# Crystal Structures of Glutaryl 7-Aminocephalosporanic Acid Acylase: Insight into Autoproteolytic Activation †

**ARTICLE** *in* BIOCHEMISTRY · MAY 2003

Impact Factor: 3.02 · DOI: 10.1021/bi027181x · Source: PubMed



**7 AUTHORS**, INCLUDING:

Jin Kwang Kim
University of California, Irvine
**31** PUBLICATIONS **245** CITATIONS
SEE PROFILE

Sung Soo Park
Korea University
**4** PUBLICATIONS **91** CITATIONS
SEE PROFILE

Kyung Hyun Kim
Korea University
**53** PUBLICATIONS **654** CITATIONS
SEE PROFILE

# Crystal Structures of Glutaryl 7-Aminocephalosporanic Acid Acylase: Insight into Autoproteolytic Activation†

Jin Kwang Kim,‡ In Seok Yang,‡ Sangkee Rhee,§ Zbigniew Dauter,|| Young Sik Lee,‡ Sung Soo Park,‡ and Kyung Hyun Kim*,‡

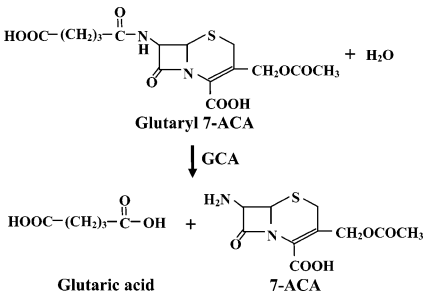
*Graduate School of Biotechnology, Korea University, Seoul 136-701, Korea, College of Life Science and Agriculture, Seoul National University, Suwon 740-644, Korea, and Synchrotron Radiation Research Section, National Cancer Institute, Brookhaven National Laboratory, Upton, New York 11973*

*Received November 15, 2002; Revised Manuscript Received February 7, 2003*

ABSTRACT: Glutaryl 7-aminocephalosporanic acid acylase (GCA, EC 3.5.1.11) is a member of N-terminal nucleophile (Ntn) hydrolases. The native enzyme is an $(\alpha\beta)_2$ heterotetramer originated from an enzymatically inactive precursor of a single polypeptide. The activation of precursor GCA consists of primary and secondary autoproteolytic cleavages, generating a terminal residue with both a nucleophile and a base and releasing a nine amino acid spacer peptide. We have determined the crystal structures of the recombinant selenomethionyl native and S170A mutant precursor from *Pseudomonas* sp. strain GK16. Precursor activation is likely triggered by conformational constraints within the spacer peptide, probably inducing a peptide flip. Autoproteolytic site solvent molecules, which have been trapped in a hydrophobic environment by the spacer peptide, may play a role as a general base for nucleophilic attack. The activation results in building up a catalytic triad composed of Ser170/His192/Glu624. However, the triad is not linked to the usual hydroxyl but the free α-amino group of the N-terminal serine residue of the native GCA. Mutagenesis and structural data support the notion that the stabilization of a transient hydroxazolidine ring during autoproteolysis would be critical during the N → O acyl shift. The autoproteolytic activation mechanism for GCA is described.

The glutaryl 7-aminocephalosporanic acid (7-ACA)[1] acylase (GCA; EC 3.5.1.11) of *Pseudomonas* sp. strain GK16 is an amide hydrolase that produces 7-ACA from glutaryl 7-ACA (Scheme 1) (*1*, *2*). It belongs to cephalosporin acylases, a group of enzymes that hydrolyze cephalosporin C. 7-ACA is a key intermediate for the synthesis of the clinically important semisynthetic cephem antibiotics (*3*−*5*). Sequence alignment analysis suggests that GCA belongs to type I cephalosporin acylase (CA-I), which shows gene structures very similar to those of penicillin G acylase (PGA) (*6*). In addition, GCA is a member of the N-terminal nucleophile (Ntn) hydrolase superfamily, including PGA, cephalosporin acylase (CA), glutamine PRPP amidotransferase (GAT), proteasome $\beta$ subunit (PRO), and glycosylasparaginase (GA) (*7*−*9*). This family is characterized by a common structural $\alpha\beta\beta\alpha$ sandwich motif and an N-terminal nucleophilic amino acid: Ser in PGA and CA, Cys in GAT, and Thr in PRO and GA. Moreover, these family member proteins autoproteolyze internal peptide bond rearrangement through an N → O or N → S acyl shift (*10*, *11*).

Scheme 1: Schematic Diagram of the Overall Reaction of GCA[a]




[a] The substrate, glutaryl 7-ACA, is hydrolyzed by glutaryl 7-ACA acylase (GCA) at the amide bond to form glutaric acid and 7-ACA.

The nascent GCA is synthesized as a 74 kDa polypeptide containing sequences coding for a signal peptide, 16 kDa α subunit, 9 amino acid spacer peptide, and 54 kDa $\beta$ subunit. The removal of the signal peptide gives rise to a 71 kDa precursor polypeptide to be processed by proteolytic cleavages into α and $\beta$ subunits in the periplasm (*12*). The precursor protein is activated through two sequential steps of autoproteolytic cleavages upon folding (Scheme 2): the first step is an intramolecular cleavage of the precursor between Gly169 and Ser170 for generation of the α subunit

† This research was supported by a grant from the Korea Basic Research Institute (R23-2001-00005). K.H.K. was supported by KOSEF for visiting Japan and by the MOST and British Council for visiting the United Kingdom. J.K.K. and I.S.Y. were supported by Brain Korea 21 from the Ministry of Education.

* Corresponding author. Tel: 82-2-3290-3444. Fax: 82-2-927-9028. E-mail: khkim@korea.ac.kr.

‡ Korea University.

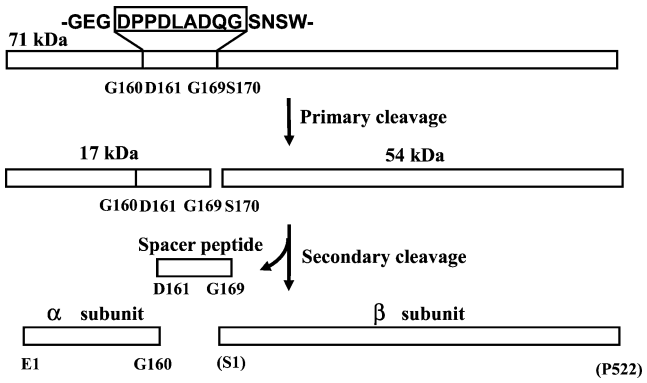
§ Seoul National University.

|| Brookhaven National Laboratory.

[1] Abbreviations: 7-ACA, 7-aminocephalosporanic acid; GCA, glutaryl 7-aminocephalosporanic acid acylase; Ntn, N-terminal nucleophile; CA, cephalosporin acylase; GAT, glutamine PRPP amidotransferase; PRO, proteasome $\beta$ subunit; GA, glycosylasparaginase; SeMet, selenomethionyl; rmsd, root-mean-square deviation; FPLC, fast-performance liquid chromatography; MALDI, matrix-assisted laser desorption ionization.

10.1021/bi027181x CCC: $25.00 
Published on Web 03/22/2003

Scheme 2: Schematic Diagram of the Two-Step Autocatalytic Processing of GCA[a]




[a] The 71 kDa precursor enzyme is activated through an intramolecular cleavage between Gly169 and Ser170 and a subsequent intermolecular cleavage between Gly160 and Asp161 to generate a native $(\alpha\beta)_2$ heterotetramer.

containing the spacer peptide and the $\beta$ subunit; the second step is an intermolecular cleavage, with the spacer peptide released, and the subunits assemble into an active heterotetramer $(\alpha\beta)_2$. Mutagenesis and reconstitution studies demonstrated that the removal of the spacer peptide is required for GCA activity and the invariant Ser170 becomes a newly generated N-terminal residue of the $\beta$ subunit of the native enzyme (*12*, *13*). Most importantly, its side chain hydroxyl group plays an essential role in both autoproteolytic activation and enzyme catalysis.

