Protein Science (1999), *8*:1971–1981. Cambridge University Press. Printed in the USA.

Crystal structure of penicillin G acylase from the Bro1 mutant strain of *Providencia rettgeri*

MICHAEL A. McDONOUGH,[1,2] HERBERT E. KLEI,[3] AND JUDITH A. KELLY[1,2]

[1]Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut 06269-3125
[2]Institute of Materials Science, University of Connecticut, Storrs, Connecticut 06269-3136
[3]Bristol-Myers Squibb, Princeton, New Jersey 08543-4000

(RECEIVED May 3, 1999; ACCEPTED June 21, 1999)

Abstract

Penicillin G acylase is an important enzyme in the commercial production of semisynthetic penicillins used to combat bacterial infections. Mutant strains of *Providencia rettgeri* were generated from wild-type cultures subjected to nutritional selective pressure. One such mutant, Bro1, was able to use 6-bromohexanamide as its sole nitrogen source. Penicillin acylase from the Bro1 strain exhibited an altered substrate specificity consistent with the ability of the mutant to process 6-bromohexanamide. The X-ray structure determination of this enzyme was undertaken to understand its altered specificity and to help in the design of site-directed mutants with desired specificities. In this paper, the structure of the Bro1 penicillin G acylase has been solved at 2.5 Å resolution by molecular replacement. The *R*-factor after refinement is 0.154 and *R*-free is 0.165. Of the 758 residues in the Bro1 penicillin acylase heterodimer (α-subunit, 205; β-subunit, 553), all but the eight C-terminal residues of the α-subunit have been modeled based on a partial Bro1 sequence and the complete wild-type *P. rettgeri* sequence. A tightly bound calcium ion coordinated by one residue from the α-subunit and five residues from the β-subunit has been identified. This enzyme belongs to the superfamily of Ntn hydrolases and uses Oγ of Serβ1 as the characteristic N-terminal nucleophile. A mutation of the wild-type Metα140 to Leu in the Bro1 acylase hydrophobic specificity pocket is evident from the electron density and is consistent with the observed specificity change for Bro1 acylase. The electron density for the N-terminal Gln of the α-subunit is best modeled by the cyclized pyroglutamate form. Examination of aligned penicillin acylase and cephalosporin acylase primary sequences, in conjunction with the *P. rettgeri* and *Escherichia coli* penicillin acylase crystal structures, suggests several mutations that could potentially allow penicillin acylase to accept charged β-lactam R-groups and to function as a cephalosporin acylase and thus be used in the manufacture of semi-synthetic cephalosporins.

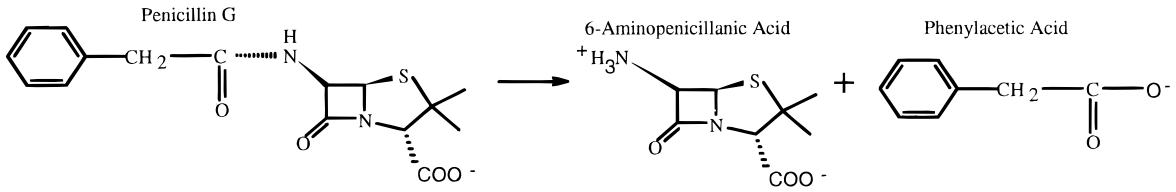
Keywords: calcium binding; Ntn hydrolase; penicillin G acylase; pyroglutamate; X-ray crystallography

Penicillin G acylase (PA) (penicillin amidohydrolase, E.C. 3.5.1.11) is found in several bacteria as well as in fungi (Vandamme & Voets, 1974). The synthesis pathway and characteristics of PA from the Bro1 mutant of *Providencia rettgeri* are similar to those of other PAs from Gram-negative bacteria (Schumacher et al., 1986; Francetic et al., 1988; Thöny-Meyer et al., 1992). The enzyme is modified by the cleavage of a 23 residue N-terminal signal sequence from the propeptide upon translocation across the plasma membrane into the periplasm. After translocation, an internal sequence of 56 residues (Klei et al., 1995) is removed from the propeptide by proteolytic cleavage resulting in the mature 85.8 kDa heterodimer (α-subunit: 23.7 kDa; β-subunit 62.1 kDa). Interestingly, the catalytic Ser is located at the N-terminus of the β-subunit created by the post-translational processing. A catalytic N-terminal nucleophile (Ntn) has been described for the nine other Ntn hydrolases with known crystal structure in the SCOP database (Murzin et al., 1995).

PA catalyzes the hydrolysis of the antibiotic penicillin G into phenylacetic acid (PAA) and 6-aminopenicillanic acid (6-APA) (see Scheme 1). Hydrolysis occurs by nucleophilic attack by the catalytic Ser Oγ on the carbonyl carbon of the β-lactam R-group. The α-amino group of the active Ser is available to potentially serve as the catalytic base analogous to the His of the catalytic triad of classical serine proteases. Substrate specificity is associated with recognition of the hydrophobic phenyl moiety of penicillin G and not the β-lactam core (Huang et al., 1963). The 6-APA product is used as a precursor in the commercial production of semisynthetic penicillins. The hope that *P. rettgeri* PA would mutate spontaneously and, through nutritional selective pressure, accept the charged β-lactam R-group of cephalosporin C, was

Reprint requests to: Judith A. Kelly, Department of Molecular and Cell Biology, University of Connecticut, 75 North Eagleville Rd. U-3125, Storrs, Connecticut 06269-3125; e-mail: kelly@uconnvm.uconn.edu.

Abbreviations: 6-APA, 6-aminopenicillanic acid; BrHEX, 6-bromohexanamide; CA, cephalosporin acylase; ESMS, electrospray mass spectrometry; MR, molecular replacement; Ntn, N-terminal nucleophile; PA, penicillin G acylase; PAA, phenylacetic acid.



Scheme 1.

not realized. However, one of the mutant strains called Bro1 exhibits a greater than 120-fold increase in specific activity toward 6-bromohexanamide (BrHEX) than the wild-type (Daumy et al., 1985b). The current structural studies were undertaken to gain an understanding of the Bro1 PA active site and to identify the structural modifications required to broaden the specificity and commercial utility of the enzyme.

