thr side chain should allow the nucleophilic attack on the carbonyl C atom of the

**FIGURE 3.** The active site area of the T179A mutant of EcAIII. 

a, stereoview of the active site in the hexagonal structure, shown in 2F$_{o}$ − F$_{c}$ electron density contoured at the 2.3σ level. Potential hydrogen bonds are indicated by broken lines. b, conformation of the 178-179 scissile peptide bond in the hexagonal (left) and orthorhombic (right) structures with emphasis on the environment of the Gly$^{178}$ carbonyl O atom. c, Ca superposition of the active sites of EcAIII. Gray, T179A mutant, hexagonal form (this work); green, mature EcAIII (Protein Data Bank code 1K2X).
Mechanism of Autocleavage of EcAIII

This proposal, linking the progress of the autocatalytic reaction with an initial rotation of the Thr$^{179}$ side chain into an eclipsed orientation, does not explain the stimulus for driving this rotation into a sterically unfavorable conformation. One possibility, suggested by Fig. 4a, would be a very short repulsive contact between the methyl group of Thr$^{179}$ and the main-chain C=O group of Ser$^{196}$. However, a rotation of the Thr$^{179}$ side chain from the gauche$^-$ to syn orientation would introduce another steric collision with the side chain of Thr$^{197}$. Considering the two steric clashes, it is probably easier to relieve the strain when it occurs with the participation of a side chain. On the other hand, the Thr$^{197}$ side chain has only limited freedom, since it is locked in its orientation by a tandem of N-H...O hydrogen bonds donated by the two peptides that follow Thr$^{197}$ in the EcAIII sequence.

This interpretation becomes more complicated when the active site of the orthorhombic form of the T179A mutant is analyzed. In this case, Gly$^{178}$ is the only residue at the subunit junction that is visible in the electron density maps. The rest of the linker (down to Leu$^{164}$ (chain A) and Glu$^{159}$ (chain B)) does not have stable conformation and is not present in the electron density maps. With the geometry of the Gly$^{178}$-Ala$^{179}$ bond seen in the oT179A structure, the structural environment for the cleavage reaction would be quite different. Most notably, in the framework of the oT179A active site, there is no orientation of the Thr$^{179}$ side chain that would bring the Oγ nucleophile closer than 3.8 Å to the scissile bond C atom, and the angle of attack is also unfavorable. Although the source of the evident conformational freedom of the linker region is not immediately obvious, it may be assumed that this flexibility is necessary for proper organization of the active site elements in the highly strained situation required for cis-cleavage of the Gly$^{178}$-Thr$^{179}$ bond by the Thr$^{179}$ OH nucleophile.

Nucleophile Activation—The gauche$^-$ and syn conformers of Thr$^{179}$ would have distinct geometry of a cis nucleophilic attack by the hydroxyl group and would also participate in different hydrogen bonds. Specifically, in the gauche$^-$ conformation, the -OH group interacts with the side chain of Thr$^{197}$, whereas in the syn conformation, it interacts with the side chain of Thr$^{230}$. The optimal direction of attack is along a line that is perpendicular to the scissile peptide plane with an average Oγ...C=O distance of 2.7 Å, as observed in a number of serine proteases (38). Thus, the reaction should begin with a rotation of the Thr$^{179}$ side chain toward an eclipsed conformation at χ = 0° (Fig. 4a), leading at the same time to the formation of a (Thr$^{179}$)Oγ-H...Oγ(Thr$^{230}$) hydrogen bond, followed by a distortion of the scissile peptide group, which would involve pyramidalization of the carbonyl C atom, oxyanion formation, and loss of π-electron bonding at the peptide N atom. An important role in this scenario is played by the side chain of Thr$^{230}$, which polarizes the O-H group of the Thr$^{179}$ nucleophile and fixes the scissile bond in its reactive orientation by flanking it with a chain of hydrogen-bonded groups, (Thr$^{179}$)Oγ-H...Oγ-H...O=C(Met$^{177}$).
Both of these residues can potentially act as a general base enhancing the nucleophilicity of Thr179 by proton abstraction. However, for the Thr179 conformation with $\chi$ between 0 and $-60^\circ$, both hydrogen bonds are possible, complicating the identification of the general base. Since Thr179 is an acceptor in two hydrogen bonds from main-chain nitrogen atoms of Gly198 and Gly199, it is not able to abstract a third proton from the Thr179 hydroxyl group. For this reason, Thr230 is a better candidate as a nucleophilic activator. The OH group of Thr230 is a proton donor in the only other hydrogen bond that it forms, namely with the main-chain O atom of Met177. Thus, it can be considered a suitable proton abstractor for the OH group of Thr179.