Precursor structures are essential to understand the mechanism of autoproteolytic reactions that proceed via an N → O or N → S acyl shift. Although three-dimensional structural information is available for some Ntn hydrolase precursors (*14*−*17*), precursor activation mechanisms are not clearly understood. In particular, the exact autoproteolytic potential is far from complete. A highly strained tight turn conformation with unusual geometry at the scissile peptide bond has been suggested to be in part a driving force for GA precursor activation (*16*). We determined the crystal structures of the selenomethionyl (SeMet) native and S170A mutant precursor GCA from *Pseudomonas* sp. strain GK16 by X-ray crystallography and refined to 2.0 and 2.5 Å resolution, respectively. In this report we describe the autoproteolytic activation potentials that include a conformational constraint in the spacer peptide, the stabilization of the hydroxazolidine transition state, and sequestering solvent molecules in a hydrophobic environment. A detailed mechanism for peptide bond rearrangement during the N → O acyl shift is also described.

## MATERIALS AND METHODS

*Preparation of Recombinant Proteins*. The pSH plasmid harboring the cloned glutaryl 7-ACA acylase gene from *Pseudomonas* sp. strain GK16 was described previously (*12*). It was used to transform Met auxotrophic *Escherichia coli* strain B834(DE3). SeMet GCA was overexpressed in a defined minimal medium (without methionine) including DL-selenomethionine. S170A mutant GCA was prepared by site-directed mutagenesis and subcloned into *E. coli* BL21(DE3) with the overexpression vector pET23d(+). GCA proteins were purified using TALON metal affinity chromatography (Clontech), Resource Q ion-exchange FPLC, and phenyl-Sepharose chromatography (Pharmacia).

*Protein Characterization*. The native GCA from *Pseudomonas* sp. GK16 is a $(\alpha\beta)_2$ heterotetramer as measured by both gel filtration chromatography and native PAGE (*12*, *18*). The active GCA tetramer from *Pseudomonas* sp. is dissociated into dimers at a low denaturant concentration and low pH, whose enzyme activity remains about 10% (*19*). Single amino acid mutation of Ser170 to Ala affects correct autocatalytic processing which occurs in cis in a folding-dependent manner, generating an enzymatically inactive precursor (*12*). SDS−PAGE analysis demonstrated a single protein band of the S170A mutant at 71 kDa, and its CD spectra are identical to those of the native GCA (data not shown), suggesting that the lack of autoprocessing is not the result of misfolding. Instead, the lack of a nucleophilic group in the S170A mutant may prevent the processing, to mimic precursor GCA (*12*, *18*).

*Crystal Preparation.* Crystals of the SeMet native and alanine mutant GCA were grown at 20 °C by the hanging-drop vapor diffusion method using ammonium sulfate as a precipitating agent. Earlier attempts to crystallize GCA resulted in tetragonal crystals with an $\alpha\beta$ dimer in the asymmetric unit (*18*). A different crystal form was then obtained, when the reservoir solution contained 1.6 M ammonium sulfate, 0.1 M HEPES (pH 7.5), and 0.1 M NaCl for the native and 1.5 M ammonium sulfate, 0.1 M Tris-HCl (pH 8.5), and 12% glycerol for the precursor GCA. Crystals obtained from the proteins were monoclinic (space group *C*2) and contained two molecules in the asymmetric unit, an $(\alpha\beta)_2$ heterotetramer positioned at the crystallographic 2-fold axis. Crystals were flash-cooled at 100 K by transferring them into a reservoir solution containing 20% glycerol. The unit cell dimensions of the precursor crystals were similar to those of the mature native enzyme, with a decrease of about 4 Å in the *a* axis and about 1 Å in *b* and *c* (Table 1).

*Data Collection.* Diffraction data were collected using synchrotron radiation sources, NSLS beamline X9B at Brookhaven National Laboratory and PLS beamline 6B at Pohang Light Source. Three wavelength data sets for SeMet GCA at the peak, at the inflection point, and remote from the absorption edge were obtained at 1.5, 2.0, and 2.0 Å resolution, respectively, and the data set for precursor GCA was obtained at 2.5 Å resolution. All data were indexed and scaled (Table 1) using DENZO and SCALEPACK (*20*).

*Experimental Phases.* Nineteen positions out of the 24 selenium atoms expected from the number of methionine residues in the amino acid sequence were initially calculated by direct and reciprocal-space Patterson searches in CNS (*21*). The complete 24 selenium sites were obtained by difference Fourier, based on dispersive as well as anomalous differences. The anomalous difference Fourier was sufficient to determine the five missing sites with outstanding correlation coefficients (higher than 12.1$\sigma$). The heavy-atom parameters were then refined to give the mean figure of merit of 0.56 using data to 2.4 Å resolution. Phase refinement was followed by DM (*22*), using solvent flipping and histogram matching, which provided a final figure of merit of 0.85 at 2.0 Å resolution. This map showed clear density for the overall structure of GCA.

Table 1: Crystallographic and Refinement Statistics

| | SeMet GCA | | | |
|---|---|---|---|---|
| | inflection ($\lambda_1$) | peak ($\lambda_2$) | remote ($\lambda_3$) | S170A precursor |
| data collection statistics | | | | |
| wavelength (Å) | 0.9810 | 0.9800 | 0.9680 | 1.0629 |
| resolution (Å) | 2.0 | 1.5 | 2.0 | 2.5 |
| observations | 436357 | 1360035 | 374135 | 248978 |
| unique reflections | 120889 | 278106 | 118490 | 59280 |
| completeness (%) | 99.5 | 97.6 | 99.2 | 99.4 |
| $R_{sym}$(%)[a] | 6.0 | 7.5 | 6.3 | 7.5 |
| $I/\sigma(I)$ | 12.9 | 9.9 | 12.2 | 9.8 |
| unit cell parameters | | $a = 229.7$ Å, $b = 69.9$ Å, $c = 113.7$ Å, $\beta = 97.6°$ | | $a = 225.4$ Å, $b = 68.8$ Å, $c = 112.8$ Å, $\beta = 97.5°$ |
| MAD phasing power[b] | 1.75/3.37 | 1.13/2.14 | 1.94/3.25 | |
| figure of merit[c] | | 0.56/0.85 | | 0.68/0.83 |
| refinement statistics | | | | |
| resolution range (Å) | | 20.0−2.0 | | 30.0−2.5 |
| no. of reflections used in refinement | | 110770 | | 56456 |
| no. of reflections used in $R_{free}$ | | 4488 (4%) | | 2290 (4%) |
| $R_{cryst}$ (%)[d] | | 18.3 | | 17.7 |
| $R_{free}$ (%)[e] | | 20.6 | | 21.8 |
| rmsd, bond distances (Å) | | 0.005 | | 0.006 |
| rmsd, bond angles (deg) | | 1.3 | | 1.3 |
| average *B*-factors (Å$^2$) | | 14.5 | | 21.1 |
| no. of water molecules | | 1256 | | 866 |
| Ramachandran plot quality | | | | |
| most favored region (%) | | 88.2 | | 87.3 |
| additionally allowed region (%) | | 11.3 | | 11.8 |
| generously allowed region (%) | | 0.2 | | 0.5 |
| disallowed region (%) | | 0.4 | | 0.3 |

[a] $R_{sym} = \sum_h \sum_i |I_i(h) - \langle I(h)\rangle| / \sum_h \sum_i I_i(h)$, where $I_i(h)$ is the $i$th measurement and $\langle I(h)\rangle$ is the mean of all measurements of $I(h)$ for Miller indices $h$. [b] Phasing power = $[|F_H(\text{calc})|^2/|\text{lack of closure error}|^2]^{1/2}$, where $F_H$ is the heavy-atom structure amplitude (average/maximum for acentric reflections). [c] Overall figure of merit was used for computing the initial map. [d] $R_{cryst} = \sum |F_o - F_c| / \sum |F_o|$ calculated for all observed data. [e] $R_{free} = \sum |F_o - F_c| / \sum |F_o|$ calculated for a specified number of randomly chosen reflections that were excluded from the refinement.