The Bro1 PA had been purified to homogeneity and crystallized in space group $P6_122$ with one molecule in the asymmetric unit. Crystals were grown by the hanging drop/vapor diffusion method. Drops of 12 mg/mL enzyme in 50 mM potassium phosphate at pH 7.0–7.5, 15–25% saturated ammonium sulfate, 12–15% v/v glycerol, and 0.02% w/v sodium azide were placed over wells of 1.0 mL of buffer with 30–45% saturated ammonium sulfate. The enzyme can exist in two active isoforms (P1 and P2) that differ by the absence or presence of a spontaneously formed pyroglutamate from the N-terminal Gln in the α-subunit (Klei et al., 1995). In this paper, the three-dimensional structure of Bro1 PA at 2.5 Å resolution is presented. The mutation responsible for the altered kinetic profile of Bro1 PA relative to the wild-type PA has been identified and is discussed. The electron density for the N-terminal pyroglutamate of the α-subunit is shown. Finally, mutations to PA considered necessary to confer cephalosporin acylase (CA) activity are presented.

Results

Experimental results

The 12–15% v/v glycerol contained in the crystallization buffer, which was used to promote crystallization (Sousa & Lafer, 1990), fortuitously served as the cryoprotectant for flash cooling. Cryocooling resulted in an increase in resolution from 2.8 Å at room temperature to 2.5 Å at 100 K. The *a* and *b* unit cell dimensions remained constant upon cooling ($a = b = 140.6$ Å). However, the *c* dimension decreased by 5% upon cooling, from 209.6 to 200.2 Å, which reduced the unit cell volume by 160,000 Å^3. The mosaicity of the crystal increased from 0.20 °C at room temperature to 0.35 °C at 100 K.

Molecular replacement solution and refinement

Map analysis began with 3.0 Å multiple isomorphous replacement maps based on a $KAuCl_4$ derivative. However, the coordinates of the *Escherichia coli* PA (Duggleby et al., 1995) were made available, so efforts to phase by MR with AMoRe (Navaza, 1994) were initiated. Three *E. coli*–based search models were evaluated: a complete atom set (5,909 atoms), a polyAla subset (3,658 atoms), and a subset selected by the program MUTATE (5,310 atoms). MUTATE used the aligned wild-type *P. rettgeri* (Ljubijankic et al., 1992) and *E. coli* PA sequences to select only common atoms for the search model. Approximately 40 N-terminal residues of both Bro1 PA subunits were determined previously (Klei et al., 1995) by Edman degradation and shown to match the wild-type sequence. The MUTATE selection results in a more conservative search model (but not as conservative as polyAla) that is a subset of the known coordinates, with the option to then add atoms missing from defined positions (example: add Cβ to the search model if Gly in the known structure is aligned with Ala in the unknown structure). The correlation coefficients from the rotation function for these three search models were 12.7, 8.0, and 13.3%, respectively. The top rotation solution from all three search models was essentially the same and correct. Even in this favorable situation, MUTATE enhanced the rotation function solution with 10% fewer atoms. Although unnecessary in this particular case, the ability to enhance the signal of the rotation function by the use of moderately conservative search models, such as the one generated by MUTATE, could conceivably be important in more difficult cases.

When a translation search was then applied to the rotation solution, the correlation coefficient improved to 51.3% with an *R*-factor of 41.9%. After rigid-body refinement of the translation solution, the final correlation coefficient and *R*-factor were 60.2 and 38.4%, respectively. Simulated annealing in X-PLOR (Brünger, 1992a) resulted in an *R*-factor of 33.6%. Inspection of the electron density map after the first cycle of refinement indicated an incorrectly modeled loop where Bro1 PA has a four-residue deletion compared to the *E. coli* PA (β490–493). Refitting of this loop based on omit maps resulted in ϕ-ψ angles that correspond to a type I β-turn. Five more cycles of manual fitting and refinement followed, in conjunction with a full series of X-PLOR omit maps that were carefully examined for incorrect modeling and possible mutation sites. The omit technique was also used to look for modeling errors (Bhat & Cohen, 1984). Water molecules were added to the structure based on the analysis of $(F_o - F_c)$ difference Fourier maps. Only peaks over 3σ and with favorable hydrogen bonds were modeled as water molecules. In four cases, 3σ $(F_o - F_c)$ peaks with tetrahedral shape were still present after refinement of the model with the water molecules so these solvent peaks were modeled as sulfate ions. Prominent peaks in the anomalous difference Fourier map also suggested the presence of anomalous scatterers such as sulfur at these positions. The present model includes 781 solvent molecules (776 H_2O molecules, 4 SO_4^{2-} ions, and 1 Ca^{2+} ion).

Structure analysis

Of the 758 residues in the Bro1 acylase, all but the eight C-terminal residues of the α-subunit were modeled. For data stronger than 2σ between 8.0–2.5 Å resolution, the current *R*-factor and *R*-free (Brünger, 1992b) are 15.4 and 16.5% respectively (Table 1). The

Table 1. *Data-collection, structural, and refinement statistics*

Data-collection statistics	
Crystal (space group $P6_122$)	Native
Unit cell dimensions (Å)	$a = b = 140.6$ $c = 200.2$
Temperature	100 K
Overall	
Resolution range (Å)	27.5–2.5
Number of observations	239,869
Number of unique reflections	39,576
R_{sym}[a] (%)	11.5
Completeness (%)	96.3
$\langle I/(\sigma(I))\rangle$	13.8
Outer shell	
Resolution range (Å)	2.6–2.5
Number of observations	[b]
Number of unique reflections	3,460
R_{sym}[a] (%)	29.1
Completeness (%)	86.3
$\langle I/(\sigma(I))\rangle$	3.5
Structural and refinement statistics	
Resolution range (Å)	8.0–2.5
$I/(\sigma(I))$ cutoff	2.0
R_{model}[c] (%) (36,540 of 37,677 reflections)	15.4
R_{free}[c] (%) (1,137 of 37,677 reflections)	16.5
RMS deviations	
Bond lengths (Å)	0.015
Bond angles (°)	1.8
Dihedral angles (°)	24.5
Number of nonhydrogen atoms/B_{avg} (Å^2)	
Protein atoms	6,018/13.1
Main-chain atoms	3,002/12.6
Side-chain atoms	3,016/13.6
Cα atoms	750/12.1
Water molecules	776/25.8
Sulfate ions	4/27.5
Calcium ions	1/16.3

[a] $R_{sym}(I) = (\sum_{hkl}\sum_i|I(i) - \langle I(hkl)\rangle|)/\sum_{hkl}\langle I(hkl)\rangle$ summed over all reflections and all observations. $I(i)$ is the i^{th} observation of the intensity of the *hkl* reflection and $\langle I(hkl)\rangle$ is the mean intensity of the *hkl* reflection.