The Oxyanion Hole—An important aspect of the autoproteolytic mechanism concerns the identification of the oxyanion hole, which stabilizes the negatively charged tetrahedral reaction intermediate. For the enzymatic reaction carried out by mature EcAIII, the oxyanion hole has been suggested to be composed of the N-H group of Gly231 and (atypically) of the mature EcAIII, the oxyanion hole has been suggested to be the free HN(His192) hydrogen bond. The proposed autocleavage mechanism includes activation of the Ser170 hydroxyl group by a water molecule fixed in the active site by four hydrogen bonds, one of which requires a peptide flip of the Gly168...HN(His192)...Ser170 bond (39). In an active site mutant analogous to the present EcAIII variant, the Gln168...Val175 region adopts a $\beta$-strand conformation and participates in antiparallel $\beta$-sheet interactions, one of which is a stabilizing (Gly169...O...HN(His192)...Ser170) hydrogen bond. The proposed autocleavage mechanism includes activation of the Ser170 hydroxyl group by a water molecule fixed in the active site by four hydrogen bonds, one of which requires a peptide flip of the Gln168...Gly169 bond. The tetrahedral intermediate is stabilized by an oxyanion hole formed by the main-chain N atom of His192. Once the $\beta$-acyl ester is formed, a second water molecule (interacting with the imidazole ring of His192) completes the final hydrolytic step. A similar interpretation of the autoproteolytic mechanism has been published for cephalosporin acylase (40). In both mutated enzymes, some distortions of main-chain geometry were observed near the scissile bond, but the geometry of the scissile peptide bond has standard stereochemistry.

On the other hand, a highly strained geometry has been considered as the driving force for the removal of the N-terminal extension during the activation of the proteasome subunit $\beta$ (41), in which the polypeptide chain at the autoproteolytic site adopts a $\gamma$-turn conformation. Analysis of a precursor structure with the nucleophilic Thr replaced by Ala has led to the conclusion that the attack of the OH group on the scissile peptide bond (Gly-Thr) is not associated with an activation of the nucleophile by any general base. The putative involvement of the adjacent Lys135 side chain postulated earlier (42) has been dismissed because the hydrogen bond network at this side chain indicates a protonated -NH$_3^+$ group (41).

In aspartylglucosaminidase, the presence of an Asp residue at the N terminus of the cleaved peptide bond (Asp151...Thr152) has suggested an alternative mechanism, in which the Asp side chain would deprotonate the Thr152 hydroxyl group prior to the nucleophilic attack (21). The reaction could be additionally favored by the slight geometric distortion of the scissile peptide bond ($\omega = 160^\circ$) observed in the T152A and T152C mutants of the enzyme (20). However, no such deviation was found in the D151N variant of the protein (21). Moreover, in the latter precursor, the Asn151 side chain adopts a conformation that excludes its interaction with the catalytic Thr residue. For bacterial aspartylglucosaminidase, the authors postulate that the Asp residue plays a dual role, namely that it enhances the nucleophilic activity of Thr152 and induces an unstable, proteolytically labile main-chain conformation. On the other hand, for the human enzyme, only the structural influence of Asp has been proposed, whereas a water molecule has been considered as a general base (43). Alternatively, His150 has been considered as a potential nucleophile activator (19, 44), although biochemical studies of several His150 mutants showed that the proteins still possessed residual autoproteolytic activity (19, 45). The Asp151 replacements had a more severe effect, but the Glu and Gly mutants still retained a partial ability to undergo autoprocessing. Interestingly, the introduction into human aspartylglucosaminidase of an Ille-Gly segment, which resembles the plant-type 1- asparaginase sequence N-terminal to the scissile Gly-Thr bond, leads to a completely dead enzyme (45). This, together with other observations, suggests that the mechanism of autocatalytic activation may be different in different classes of Ntn-hydrolases, in which case the Ockham’s razor-motivated efforts to formulate a universal theory would be ill fated.