*Model Building and Refinement.* Density-modified phases from CNS and 2.0 Å resolution data were used to initiate building of dummy models in ARP/wARP (*23*). The main chain was almost completely traced, with the exception of terminal residues and loop regions, and its side chains were placed into the density. Manual building of remaining residues and their side chains was conducted in O (*24*). The initial model was composed of two $\alpha\beta$ dimers in the asymmetric unit that are related by a translation of $^1/_2$ along the *c* axis and a rotation of 2.2°. An $(\alpha\beta)_2$ heterotetrameric molecule could be generated by crystallographic 2-fold symmetry. Crystallographic refinement was subsequently carried out using CNS (*21*). After a few rounds of model rebuilding, the final model of the heterotetramer comprises residues Gln8−Gly160 of the α subunit, residues Ser1−Pro522 of the $\beta$ subunit, and 1256 water molecules in the asymmetric unit (Table 1). The stereochemistry of the refined structure was analyzed by PROCHECK (*25*), revealing that four prolines (Pro131 of the α subunit and Pro253, Pro379, and Pro466 of the $\beta$ subunit) adopt a cis peptide conformation and one residue, Phe177 of the $\beta$ subunit, is in a disallowed region of the Ramachandran plot.

The S170A mutant precursor structure was solved by molecular replacement using the refined polyalanine model of the native GCA as a template, which was carried out in CNS (*21*). Density modification resulted in a figure of merit of 0.83 with 2.5 Å resolution data. The map was clearly interpretable for the entire structure. Crystallographic refinement was carried out using CNS (*21*). Refinement statistics and geometry of the model that comprises residues Pro6−Pro691 of the unprocessed single polypeptide chain and 866 water molecules are also shown in Table 1. As shown in its native structure, four proline residues, Pro131 (Pro131),[2] Pro422 (Pro253), Pro548 (Pro379), and Pro635 (Pro466), assume a cis peptide conformation and Phe346 (Phe177) was in a disallowed region of the Ramachandran plot. We rebuilt the Ser residue in place of Ala in the precursor structure. The side chain orientation of Ser170 was modeled to reflect the native conformation and corresponded to one of the most frequently observed rotamer conformations. Two other most frequently observed rotamer orientations caused steric hindrance with neighboring atoms at the autoproteolytic site. The modeled side chain directs its OG toward the scissile carbonyl carbon atom in the Gly169−Ser170 peptide bond, so that it could follow the preferred orientation of a nucleophilic addition reaction (*26*). The rebuilt model was further subjected to energy minimization using CNS (*21*). The coordinates of the SeMet native and precursor GCA have been deposited in the Protein Data Bank (PDB ID codes 1OR0 and 1OQZ).

## RESULTS AND DISCUSSION

*Overall Structures.* The native GCA heterodimer folds into a compact structure containing the $\alpha\beta\beta\alpha$ motif. This structural core is formed by two $\beta$ sheets packed against each other, which are sandwiched by two layers of α helices, constructing a heterotetramer. The precursor GCA has a

[2] The precursor has been numbered as a single chain (residues 1−691) comprising the α subunit (1−160) and the spacer region (161−169) followed by the $\beta$ subunit (170−691). The amino acid residue of the native is numbered in parentheses, if applicable.

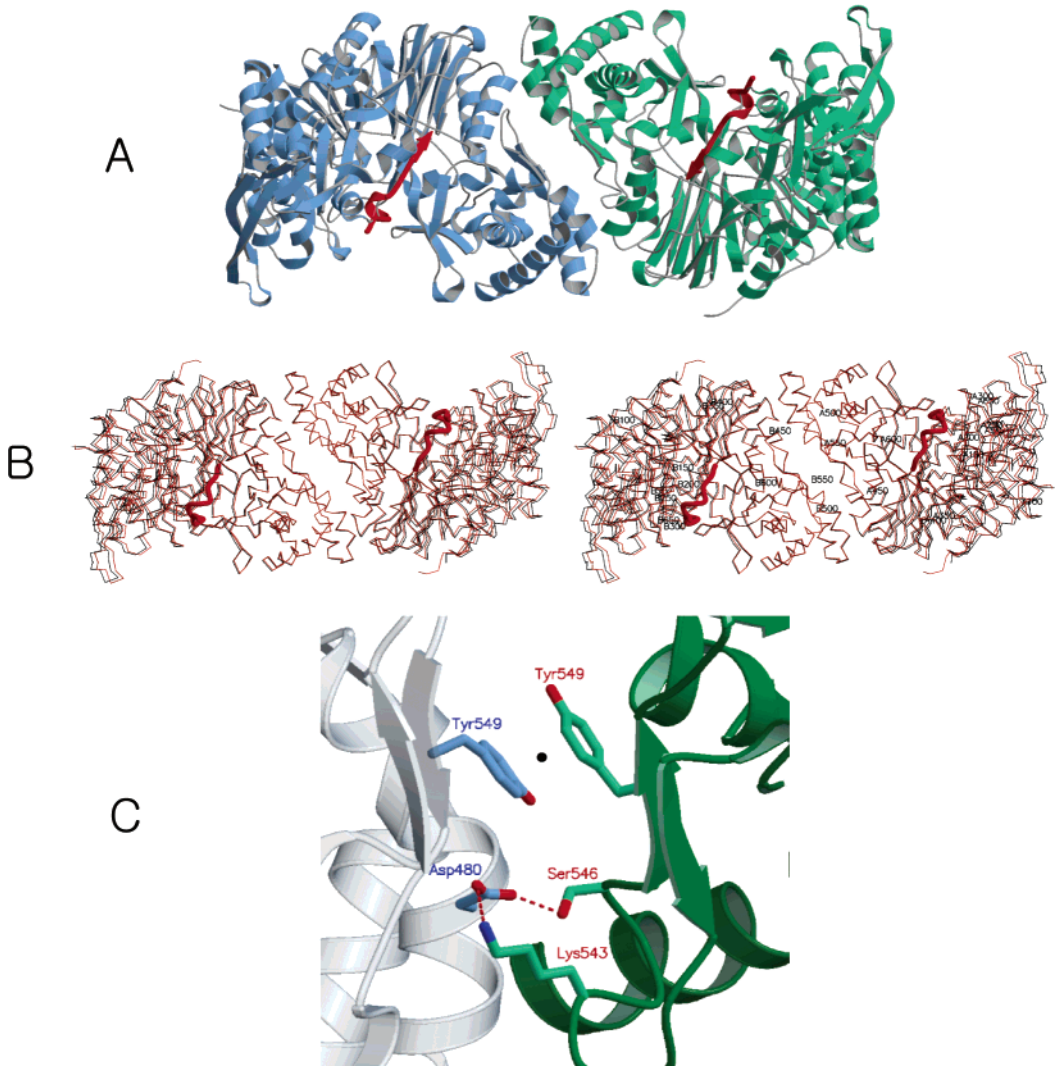



FIGURE 1: (A) Dimer structure of the GCA precursor from *Pseudomonas* sp. GK16. The precursor consists of two monomers related by a crystallographic 2-fold axis. The spacer peptide which is autoproteolytically cleaved out is colored in red. (B) Superimposed structures of the native (black) and precursor GCA (brown). The spacer peptide is shown with a thicker line in red. Every 50th residue is numbered. (C) One of the interface contact regions between two monomers. The position of the crystallographic 2-fold axis relating the two monomers passing between the two aromatic rings of Tyr549 is shown as a black mark.