[b] The total number of observations in the outermost shell was not output by HKL.

[c] $R_{model}(F) = \sum_h||F_o(h)| - |F_c(h)||/\sum|F_o(h)|$ where F_o and F_c are the observed and calculated structure-factor amplitudes, respectively, for the reflection with Miller indices $h = (h,k,l)$. R_{free} is calculated with the same expression as R_{model} from the random subset of reflections excluded from the atomic refinements (Brünger, 1992b).

deviation of bond lengths and bond angles from stereochemical ideality are 0.015 Å and 1.8°, respectively. The overall *B*-factor for all nonhydrogen protein atoms is 14.6 Å^2. The structure validation program PROCHECK (Laskowski et al., 1993) shows that all residues are located within acceptable regions of the Ramachandran plot. An overall average *G*-factor from PROCHECK of 0.17 indicates that, on average, the deviation of dihedral angles, main-chain bond angles, and main-chain bond lengths of the model to the values of previously determined structures is well within normal limits. Changes proposed by the structure validation program WHATCHECK (Vriend, 1990), such as side-chain flips to satisfy hydrogen bonding, were incorporated into the model if they were consistent with the electron density.

Discussion

Description of the structure

The α- and β-subunits of the Bro1 PA are highly intertwined as one might expect from a protein synthesized as a single polypeptide chain (Fig. 1). Together the α- and β-subunits are made up of five distinct structural regions: a β-sandwich, two predominantly helical regions, a partial β-barrel, and a small β-strand region (Fig. 2A,B). Hereafter, the secondary structural elements are designated (H) for α-helices, (B) for β-strands, and (3_{10}) for 3_{10} helices. An α or β before the structural element symbols is used to indicate the subunit. The active site is in a deep cleft formed by residues from all five structural regions. In both Bro1 and *E. coli* PA structures, a hydrophobic pocket is located at the base of the cleft and is believed to confer specificity for the penicillin G phenylacetic R-group (Duggleby et al., 1995).

Of the five structural regions, the largest is the β-sandwich (Fig. 2, yellow). This region contains the catalytic Ser (Serβ1) and also contributes residues to the hydrophobic specificity pocket. The β-sandwich consists of a six-stranded antiparallel β-sheet (βB37, βB33–34, βB3, βB1, and βB22) opposed by a 10-stranded mixed β-sheet (βB38, αB1–2, βB4–7, βB16–17, and βB21). A pair of stabilizing α-helices is located on each face of the β-sandwich (βH1–2 and βH3–4). This region is composed of residues from the β-subunit with the exception of two β-strands from the N-terminal end of the α-subunit (αB1–2). The placement of αB1 and αB2 within the 10-stranded β-sheet positions the N-terminus of the α-subunit near the C-terminus of the β-subunit, β557, a multiconformational Arg, which, despite its proximity, does not interact with the pyroglutamate of the α-subunit.

Surrounding the core β-sandwich are the four other regions. A mostly helical region (Fig. 2, red) is situated on the six-stranded face of the β-sandwich. It is composed of residues from the β-subunit with the exception of two α-helices from the C-terminal end of the α-subunit (αH9–10). An entirely α-helical region (Fig. 2, orange), which contains the majority of the residues from the α-subunit, forms an interesting cluster of eight α-helices (αH1–8) arranged in a spiral with the eighth α-helix (αH8) passing through the center of the spiral. The C-terminus of αH8 is involved in the hydrophobic specificity pocket. This region packs against the 10-stranded face of the β-sandwich. The spiral helical region is also flanked by another β-strand region that approximates a β-barrel (Fig. 2, green). Two antiparallel β-strands (βB11–12) of the partial β-barrel pack alongside helix αH4. A small β-strand region (Fig. 2, blue), with residues that contribute to the hydrophobic specificity pocket, flanks both the β-sandwich and the spiral helical region. Three of the 41 prolines are in the *cis* conformation (Proβ29, Proβ366, and Proβ501). *Cis*-Proβ29 and *cis*-Proβ501 are located in loops while *cis*-Proβ366 initiates the helix $\beta 3_{10}1$.

Ntn-hydrolase fold

Structural homology has been found between the β-sandwich region of *E. coli* PA (Brannigan et al., 1995; Duggleby et al., 1995), and now *P. rettgeri* Bro1 PA, and other amidohydrolases with known crystal structures (Table 2). Proteins with this distinctive motif are referred to as Ntn hydrolases (Murzin, 1996). The proteins of this superfamily possess similar characteristics such as an N-terminal nucleophile and autocatalytic processing (Brannigan et al., 1995; Murzin, 1996). While there is no conclusive evidence