The interpretation of the structural data obtained for a series of aspartylglucosaminidase precursors is complicated by the fact that the self-processing site adopts different conformations leading to nonconserved interactions. For example, in the T151A variant, the carbonyl O atom of Asp151 is hydrogen-bonded to the side chain of Ser50 (equivalent of Asn67 in the EcAIII structure), but in the T152C variant, this O atom interacts with the OH group of Thr179 (Thr199 in EcAIII). In the W11F structure, the Asp151 O atom participates in an additional hydrogen bond with a water molecule, also found in the present EcAIII mutant structure (Wat1). The latter observation has led to the conclusion that Thr179 and the water molecule
create the oxyanion hole for the autoproteolytic reaction. A similar tandem has been proposed to define the oxyanion hole in splicing proteins (46).

None of these interpretations can be directly applied to EcAIII. First of all, it is difficult to unambiguously identify the general base activating the nucleophile. However, the most likely residue to abstract proton from the Thr179 nucleophile is Thr230. This suggests that the general base residue does not have to be close to the nucleophilic residue in the primary structure, as in aspartylglucosaminidases, where it is assumed to immediately precede the nucleophile, but that it can be contributed by a suitable element of tertiary structure.

Mechanism of Autocleavage of EcAIII

On the other hand, mutagenesis studies of aspartylglucosaminidases have established that substitution of the equivalent of Thr230 by Ala does not abolish autoproteolysis, although it makes it less efficient (19, 47). Moreover, it was reported that although this mutation completely inhibits the catalytic activity of the human enzyme (47), the bacterial homolog retains 30% of its original activity (19). As a conclusion, it has been postulated that in aspartylglucosaminidases, the Thr residue is required for proper positioning of Asp151 (20).

With regard to Thr197, an alanine substitution of its equivalent in aspartylglucosaminidase resulted in a processing-defective human enzyme (47), whereas for the bacterial protein, only the processing rate was affected, whereas the reaction product was unchanged (19). Those observations suggest that none of these Thr residues is a universal and absolute prerequisite for the autoproteolytic process.

The role of the Wat1 molecule, interacting with the Gly178 O atom, may be to polarize the scissile peptide carbonyl bond and thus enhance its electrophilicity, making it a better target for the nucleophilic attack. This would mirror the conclusion drawn for the splicing proteins, namely that elements creating the oxyanion hole for the stabilization of the tetrahedral intermediate may also be crucial before it is formed by promoting a favorable charge distribution. However, there are no reasons to believe that in the case of EcAIII, when the oxyoxazolidine ring is created and the local geometry changes, the Gly178 O atom would still interact with Wat1, since the direction of the nucleophilic attack probably pushes the carbonyl O atom away from this water molecule. In consequence, the only element that definitely forms part of the oxyanion hole for the autocleavage reaction is the Asn67 side chain.

The proposal that the autoproteolytic reaction is driven by stereochemical distortion of the scissile bond seems plausible, because in all Ntn-hydrolases, the N-terminal residue of the cleavable peptide facilitates distortions. However, in the T179A EcAIII structures, the Gly178-Ala179 bond has standard conformation, which can be either the innate feature of the protein or an artifact introduced by the mutation. The geometric analysis of the nucleophilic attack (see "Conformation of Thr179") suggests that some conformational transformations at the autocleavage site are necessary for optimal conditions.