structure very similar to that of the native GCA, except for the spacer peptide that connects the C-terminal end of the α subunit and the N-terminal end of the $\beta$ subunit of the native enzyme (Figure 1A). The root-mean-square deviation (rmsd) upon superposition of the native $(\alpha\beta)_2$ heterotetramer and the precursor dimer structures is 0.68 Å for 1346 common Cα atoms (Figure 1B). One monomer of the precursor is essentially identical to that of the native $\alpha\beta$ heterodimer (rmsd = 0.24 Å for 673 Cα atoms). The sandwich structural motif is highly conserved in the Ntn hydrolase superfamily, as also demonstrated in recently determined structures of glutarylamidase from *Pseudomonas* sp. and CA from *Pseudomonas diminuta* (*7*−*9*). The rmsd upon superposition of the native GCA and glutarylamidase structures is merely 0.81 Å for 1345 common Cα atoms. A superposition of one monomer of the precursor GCA and the monomeric CA precursor structure (*14*) is 0.52 Å for 679 common Cα atoms.

The two monomers of the precursor form a tight contact, shielding solvent-accessible surface area of 1935 Å$^2$. Dimerization of the precursor sequesters a slightly larger surface area, compared with 1868 Å$^2$ in the native structure. Three helical regions (Arg449−Asn454, Pro479−Ile487, and Arg498−Ala501) and portions of the loop (Trp421−Pro422, Lys543−Ser546, and Val551-Pro554) and strand (Ala536−Thr537 and Tyr549−Val551) regions participate in this monomer−monomer contact to which van der Waals interactions contribute substantially. In particular, two aromatic rings of Tyr549 from each monomer are packed at a distance of about 3.6 Å, where the crystallographic 2-fold axis passes through (Figure 1C). The side chain of Asp480 of one monomer forms a salt bridge with that of Lys543 and hydrogen bonds to Ser546 OG of the other monomer. These structural features are also found in the native structure, and the residues at the interface are highly conserved in the type I CA (CA-I) enzymes (*6*). The surfaces of the spacer peptide region in the precursor exhibit negative potentials, whereas the native enzyme shows a partly positive surface potential in the absence of the spacer peptide and a solvent-accessible hole to the active site (data not shown). A negative potential patch is also found on the diametrically opposite end of the spacer surface. Precursor dimerization has been implicated as an important factor for the autoproteolysis of GA that is a member of Ntn hydrolase (*28*, *29*). His124 at the interface between two heterodimers of GA is crucial for the thermo-

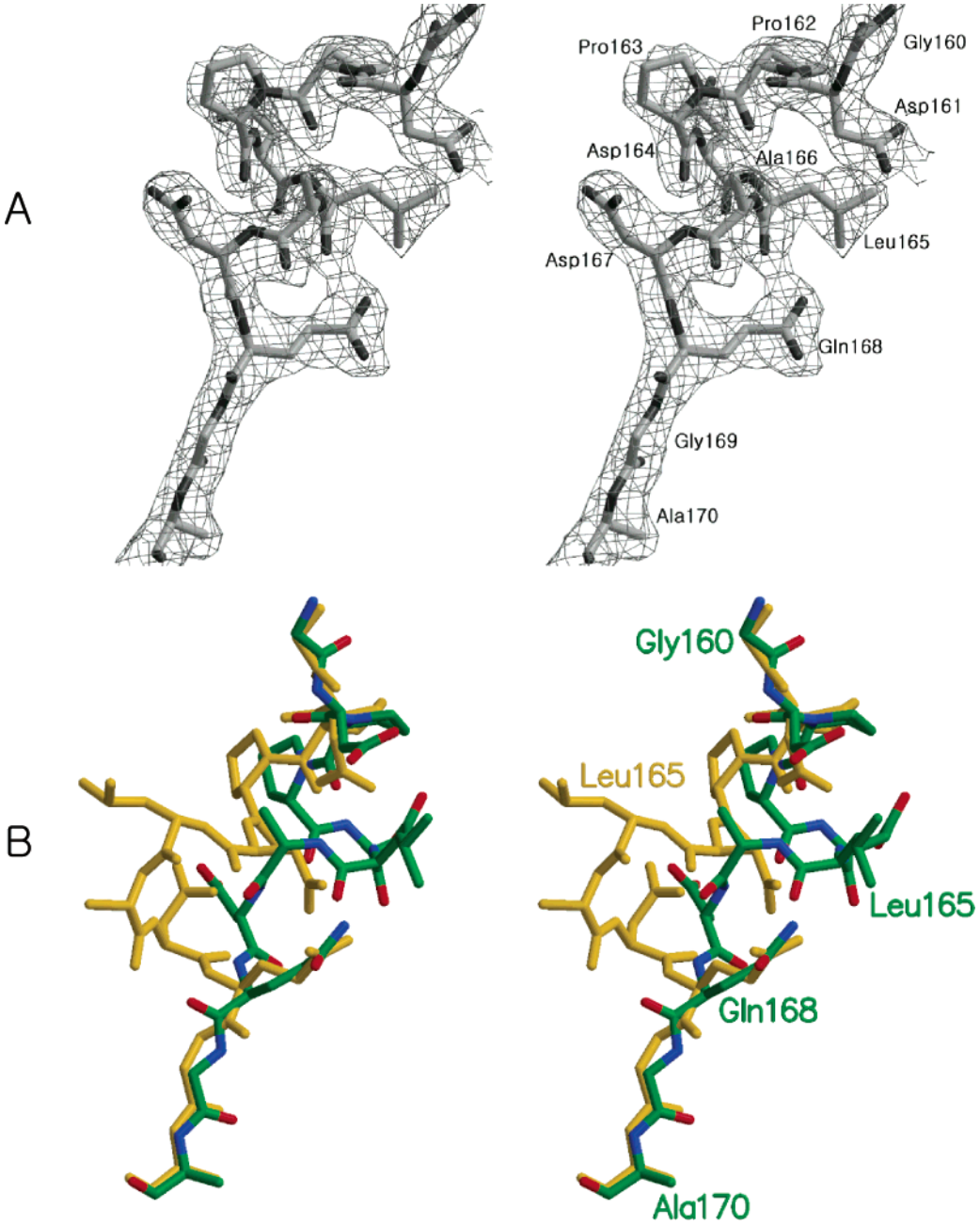



FIGURE 2: (A) Stereoview of the electron density map and a model for the spacer peptide of the precursor GCA. The map is of the $2F_o - F_c$ type at 2.5 Å resolution and contoured at $1.2\sigma$. (B) Superimposition of the spacer peptides of the dimeric GCA (C, N, and O atoms are colored in green, blue, and red, respectively) and the monomeric CA (dark gold).

dynamically stable and mature oligomeric structure. In precursor GCA, the negative potential patches may guide its monomers in side-by-side orientation using electrostatic repulsion. The interactions of the aromatic $\pi-\pi$ stacking and salt bridges at the interface might be important for the correct positioning of dimerization of the GCA precursor and thus for the assembly of the native heterotetramer.