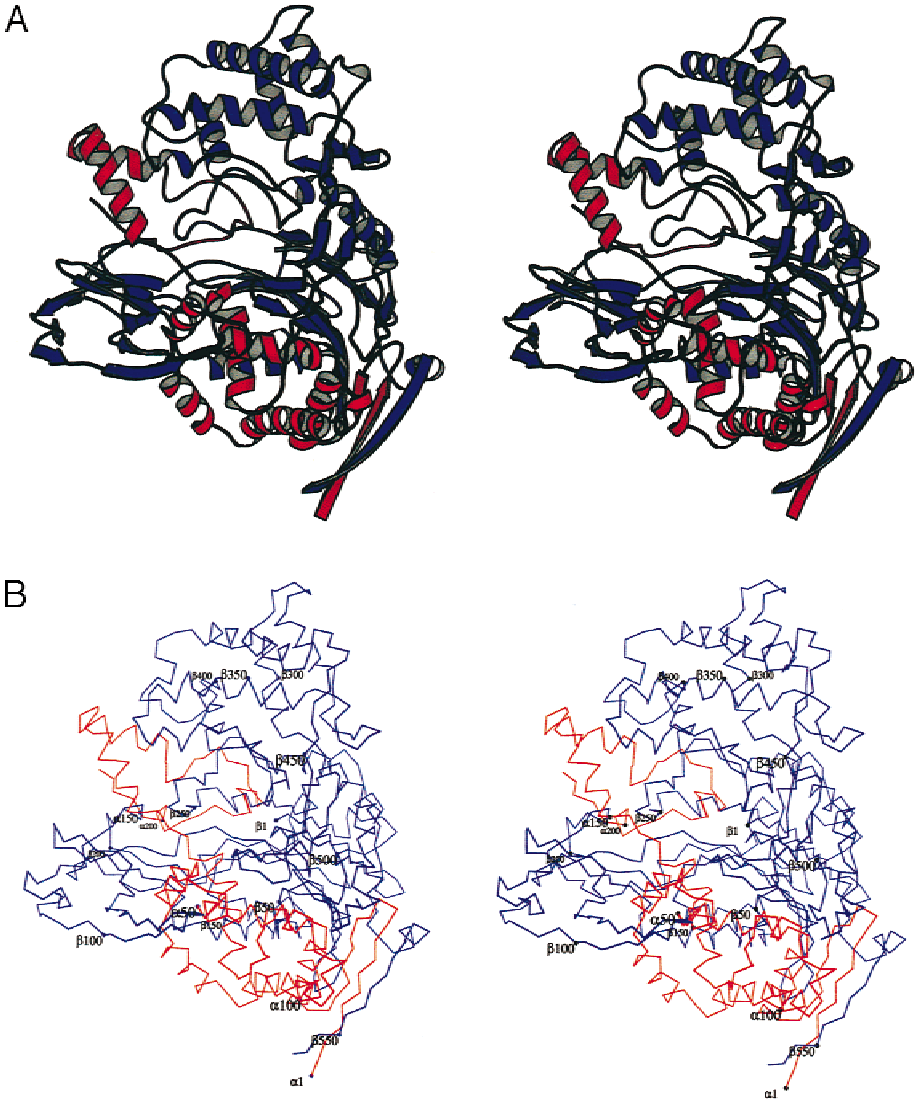


Fig. 1. Stereo representations of the Bro1 PA created with the program MOLSCRIPT (Kraulis, 1991). The α- and β-subunits are colored red and blue, respectively. **A:** Ribbon diagram. **B:** Cα trace labeled every 50 residues. The size of the labels is adjusted to convey depth perpendicular to the page.

for autocatalysis of Bro1 PA, the mechanism for PA autocatalysis suggested by Brannigan et al. (1995) based on the *E. coli* PA structure remains plausible in light of the Bro1 PA structure. The location of the N-terminal nucleophilic residue in each of these proteins is on a β-strand within a β-sheet that is part of a β-sandwich bordered by two α-helices on both sides. Ntn hydrolases can have either Ser, Cys, or Thr as their nucleophilic residue. They also have remarkably different substrates. As expected, results from the three-dimensional structural homology search program DALI (Holm & Sander, 1995) indicate that the Bro1 PA β-subunit is homologous to the *E. coli* PA β-subunit, and each of the other Ntn hydrolase families classified by SCOP (Murzin et al., 1995).

Catalytic residues and mechanism

There are four key residues conserved in the active site of the *P. rettgeri* and *E. coli* PA structures: the catalytic Serβ1, plus Glnβ23, Alaβ69, and Asnβ241. In both PA structures, a water molecule links the nucleophilic Serβ1 Oγ and its α-amino group making Oγ even more nucleophilic. In the Bro1 PA structure, this water also hydrogen bonds with the main-chain amide of Glnβ23. The Serβ1 α-amino group hydrogen bonds with the main-chain carbonyl oxygens of both Glnβ23 and Asnβ241 giving the bridging water oxygen a more basic character. The amide bond of the substrate is broken when the nucleophilic Oγ of Serβ1 attacks the carbonyl oxygen of the substrate forming an acyl-enzyme intermediate. Bond breakage is facilitated by the positioning of the substrate side-chain carbonyl in an oxyanion hole formed by the main-chain amide nitrogen of Alaβ69 and Nδ of Asnβ241 (Fig. 3) (Duggleby et al., 1995).

Description of hydrophobic specificity pocket

Some published reports attribute the specificity of PA for penicillin G to only the α-subunit (Daumy et al., 1985a) while others attribute