Another question concerns the water molecule in the decylation step. This issue has not been discussed for aspartylglucosaminidases. In the T179A EcAIII structure, we do not see any water molecule within 3.5 Å of the carbonyl C atom of Gly178. The closest water molecules are the potential element of the oxyanion hole (Wat1) that is hydrogen-bonded to the carbonyl O atom of Gly178 and another water molecule that interacts with Gly178 and Asn67 (Wat2; Fig. 3b). The latter molecule is located 3.7 Å from the scissile bond N atom, so it could participate in the acylation step by protonating the leaving group (i.e. the α-nitrogen atom of Thr179).

Conclusions—Wild-type E. coli asparaginase/isoaspartyl peptidase, EcAIII, in common with all Ntn-hydrolases, undergoes autoproteolytic activation, splitting itself with the liberation of Thr179 at the newly created N terminus as the catalytic nucleophile. To elucidate the mechanism of autoproteolysis, we have created a recombinant mutant of the protein with the catalytic Thr residue substituted by alanine. The mutated protein was expressed in an E. coli strain that also produces small amounts of native EcAIII. The fact that the purified mutant remained unprocessed suggests that the maturation reaction has cis character and that Thr179 is essential for autocatalysis. The single chain of the mutated protein has the same fold as the two subunits of the mature enzyme. Although a small fragment of the linker peptide joining the two subunits is still disordered, its key fragment, N-terminal to the 178-179 cleavage site, has very good definition in electron density. In contrast to some other Ntn-hydrolase precursors (20, 41), the cleavable Gly178-Ala179 peptide bond does not show any stereochemical distortions. To understand the autocatalytic reaction, one has to (i) consider the conformation of the nucleophilic Thr179 residue, (ii) propose a suitable general base activator of the -OH nucleophile, and (iii) identify a region of the molecule that will act as the oxyanion hole stabilizing the negatively charged O atom of the tetrahedral intermediate in the course of the reaction. Since the Thr179 rotamer visible in mature EcAIII is not compatible with a nucleophilic attack on the carbonyl C atom in front of it, we postulate that the reaction begins with a rotation of the Thr179 side chain, which may be coupled with a distortion of the scissile peptide bond. In its attacking position, the Thr179 OH group would become a donor in a hydrogen bond with Thr230, suggesting that the latter residue may act as a general base. The tetrahedral intermediate with sp3 configuration of the Gly178 C atom and negatively charged O atom can be stabilized in an oxyanion hole formed by the side-chain amide nitrogen of Asn67 and a water molecule. This water molecule is probably also needed at the initial stage of the reaction, to polarize the Gly178 C=O bond and make the peptide more susceptible to hydrolysis. It is important to note that the oxyanion hole of the autoproteolytic reaction is entirely different from the oxyanion hole utilized by the mature enzyme. It is also of note that Asn67 is part of a rigid Na+ binding loop that remains invariant regardless of the maturation stage of the protein. However, the Asn67 residue seems to serve a dual role. Although in the precursor it functions as an element of the oxyanion hole, in the mature enzyme it fixes the Thr179 nucleophile in the active conformation. The above view of the autoproteolytic mechanism is applicable to plant-type Ntn-hydrolases. The lack of universal conservation of the residuesflanking the scissile bond suggests that there may be different variants of this mechanism in different groups of Ntn-hydrolases.
Acknowledgments—The help of Grzegorz and Anna Bujacz with the x-ray diffraction experiments is gratefully acknowledged. Some of the calculations were carried out in the Poznan Metropolitan Supercomputing and Networking Center.