*Conformation of the Autoproteolyzed Spacer Peptide of Precursor GCA.* A novel feature of the nine amino acid spacer peptide is that it adopts one and a half turns of α-helical conformation from Pro162 to Ala166 and a $\beta$ strand conformation from Gln168 to Ala170 (Figure 2A), which extends to Val175. It was noted earlier in GA that the interlayer loops cluster at the functional side of the structure for autoproteolysis and hydrolase activity (*16*). Two interlayer loops are found at the functional side of GCA, and the spacer peptide including the autoproteolytic site is located in the longer interlayer loop. The overall *B*-factors for this spacer region are between 15 and 35 Å$^2$, with the higher mobility near the loop region. Superimposition of the dimeric GCA and monomeric CA spacer peptides reveals significant differences (Figure 2B), although the amino acid sequences are identical in this region. The CA spacer peptide shows a bend conformation from Asp164 to Asp167, whereas the GCA spacer peptide has α-helical conformation from Pro162 to Ala166. The differences in the $\phi$ and $\psi$ values of the spacer peptides between GCA and CA also demonstrate the major changes of the angles in three segments of Gly158−Asp161, Pro163−Ala166, and Gln168−Gly169 (data not shown). These segments are involved in secondary processing, geometric constraints, and peptide flipping, respectively (see below).

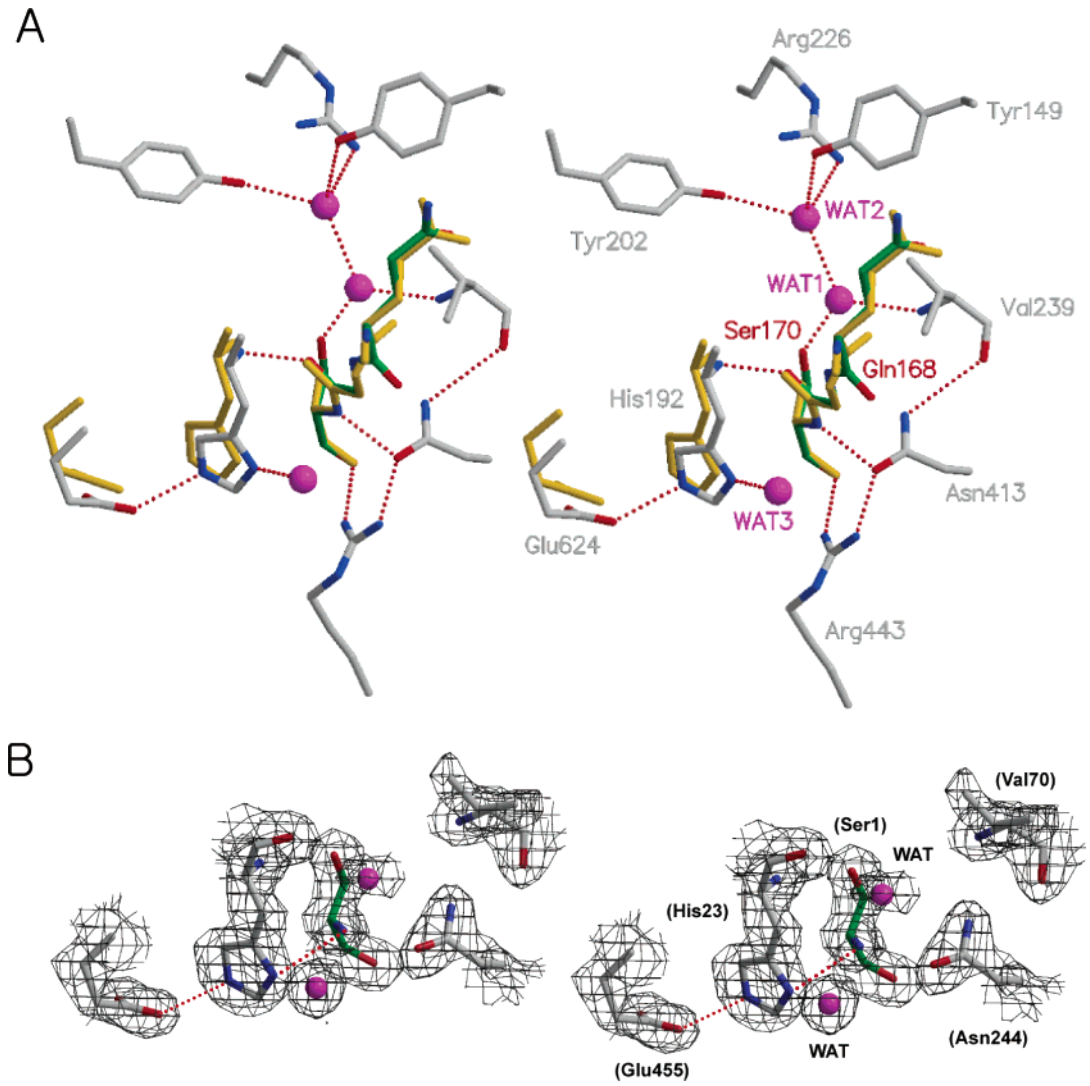



FIGURE 3: (A) Stereoview of the autoproteolytic site of the precursor GCA (C, N, and O atoms are colored in gray, blue, and red, respectively, and C atoms of the precursor GCA spacer peptide are colored in green). The catalytic triad and part of the spacer peptide of the precursor CA are superimposed (dark gold). The Ala170 residue was modeled as Ser170 on the basis of the native conformation, and its side chain corresponded to one of the most frequently observed rotamer conformations. Solvent molecules are colored in magenta, and dotted lines indicate putative hydrogen bonds. (B) Stereoview of the electron density map and a model of the catalytic triad found in the native GCA. This map is of the $2F_o - F_c$ type at 2.0 Å resolution and contoured at 1.1$\sigma$. Dotted lines indicate putative hydrogen bonds for the triad, and solvent molecules are colored in magenta.

Antiparallel $\beta$-sheet is formed between the strand from Asp167 to Ala170 and the strand from Pro191 to Ser194. However, it is absent in the native structure, and the site is occupied by ordered solvent molecules, bound to the main chain atoms of the flanking strand. The main chain interaction between Asp167 and Ser194 is facilitated by one of the bound solvent molecules. The amide N of His192 forms a hydrogen bond with the carbonyl oxygen at the scissile bond between Gly169 and Ser170 in the precursor structure, involved in extensive $\beta$ sheet hydrogen-bonding interactions.

*Autoproteolytic Site of Precursor GCA.* Comparison between the native and precursor molecules shows that the substrate binding cavity is shielded by the spacer peptide at the autoproteolytic site. The Ala170 residue of the S170A mutant precursor GCA was modeled as Ser170, to reflect the native conformation and correspond to one of the most frequently observed rotamer conformations. Ser170, an essential nucleophile in autoproteolysis as well as catalysis, is entirely buried in the precursor structure (Figure 3A). After the primary autocleavage, however, it becomes located at the end of the $\beta$ strand in the extended $\beta$ sheet core and thus exposed to solvent, with a solvent-accessible surface area of about 12 Å$^2$ in the native structure (compared to 55−65 Å$^2$ for an entirely exposed Ser). One of the solvent molecules is found to bridge these amino and hydroxyl groups in the native structure, which appears to play a role in acylase activity. In contrast, its accessibility is blocked by the presence of Gly169 in the precursor structure. Therefore, the differential burial of the residue Ser170 and the replacement of the residues from Asp167 to Gly169 with ordered solvent molecules may gain an access to the substrate recognition cleft and provide a catalytic advantage to the native enzyme.