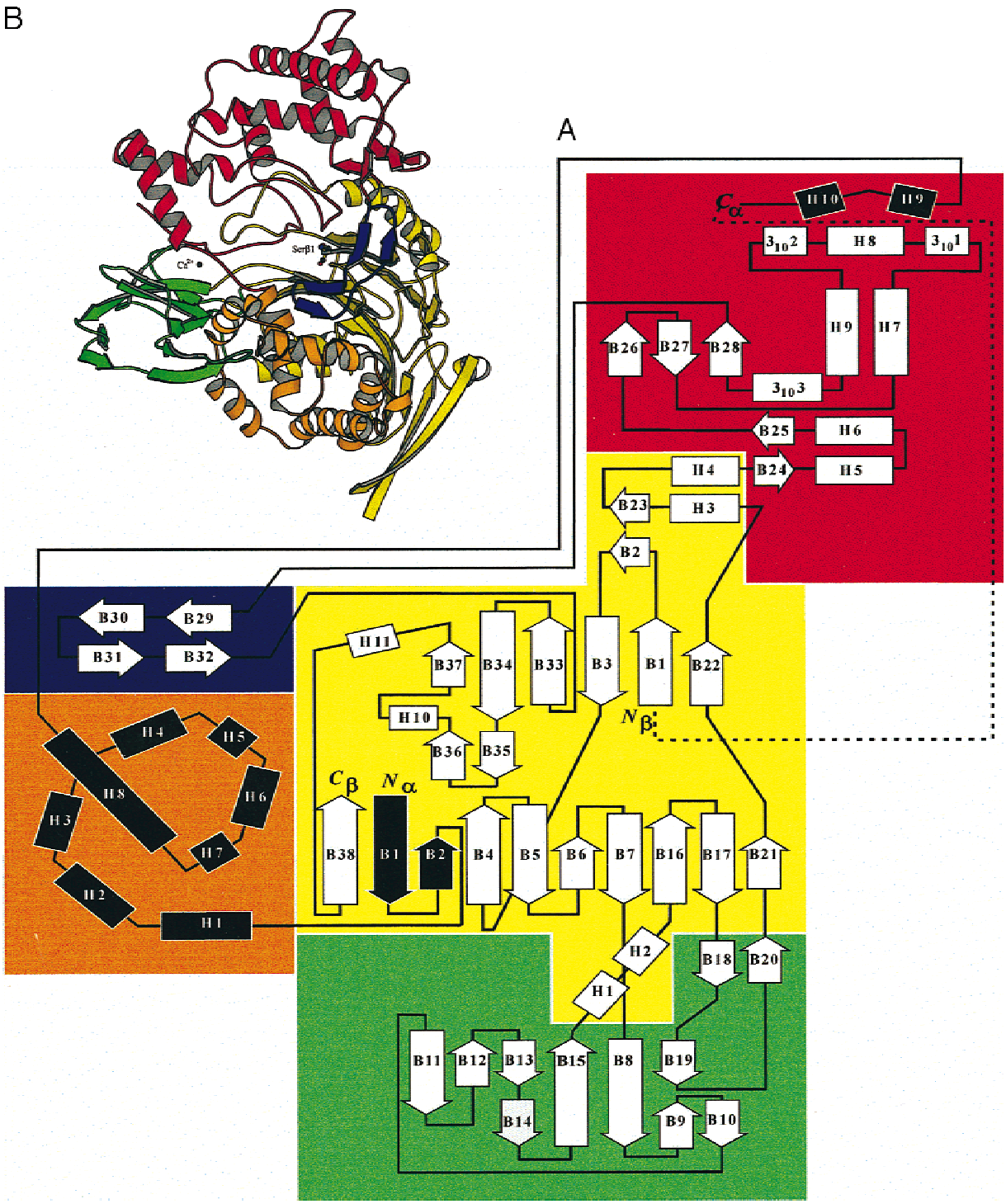


Fig. 2. Color-coded illustrations of the major structural regions of Bro1 PA: β-sandwich (yellow), two helical regions (red and orange), partial β-barrel (green), and a small β-strand region (blue). **A:** Numbered secondary structure elements are designated as rectangles for α-helix (H) and 3_{10}-helix (3_{10}), and arrows for β-strand (B). The structural elements for the α- and β-subunits are shaded black and white, respectively. The connection between the α- and β-subunits made by an internal sequence prior to post-translational removal from the propeptide is shown as a dotted line. **B:** Ribbon diagram produced with the program MOLSCRIPT (Kraulis, 1991).

some specificity to the β-subunit (Roa et al., 1994). This paper makes it clear that the hydrophobic specificity pocket is made up of residues from both the α- and β-subunits (Table 3). The single mutation observed in the electron density map of Bro1 is found in the hydrophobic specificity pocket. Residue α140 was initially modeled as Met based on the wild-type *P. rettgeri* sequence. However, when $(F_o - F_c)$ difference Fourier maps were contoured at 3σ, negative electron density was seen around the sulfur atom of

Table 2. *Summary of Ntn hydrolases with known crystal structure derived from SCOP Release 1.39*

Ntn hydrolase family	Penicillin acylase, catalytic domain	Proteasome subunits	Class II Glutamine amidotransferases	(Glycosyl)asparaginases
Biological source (PDB id)[a]	*P. rettgeri*[1] (1cp9) *E. coli*[2] (1pnk)	*Thermoplasma acidophilum*[3] (1pma) *Saccharomyces cerevisiae*[4] (1ryp) *E. coli*[5] (1ned)	*B. subtilis*[6] (1gph) *E. coli*[7] (1ecb) *E. coli*[8] (1gdo)	Human[9] (1apy) *Flavobacterium meningosepticum*[10] (1ayy)
Biologically active quaternary structure	Heterodimer α,β	[3,4]Hetero28mer $14\alpha,14\beta$ [5]Dodecamer	[6,7]Homotetramer [8]Monomer (fragment of whole enzyme)	[9](Crystallized as) Heterotetramer $2\alpha,2\beta$ [10]Heterodimer
N-terminal nucleophilic residue	Serine	Threonine	Cysteine	Threonine
Reaction catalyzed	Hydrolyzes penicillin G into PAA and 6-APA	Nonspecific peptide bond hydrolysis	[6,7]Transfers the side-chain amine of L-glutamine to 5-phosphoribosyl-1-pyrophosphate to form phosphoribosyl-1-amine, glutamate, pyrophosphate, and H_2O. [8]Hydrolyzes L-glutamine to L-glutamate and ammonia.	[9]Hydrolyzes the N-glycosidic linkage between glycan and L-asparagine to form L-aspartic acid and 1-aminoglycan. [10]Cleaves the link between the asparagine and the N–acetylglucosamine of N-linked oligosaccharides.

[a]Only one representative PDB id is given for each structure, other complexed entries exist in the PDB.

Metα140. Previous electrospray mass spectrometry (ESMS) on CNBr digest fragments of the Bro1 α-subunit gave a molecular weight for the fragment containing Metα140 that was inconsistent with the wild-type sequence (Klei et al., 1995). Substituting Ile, Leu, or Asn for Metα140 yields the ESMS molecular weight. Anomalous difference maps contained peaks for the sulfur atoms of 12 of the 14 Met in the *P. rettgeri* sequence. No anomalous difference peak was seen near residue α140. Omit maps clearly show density for this residue inconsistent with Met. The shape of the electron density is consistent with Leu or Asn, but not Ile. Based on the hydrophobic nature of the specificity pocket and the fact that the Bro1 mutant readily uses BrHEX as a substrate, α140 was modeled as Leu. Changes in substrate specificity due to single residue mutations of the equivalent Met in the *E. coli* (Metα142) (Williams & Zuzel, 1985) and *Kluyvera citrophila* (Metα168) (Martín et al., 1990) enzymes are also documented.