REFERENCES

1. Evans, T. C., and Xu, M.-Q. (2002) Chem. Rev. 102, 4869–4883
2. Brannigan, J. A., Dodson, G., Duggleby, H. J., Moody, P. C., Smith, J. L., Tomchick, D. R., and Murzlin, A. G. (1995) Nature 378, 416–419
3. Oinonen, C., Tikkanen, R., Rouvinen, J., and Peltonen, L. (1995) Nat. Struct. Biol. 2, 1102–1108
4. Guo, H. C., Xu, Q., Buckley, D., and Guan, C. (1998) J. Biol. Chem. 273, 20205–20212
5. Duggleby, H. J., Tolley, S. P., Hill, C. P., Dodson, E. J., Dodson, G., and Moody, P. C. F. (1995) Nature 373, 264–268
6. Suresh, C. G., Pundle, A. V., Sivaraman, H., Rao, K. N., Brannigan, J. A., McVey, C. E., Verma, C. S., Dauter, Z., Dodson, E. J., and Dodson, G. G. (1999) Nat. Struct. Biol. 6, 414–416
7. Kim, Y., Yoon, K., Khang, Y., Turley, S., and Hol, W. G. (2000) Structure Fold. Des. 8, 1059–1068
8. Khan, J. A., Dunn, B. M., and Tong, L. (2005) Structure 13, 1443–1452
9. Michalska, K., Michalska, K., Brzezinski, K., Kisiel, A., Podkowski, J., Bonthron, D. T., Krowarsch, D., Otlewski, J., and Jaskolski, M. (2004) Biol. Chem. 385, 13397
10. Michalska, K., Brzezinski, K., and Jaskolski, M. (2005) J. Biol. Chem. 280, 28484–28491
11. Michalska, K., Bujacz, G., and Jaskolski, M. (2006) J. Biol. Chem. 360, 105–116
12. Michalska, K., and Jaskolski, M. (2006) Acta Biochim. Polon. 53, 627–640
13. Oinonen, C., and Rouvinen, J. (2000) Protein Sci. 9, 2329–2337
14. Seemu¨ller, E., Lupas, A., and Baumeister, W. (1995) Nature 373, 416–419
15. Dodson, G., Wlodawer, A., and Ton, L. (2005) Structure 13, 1443–1452
16. Müller, P., Kopke, S., and Sheldrick, G. M. (2003) Acta Crystallogr. D 59, 32–37
17. Ignatova, Z., Wischniewski, F., Notbohm, H., and Kasche, V. (2005) J. Mol. Biol. 348, 999–1014
18. Hewitt, L., Kasche, V., Lummer, K., Lewis, R. J., Murshudov, G. N., Verma, C. S., Dodson, G. G., and Wilson, K. S. (2000) J. Mol. Biol. 302, 887–898
19. Radisky, E. S., and Koshland, D. E. Jr. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 10316–10321
20. Kim, J. K., Yang, I. S., Rhee, S., Dauter, Z., Lee, Y. S., Park, S. S., and Kim, K. H. (2003) Biochemistry 42, 4084–4093
21. Yoon, J., Oh, B., Kim, K., Park, J., Han, D., Kim, K. K., Cha, S. S., Lee, D., and Kim, Y. (2004) J. Biol. Chem. 279, 341–347
22. Ditzel, L., Huber, R., Mann, K., Heinemeyer, W., Wolf, D. H., and Groll, M. (1998) J. Mol. Biol. 279, 1187–1191
23. Schmidtke, G., Kraft, R., Kostka, S., Henklein, P., Frömmel, C., Löwe, J., Huber, R., Kloetzel, P. M., and Schmidt, M. (1996) EMBO J. 15, 6887–6898
24. Saarelä, J., Oinonen, C., Jalkanen, A., Rouvinen, J., and Peltonen, L. (2004) Biochem. J. 378, 363–371
25. Guan, C., Cui, T., Rao, V., Liao, W., Benner, J., Lin, C. L., and Comb, D. (1996) J. Biol. Chem. 271, 1732–1737
26. Saarelä, J., Laine, M., Tikkanen, R., Oinonen, C., Jalkanen, A., Rouvinen, J., and Peltonen, L. (1998) J. Biol. Chem. 273, 25320–25328
27. Schnitzberger, F., Kilkenny, M. L., Loble, C. M., Webb, M. E., Vinkovic, M., Matak-Vinkovic, D., Witty, M., Chiragadze, D. Y., Smith, A. G., Abell, C., and Blundell, T. L. (2003) EMBO J. 22, 6193–6204
28. Tikkanen, R., Riikonen, A., Oinonen, C., Rouvinen, R., and Peltonen, L. (1996) EMBO J. 15, 2954–2960