One interesting feature of the autoproteolytic site of the precursor GCA is that the carboxylic group of Glu624 makes a hydrogen bond with the imidazole ring of His192 (3.0 Å). However, the hydrogen bond between Ser170 and His192 does not appear to exist in the precursor (3.6 Å), whereas it is formed in the native structure (3.1 Å) to underline a catalytic triad of Ser-His-Glu, as observed in glutarylamidase

(*7*) (Figure 3B). Contrary to our expectations of the catalytic triad of serine proteases (*30*), there is a difference in the specific nature of the triad found in native GCA. The hydrogen bond between Ser1 and His23 involves the free α-amino group of the serine residue, not the usual hydroxyl group. The side chain of His23 in the native structure may contribute to reduce the p$K_a$ of the α-amino group of Ser1, which then acts as a general base to deprotonate its own hydroxyl oxygen involved in acylase catalysis (*7*, *13*). Our structures thus suggest that this catalytic triad is not required for precursor activation but is an outcome of the activation. It is interesting that the main chain nitrogen of Ser170 (Ser1) interacts with the side chain of Asn413 (Asn244) which hydrogen bonds to Arg443 (Arg274) in the native structure and the precursor structure with the modeled Ser170 (Figure 3A). This hydrogen bond network may support a structural framework of the autoproteolytic site and further the active site of GCA, whereas the catalytic triad on the opposite side provides a functional role.

*Autoproteolytic Site Waters and Activation Potential*. One structural feature that drew our attention was the presence of WAT1 and WAT2 trapped inside by the spacer peptide at the autoproteolytic site. The solvent molecules underneath the spacer peptide are sequestered in a hydrophobic environment, surrounded by mainly nonpolar amino acid residues (Leu193, Phe227, Val239, and Phe346). In the monomeric CA precursor structure, a solvent molecule forms hydrogen bonds with the main chain NH of Val239, the OD1 of Asn413, the carbonyl oxygen of Gln168, and the OG of Ser170. It is stabilized by four hydrogen bonds in pseudotetrahedral geometry and provides a base role to assist the nucleophilic attack by the hydroxyl group of the serine (*14*). In dimeric precursor GCA, there are no charged or hydrophilic groups within 3.4 Å of the hydroxyl group of Ser170, except for WAT1. This solvent molecule, close to the hydroxyl group of Ser170 (2.6 Å), sits 3.0 Å from the Val239 amide nitrogen and 3.1 Å from WAT2 (Figure 3A). WAT2 is hydrogen bonded to Tyr149, Tyr202, and Arg226. However, the carbonyl oxygen of Gln168 is oriented away from WAT1. A peptide flip between Gln168 and Gly169, as modeled in the monomeric CA, would reorient the carbonyl group closer to WAT1 for a possible hydrogen bond. It would also bring Gln168 from a $\beta$ strand to an α helix area of the Ramachandran plot (from 140.6° to −61.2° in $\psi$). However, this peptide flip is less favorable in the current GCA precursor structure, demonstrated by a negative peak in the difference Fourier map. Interestingly, the α-helix (Pro162−Ala166) in the spacer peptide appears to contain a conformational strain on the backbone, in that the $\phi$ angle of Leu165 ($\phi$ −103.0° and $\psi$ −46.0°) deviates significantly from ideality. The backbone angle N−Cα−C ($\tau$ angle) of Asn171 also deviates more than 3$\sigma$ from ideality, but the planarity of the scissile peptide bond is refined close to ideality. One unusual turn of the GA precursor creates conformational strain on the backbone, in that the planarity of the scissile peptide bond between Asp151 and Thr152 and the $\tau$ angle of Asp151 deviate significantly from ideality (*16*). In the monomeric CA precursor, the $\tau$ angles of Gln168 and Gly169 deviate more than 3$\sigma$ from ideality (*14*).

A highly strained and distorted trans configuration at the autocleavage site was suggested to generate the activation potential for autoproteolytic cleavage (*16*). The pro segment of the PRO T1A mutant $\beta$ subunit adopts a well-defined $\gamma$ turn conformation at the autoproteolytic site (*17*), and the modeled Thr1 side chain is positioned over the $\gamma$ turn bulge. Thus, addition of the nucleophilic hydroxyl group to the autocleavage site may proceed by very small motions. We searched for conformations similar to the spatial motif composed of the spacer peptide (residues Asp161−Gly169) in the Protein Data Bank database (*31*). The best matches were residues 61−69 of 1-aminocyclopropane-1-carboxylate synthase (*32*; PDB code 1b8g; rmsd = 0.70 Å), residues 110−118 of homogentisate 1,2-dioxygenase (*33*; PDB code 1eyb; rmsd = 0.80 Å), and residues 69−77 of bacterial xylanase (*34*; PDB code 1exg; rmsd = 0.84 Å). However, no backbone distortion was found in these conformations. The conformational constraint in precursor GCA may be confined within the spacer peptide during precursor folding. Conformational changes in proteins are made possible by their intrinsic flexibility. These changes may occur with only relatively small expenditure of energy. It should be noted that this spacer region has an intrinsic flexibility with high *B*-factors. Furthermore, we cannot rule out the possibility that the current peptide bond geometry between Gln168 and Gly169 may have been created by the S170A mutation. Nevertheless, the plausible peptide flip and geometric constraints seem to be crucial for the activation of the GCA precursor polypeptide. In this connection, it is very likely that a high-energy state of the spacer peptide due to conformational constraints may trigger the autoproteolytic cleavage in GCA.

On the opposite side of WAT1 with respect to Ser170, another water molecule, WAT3, forms a strong hydrogen bond with ND1 of His192 (2.7 Å) (Figure 3A). Its position is well conserved in both the native and precursor structures, and its electron density level is comparable to that of protein side chains. We postulate that WAT3 may carry out the second nucleophilic attack on the ester intermediate. First, mutant protein H192S in GCA has been shown to completely impair autocleavage (unpublished results), suggesting an important role of the imidazole ring and plausibly WAT3. Second, all mutants at position 169 in the acylase from *Pseudomonas* sp. 130 belonging to CA-I showed no processing or acylase activity, suggesting that glycine at position 169 is the critical and primary determinant of the precursor processing (*6*). In fact, any side chain would create a steric clash with WAT3 in our precursor structure. Third, based on the comparison of solvent molecules at the autoproteolytic site between the precursor and native GCA structures, it is likely that WAT1 is able to donate a proton to the main chain NH of Ser170, a leaving group. In addition, if we consider a dynamic rearrangement during the peptide flip and the following N → O acyl shift, our structures strongly suggest that the carbonyl carbon of the ester intermediate becomes closer to WAT3. Therefore, WAT3 may play a role as a second nucleophile for the ester intermediate breakdown, with the help of the imidazole ring of His192 as a general base. It can be inferred that WAT1 may play a role as a base to initiate a primary nucleophilic attack and then the proximal WAT3/histidine/glutamic acid pseudotriad then carries out a flank attack on the ester intermediate.

*Transition State Complementarity in Cis Autoproteolysis.* During the course of the cis-autoproteolytic reaction, a hydroxazolidine transition state is produced and partly