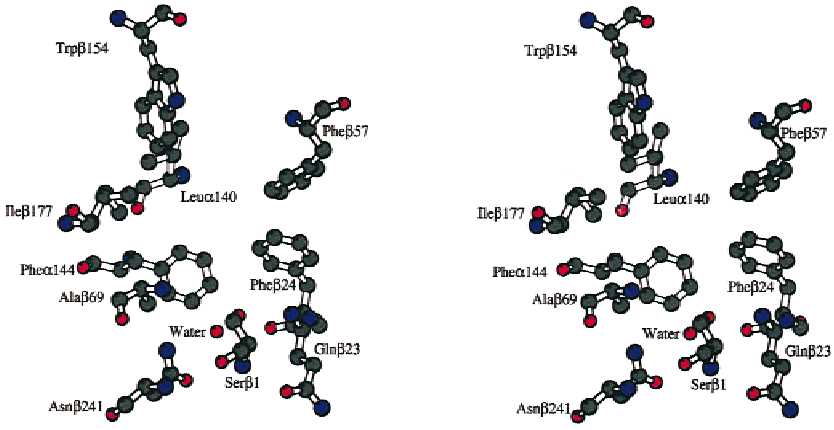


Fig. 3. Stereo view of the active site. All active site residues and select hydrophobic specificity pocket residues are shown (see Table 3). Created with the program MOLSCRIPT (Kraulis, 1991).

Table 3. *Summary of key residues in Bro1 PA*

Residue	Location[a]	Description
Serβ1	A	Oγ is the N-terminal nucleophile.
Glnβ23	A	Main-chain N H-bonds to Serβ1 Oγ, bridging catalytic water molecule.
Alaβ69	A,S	Main-chain carbonyl O H-bonds Serβ1 N-terminal N. Main-chain N helps form the oxyanion hole. Aligns with Hisβ70 in CA.[b]
Asnβ241	A	Nδ helps form the oxyanion hole.
Leuα140	S	Hydrophobic. Suspected mutation from wild-type Met. Aligns with Leuα163 in CA.[b]
Pheα144	S	Hydrophobic. Aligns with Valα167 in CA.[b]
Pheβ24	S	Hydrophobic. Aligns with Argβ24 in CA.[b]
Tyrβ31	S	Hydrophobic.
Proβ49	S	Hydrophobic.
Tyrβ52	S	Hydrophobic. Aligns with Valβ52 in CA.[b]
Leuβ56	S	Hydrophobic.
Pheβ57	S	Hydrophobic.
Trpβ154	S	Hydrophobic. Aligns with Aspβ177 in CA.[b]
Ileβ177	S	Hydrophobic. Aligns with Aspβ177 in CA.[b]
Gluα150	C	Oϵ1 is 2.4 Å from calcium ion.
Aspβ73	C	Oδ1 is 2.5 Å from calcium ion. Oδ2 is 2.5 Å from calcium ion.
Valβ75	C	Main-chain O is 2.6 Å from calcium ion.
Aspβ76	C	Oδ is 2.5 Å from calcium ion.
Proβ205	C	Main-chain O is 2.5 Å from calcium ion.
Aspβ252	C	Oδ is 2.6 Å from calcium ion.

[a]A = active site; S = specificity pocket; C = calcium binding site.
[b]CA = cephalosporin acylase from *Pseudomonas sp*. strain SE83 (acyII) (Matsuda et al., 1987).

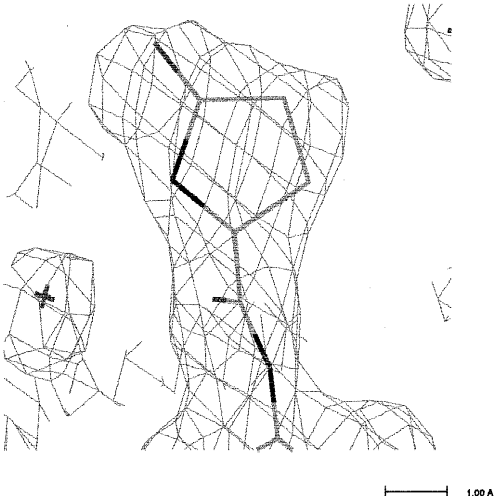


Fig. 4. ($2F_o - F_c$) electron density of the pyroglutamate at the N-terminus of the α-subunit. The density was contoured at 1σ and displayed with the program CHAIN (Sack, 1988).

N-terminal pyroglutamate

Initially, the N-terminal residue of the α-subunit could not be identified in the electron density. However, as the refinement progressed, omit maps clearly showed density consistent with pyroglutamate (Fig. 4). This result was unexpected since the sample of the enzyme used in the crystallization was shown to have Gln at the N-terminus (P1 isoform) at the time of purification (Klei et al., 1995). Apparently, conversion to pyroglutamate occurred during the course of crystallization and (or) crystal storage in the phosphate buffer at pH 7.0–7.5. Spontaneous pyroglutamate formation can occur at slightly acidic to neutral pH at room temperature over the course of several days (Yuuki et al., 1989). The Bro1 crystallization conditions favor cyclization, and no obvious interactions or crystal contacts are present that could prevent the reaction. An N-terminal Glu can also spontaneously cyclize to form pyroglutamate. The N-terminus of the *E. coli* PA is Glu, but this residue was not included in the *E. coli* coordinates (Duggleby et al., 1995). It is possible that the N-terminal glutamate of the *E. coli* PA also forms pyroglutamate. Cyclization of Glu is essential for the activity of some neuropeptides (Moret & Briley, 1988). However, in the case of the Bro1 PA, it does not affect activity and appears to be simply the result of nonenzymatic cyclization of the N-terminus.

Calcium binding site

While Bro1 has no known requirement for calcium, a calcium binding site was located by difference Fourier during the latter stages of refinement when solvent molecules were being added to the structure. The calcium site was the highest peak in the difference Fourier map. The site was also confirmed in an anomalous difference Fourier map. The calcium binding site is located between αH9 and βB8. The calcium ion is octahedrally coordinated by seven oxygens located within 2.4–2.6 Å and contributed by six residues: both carboxyl oxygens of Aspβ73; a carboxyl oxygen from Gluα150, Aspβ76, and Aspβ252; and the backbone carbonyl oxygens of Valβ75 and Proβ205 (Fig. 5). A calcium ion was also found in the *E. coli* PA structure (Duggleby et al., 1995).

Assignment of a calcium ion to this site was supported by the valence calculation of potential calcium binding sites in the program VALE (Nayal & Di Cera, 1994). Nearly all grid points (98% on a 0.12 Å grid) within 1 Å of the calcium ion had a calculated valence greater than or equal to 1.4 and the average valence for all grid points within 1 Å of the calcium ion was 1.70. An average value of 1.70 has been found to be statistically significant when compared with a database of known calcium-containing structures (Nayal & Di Cera, 1994). Substitution of other cations (Mg^{+2}, Na^{+}, and K^{+}) resulted in a widely dispersed array of grid points with a valence greater than or equal to 1.4, indicating poor valence compatibility for these ions with the Bro1 PA structure.

Comparison with E. coli PA structure

The wild-type *P. rettgeri* PA shows 64% sequence identity with the *E. coli* PA. The sequence homology rises above 75% if conservative changes are taken into account. There is one amino acid difference between the Bro1 and *E. coli* PA hydrophobic specificity pockets. Metα140 of wild-type *P. rettgeri* (Metα142 of *E. coli*) is mutated to Leu in Bro1 PA. This mutation accounts for the difference between the calculated mass of the wild-type α-subunit and the ESMS mass of the Bro1 α-subunit. It is consistent with the Bro1 electron density and could explain the altered Bro1 substrate profile due to the difference in the length of the Leu side chain vs. the length of the Met side chain found in both the *E. coli* and wild-type *P. rettgeri* PAs (Fig. 6). Another difference between the

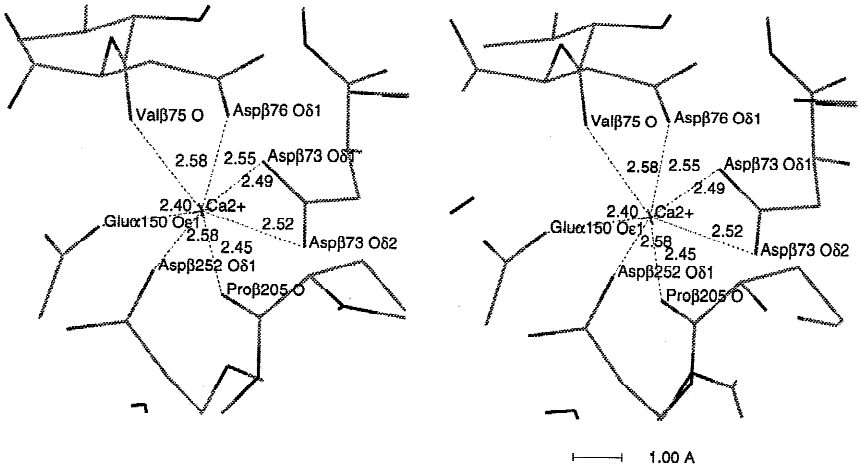


Fig. 5. Stick representation of the calcium binding site created with the program CHAIN (Sack, 1988). Interactions between the oxygen ligands and the central Ca^{2+} ion are shown along with the interatomic distances.

two sequences occurs in a loop between two antiparallel β-strands (βB33–34) located in the β-sandwich in which the *E. coli* PA contains four additional residues.

Sequence alignment and structure superposition with other PAs/CA

The sequence alignment (Konstantinovic et al., 1994) of PAs from wild-type *P. rettgeri, E. coli, K. citrophila, Arthrobacter viscosus*, and CA from *Pseudomonas sp.* strain SE83 (acyII) was compared to the superposition of the three-dimensional structures of the *P. rettgeri* and *E. coli* PAs. From this comparison, it was possible to identify the key catalytic residues, the calcium binding residues, and residues that could be responsible for substrate specificity in the PAs and CA. The four active site residues (Serβ1, Glnβ23, Alaβ69, and Asnβ241) are strictly conserved in all of the PAs. The catalytic Ser is conserved in CA as well. A second Asn residue (Asnβ242) is also conserved in all four of the PAs. It is likely that this homology also extends to CA even though the alignment by Konstantinovic et al. (1994) positions CA differently in this region. By shifting the four residue gap in CA from between residues Valβ239 and Thrβ240 to between residues Valβ245 and Valβ246, Asnβ242 and Asnβ243 become aligned with the conserved Asns in the PAs. Another modification is also included where the *P. rettgeri* Bro1 PA α-subunit was shown to begin with Gln-Ser-Thr (Klei et al., 1995). PA Glnβ23 and Alaβ69 are paired with CA Hisβ23 and Hisβ70, respectively. These changes are presumed tolerable because it is only the main-chain amides of these two residues that are required to help form the oxyanion hole. The gaps after residues PA Ileβ118 and CA Gluβ121 are closed with no significant changes to the alignment of key residues. The modified alignment of Bro1 PA and CA (Matsuda et al., 1987) is shown in Figure 7, with key residues (Table 3) highlighted.

All of the calcium binding ligands of PA are conserved in CA except for Gluα150. It is interesting to note that Gluα150 is the only residue from the α-subunit of PA that participates in calcium binding. Based on the alignment (Fig. 7), Gluα150 is equivalent to Argα173 in CA thus reducing the number of protein atoms serving as calcium ligands from seven to six. None of the residues around this single difference in CA contain a side-chain oxygen that might serve as an alternate calcium ligand. It is possible the α-subunit of CA could adopt a different conformation that would allow a main-

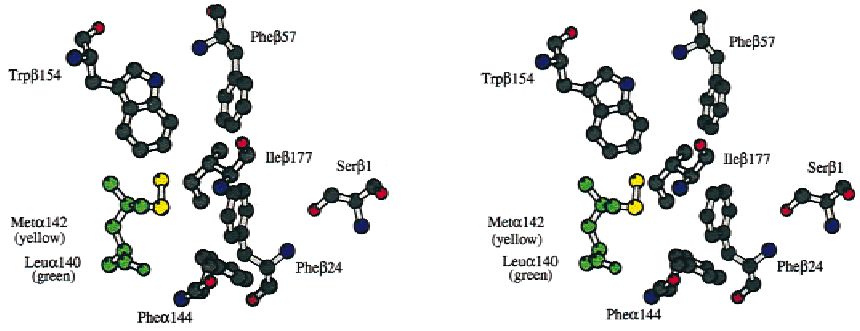


Fig. 6. Stereo view of key residues in the hydrophobic specificity pocket created with MOLSCRIPT (Kraulis, 1991). The mutation of Metα140 is believed to be largely responsible for the different substrate profiles of the Bro1 and wild-type PAs from *P. rettgeri*. The model of Bro1 Leuα140, which is consistent with the electron density, is shown in green. Sδ and Cϵ from Metα142 of the superimposed *E. coli* PA structure (Duggleby et al., 1995) are shown in yellow. Atoms from residues other than α140 are colored by atom type.