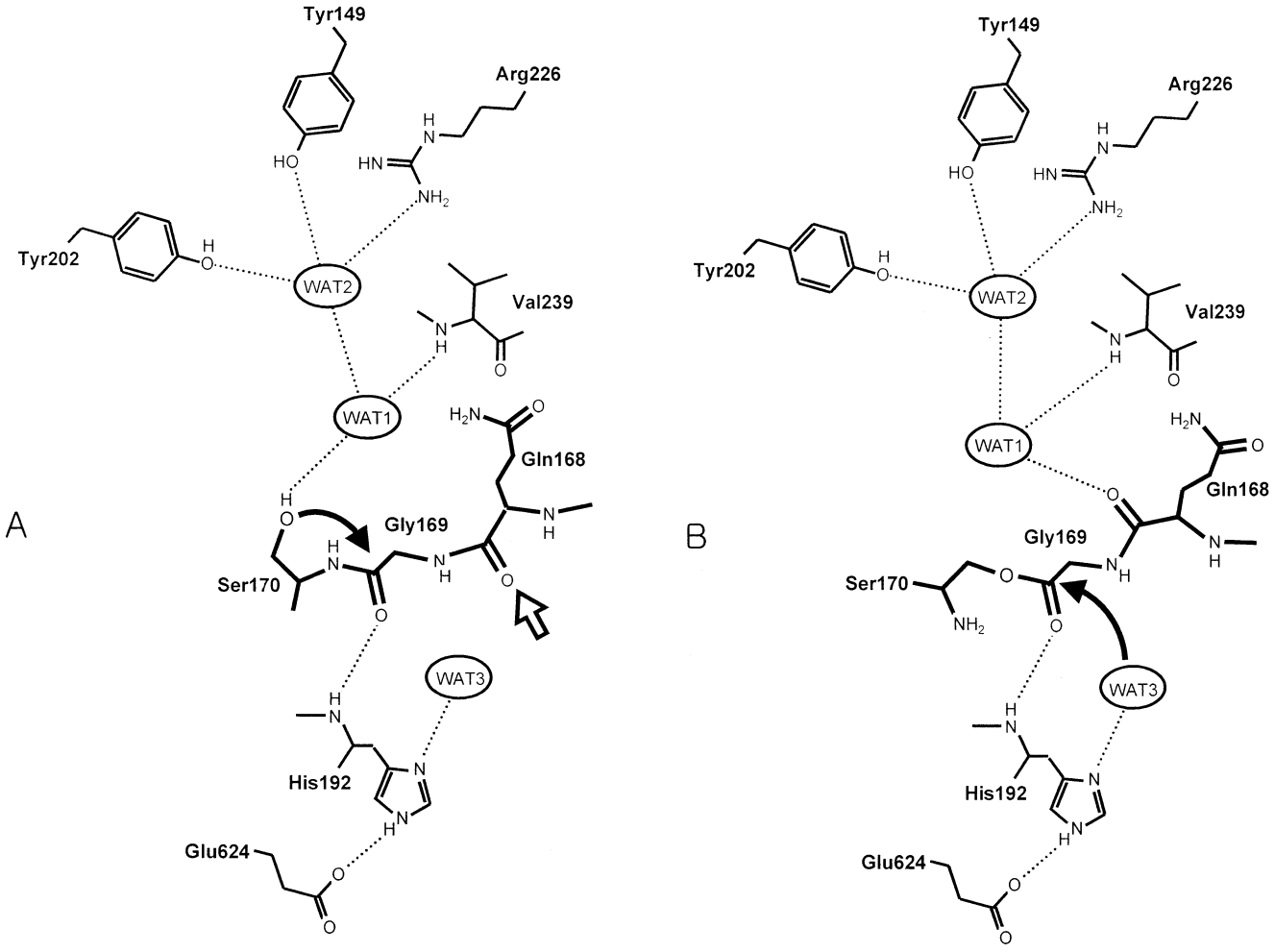



FIGURE 4: Proposed autoproteolytic mechanism for the precursor GCA. The model is based on the refined structures of native and precursor GCA, with some modification (see the text). Dotted lines indicate hydrogen bond interactions. (A) The N → O acyl shift producing an ester linkage. An open arrow represents the plausible peptide flip site and a filled arrow a nucleophilic attack. (B) Hydrolysis of the ester intermediate, finishing the cleavage of the C-terminus of the spacer peptide.

stabilized by the oxyanion hole consisting of the main chain nitrogen of His192. This is featured in the hydrogen bond between Gly169 and His192 in the antiparallel $\beta$ sheet conformation of the precursor, one of the key differences between the native and precursor GCA. This transient hydroxazolidine is composed of Ser170 (backbone N, $C_\alpha$, $C_\beta$, and $O_\gamma$) and the carbonyl carbon of Gly169, making a putative five-membered ring. Our structure demonstrates that the side chain OD1 of Asn413 is very close to form a strong hydrogen bond with the nitrogen atom of the hydroxazolidine (2.6 Å). It is generally conceived that the transition state is stabilized by an oxyanion hole in cis-autoproteolysis, as observed in serine protease. In addition to this oxyanion hole, the stabilization of the hydroxazolidine ring by Asn413 OD1 would be critical for the acylation step during the N → O acyl shift. Mutation data have been shown to support this notion: the N413D mutant protein is autoproteolyzed as native GCA, whereas the N413A mutant severely blocks the autocleavage (*3*). Probably more importantly, the immediate amino-terminal residue of the autoproteolytic nucleophile is conserved as glycine in PRO, PGA, GCA, and GAT of the Ntn hydrolases, except for that in GA. Other internal autoproteolytic proteins such as inteins (*35*), aspartate decarboxylase (*36*), and hedgehog protein (*37*) have the same conserved glycine residue as an immediate amino-terminal residue. In case of GA, which has aspartate instead of glycine, the crystal structure shows that the carboxyl side chain has a good internal hydrogen bond with the amide N of Thr153, a nucleophilic residue in GA (*16*), another indication of the stabilization of the putative transient hydroxazolidine ring. Structural features that specifically stabilize the hydroxazolidine as well as oxyanion may be likely candidates for the transition state complementarity in cis-autoproteolysis. Further studies are necessary to determine whether this stabilization is a common feature in autoproteolysis.

*Autoproteolytic Mechanism*. On the basis of our structures, we propose an activation mechanism of precursor GCA during autoproteolysis that utilizes an N → O acyl shift. Initially, WAT1, a sequestered solvent molecule underneath the spacer peptide, is held by hydrogen bonds in pseudotetrahedral geometry. One of the hydrogen bonds could be created by a peptide flip under the influence of the conformational constraints within the spacer peptide. WAT1 is positioned to act as a general base and deprotonate the hydroxyl group of Ser170 to enhance its nucleophilicity (Figure 4A). The nucleophilic attack on the carbonyl carbon of Gly169 by Ser170 OG would result in producing a transition state of the hydroxazolidine ring stabilized by Asn413 OD1 as well as an oxyanion hole of the backbone

nitrogen of His192. Collapse of the transient hydroxazolidine shifts the scissile bond from an amide to an ester bond intermediate (N → O acyl shift), completing the acylation step of the primary processing. A bound solvent molecule WAT3 starts the deacylation step to carry out the second nucleophilic attack on the carbonyl carbon of the newly formed ester intermediate (Figure 4B). This water molecule is hydrogen bonded to the side chain of His192 interacting with the side chain of Glu624, mimicking a catalytic triad to increase the nucleophilicity of the second nucleophile. After the second nucleophilic attack, the stabilization of a tetrahedral intermediate is achieved by the same oxyanion hole of His192. Finally, the ester intermediate will be collapsed, resulting in a free N-terminal serine residue and the C-terminus of Gly169. The spacer peptide is now linked at one end to the C-terminus of the α subunit and probably leaves the center of the autoproteolytic site freely around the active site cleft of the native enzyme. Once the mobile end of the spacer peptide is moving out of the autoproteolytic site, it is assumed to be ready for the secondary autoproteolysis.

After the primary autocleavage of precursor GCA, the secondary cleavage site is located at the distal peptide bond, which lies ~24 Å away from the primary autoproteolytic site. It is unlikely that the terminal Ser can perform a direct nucleophilic attack to this distal peptide bond. The secondary processing site of precursor GCA was identified by MALDI mass spectrometry, revealing that it is located between Gly160 and Asp161, releasing the spacer peptide of 9 amino acids (DPPDLADQG) (*3*). The CA precursor structure from *P. diminuta* revealed a spacer peptide of 11 amino acids (EGDPPDLADQG) (*14*). It is more interesting to find that the spacer peptide of CA from *Pseudomonas* sp. 130 has 10 amino acid residues (GDPPDLADQG) (*8*) and that from *Pseudomonas* C427 has 8 amino acids (PPDLADQG) (*38*). All of these enzymes, however, showed their activity after the removal of their spacer peptide, suggesting that a slight change in spacer length had no effect on the activity of the acylase (*8*). Moreover, the secondary processing site is located in the loop region with high temperature factors, fully exposed to the solvent. The spacer peptide of variable lengths could be accounted for by this high *B*-factor during the secondary processing, which may not be necessarily specific. The discharge of the spacer peptide blocking the substrate binding pocket and the consequent introduction of solvent molecules are prerequisite for cephalosporin acylase activity.

## CONCLUSION

This study reveals an $(\alpha\beta)_2$ quaternary structure which has been observed in solution and crystals of native GCA and a dimeric structure of its precursor. Stacking of aromatic rings in a hydrophobic surface which otherwise are exposed to solvent regions would correctly position the monomeric precursor, together with negative potential patches. The structural difference between native and precursor GCA provides the structural basis for the stabilization of a transient hydroxazolidine and the identification of the second nucleophile based on the WAT/histidine/glutamic acid pseudotriad. Geometry distortions found in the spacer peptide of the precursor could be an important triggering factor in autoproteolysis. Precursor protein folding may generate a high-energy conformational constraint within the spacer peptide, which would then introduce a peptide flip at the autoproteolytic site. A sequestered solvent molecule held by hydrogen bonds near the hydroxyl group of Ser170 may play a role as a general base for autoproteolysis. Therefore, the existence of the spacer peptide in precursor GCA provides two major structural roles in autoproteolysis: (1) the confinement of conformational constraints at the autoproteolytic site and (2) the sequestration of solvent molecules.

## ACKNOWLEDGMENT

The authors thank Prof. Tom Blundell, Dr. J. Patel, and Dr. D. Chirgadze for comments on this project.

## REFERENCES

1. Matsuda, A., and Komatsu, K. (1985) *J. Bacteriol. 163*, 1222−1228.
2. Lee, Y. S., Yang, H. C., and Park, S. S. (1996) *J. Microbiol. Biotechnol. 6*, 375−380.
3. Sykes, R. B., Cimarusti, C. M., Bonner, D. P., Bush, K., Floyd, D. M., Georgopapadakou, N. H., Koster, W. H., Liu, W. C., Parker, W. L., Principe, P. A., Rathnum, M. L., Slusarchyk, W. A., Trejo, W. H., and Wells, J. S. (1981) *Nature 291*, 489−491.
4. Huber, F., Chauvette, R., and Jackson, B. (1972) in *Cephalosporins and penicillins, Chemistry and biology* (Flynn, E., Ed.) pp 27−48, Academic, New York.
5. Ichikawa, S., Murai, Y., Yamamoto, S., Shibuya, Y., Fujii, T., Komatsu, K., and Kodaira, R. (1981) *Agric. Biol. Chem. 45*, 2225−2229.
6. Li, Y., Chen, J., Jiang, W., Mao, X., Zhao, G., and Wang, E. (1999) *Eur. J. Biochem. 262*, 713−719.
7. Fritz-Wolf, K., Koller, K., Lange, G., Liesum, A., Sauber, K., Schreuder, H., Aretz, W., and Kabsch, W. (2002) *Protein Sci. 11*, 92−103.
8. Kim, Y., Yoon, K. H., Khang, Y., Turley, S., and Hol, W. G. J. (2000) *Structure 8*, 1059−1068.
9. Brannigan, J. A., Dodson, G., Duggleby, H. J., Moody, P. C. E., Smith, J., Tomchick, D. R., and Murzin, A. G. (1995) *Nature 378*, 416−419.
10. Lee, Y. S., and Park, S. S. (1998) *J. Bacteriol. 180*, 4576−4582.
11. Lee, Y. S., Kim, H. W., and Park, S. S. (2000) *J. Biol. Chem. 275*, 39200−39206.
12. Paulus, H. (1998) *Chem. Soc. Rev. 27*, 375−386.
13. Perler, F. B. (1998) *Nat. Struct. Biol. 5*, 249−252.
14. Kim, Y., Kim, S., Earnest, T. N., and Hol, W. G. J. (2002) *J. Biol. Chem. 277*, 2823−2829.
15. Hewitt, L., Kasche, V., Lummer, K., Lewis, R. J., Murshudoc, G. N., Verma, C. S., Dodson, G. G., and Wilson, K. S. (2000) *J. Mol. Biol. 302*, 887−898.
16. Xu, Q., Buckley, D., Guan, C., and Guo, H. (1999) *Cell 98*, 651−661.
17. Ditzel, L., Huber, R., Mann, K, Heinemeyer, W., Wolf, D. H., and Groll, M. (1998) *J. Mol. Biol. 279*, 1187−1191.
18. Kwon, T. H., Rhee, S., Lee, Y. S., Park, S. S., and Kim, K. H. (2000) *J. Struct. Biol. 5*, 79−81.
19. Battistel, E., Bianchi, D., Bortolo, R., and Bonoldi, L. (1998) *Appl. Biochem. Biotechnol. 69*, 53−67.
20. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
21. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
22. Abrahams, J. P., and Leslie, A. G. W. (1996) *Acta Crystallogr. D52*, 30−42.
23. Perrakis, A., Sixma, T. A., Wilson, K. S., and Lamzin, V. S. (1997) *Acta Crystallogr. D53*, 448−455.
24. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
25. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thorton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.

26. Burgi, H. B., Dunitz, J. D., and Shefter, E. (1973) *J. Am. Chem. Soc. 95*, 5065−5067.
27. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins 11*, 281−296.
28. Guo, H. C., Xu, Q., Buckley, D., and Guan, C. (1998) *J. Biol. Chem. 273*, 20205−20212.
29. Saarela, J., Laine, M., Tikkanen, R., Oinonen, C., Jalanko, A., Rouvinen, J., and Peltonen, L. (1998) *J. Biol. Chem. 273*, 25320−25328.
30. Dodson, G., and Wlodawer, A. (1998) *Trends Biochem. Sci. 23*, 347−352.
31. Kleywegt, G. J. (1999). *J. Mol. Biol. 285*, 1887−1897.
32. Capitani, G., Hohenester, E., Feng, L., Storici, P., Kirsch, J. F., and Jansonius, J. N. (1999) *J. Mol. Biol. 294*, 745−756.
33. Titus, G. P., Mueller, H. A., Burgner, J., De Cordoba, S. M., Penalva, M. A., and Timm, D. E. (2000) *Nat. Struct. Biol. 7*, 542−546.
34. Xu, G. Y., Ong, E., Gilkes, N. R., Kilburn, D. G., Muhandiram, D. R., Harris-Brandts, M., Carver, J. P., Kay, L. E., and Harvey, T. S. (1995) *Biochemistry 34*, 6993−7009.
35. Perler, F. B., Olsen, G. J., and Adam, E. (1997) *Nucleic Acids Res. 25*, 1087−1093.
36. Albert, A., Dhanaraj, V., Genschel, U., Khan, G., Ramjee, M. K., Pulido, R., Sibanda, B. L., von Delft, F., Witty, M., Blundell, T. L., Smith, A. G., and Abell, C. (1998) *Nat. Struct. Biol. 5*, 289−293.
37. Hall, T. M. T., Porter, J. A., Young, K. E., Koonin, E. V., Beachy, P. A., and Leahy, D. J. (1997) *Cell 91*, 85−97.
38. Ishii, Y., Saito, Y., Fujimura, T., Isogai, T., Kojo, H., Yamashita, M., Niwa, Y., and Kohsaka, M. (1994) *J. Ferment. Bioeng. 77*, 591−597.

BI027181X