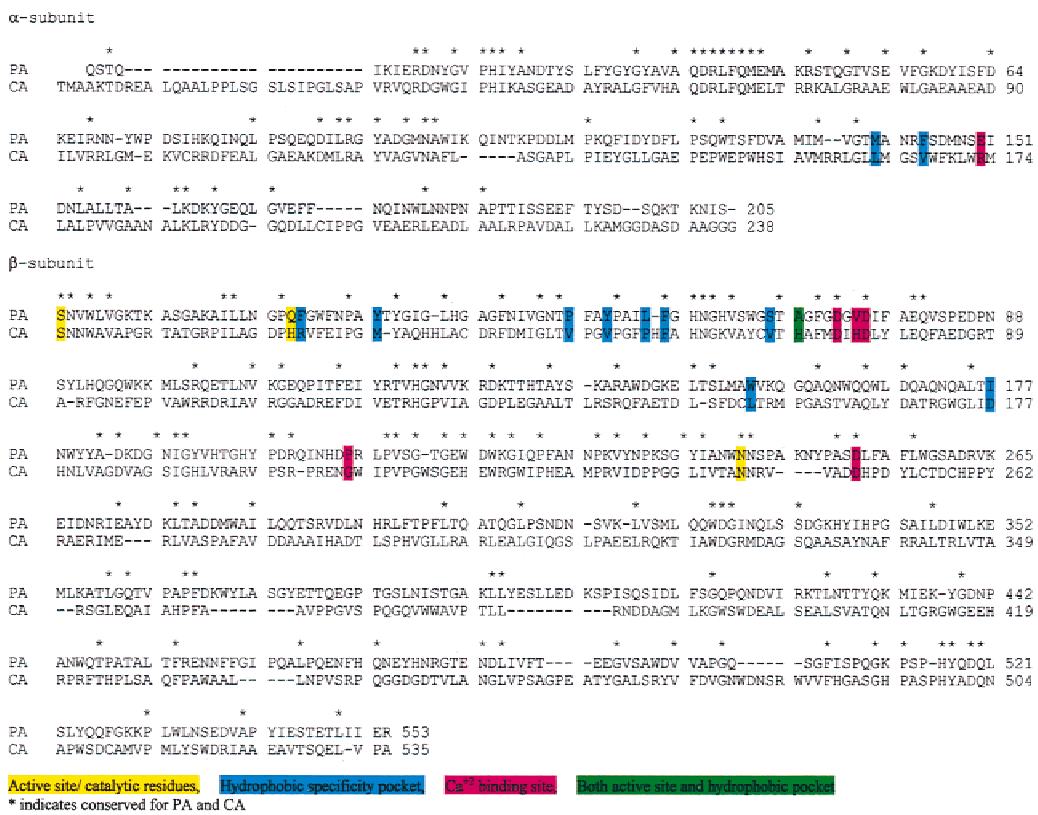


Fig. 7. Amino acid sequence alignment of wild-type *P. rettgeri* PA (Ljubijankic et al., 1992) with *Pseudomonas* sp. SE83 CA (acyII) (Matsuda et al., 1987) derived from Konstantinovic et al. (1994) and modified as described in the text. Key residues are highlighted (see Table 3). Identity is indicated by an asterisk.

chain carbonyl oxygen to replace the missing ligand or a water molecule may move in to serve that role.

Another interesting aspect of this comparison is the type of residues found in the specificity pocket of the PAs vs. those of the CA (Fig. 8). The specificity pockets of all four PAs are lined solely with hydrophobic residues. However, in the CA sequence, three residues modeled in the pocket are hydrogen bond donor/acceptors: Argβ24, Hisβ70, and Aspβ177 (Pheβ24, Alaβ69, and Ileβ177, respectively, in PA). It would be informative to test if the positive charge introduced by the mutation Pheβ24Arg makes 6-carboxyl-hexanamide, a close analog of the cephalosporin C β-lactam R-group, a viable substrate for PA. The combination of the mutations Pheβ24Arg and Ile1β177Asp could provide an environment in which the charged side chain of cephalosporin C has satisfactory electrostatic/hydrogen bonding interactions. When residue Alaβ69 is modeled as His in the *E. coli* and *P. rettgeri* Bro1 PA structures, its side-chain overlays a tightly bound water molecule in the active site of PA that is believed to play a catalytic role (Duggleby et al.,

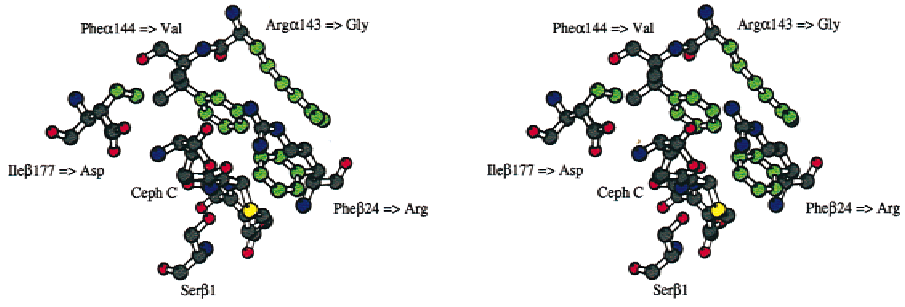


Fig. 8. Hypothetical model of selected *Pseudomonas sp.* SE83 CA (acyII) specificity pocket residues with cephalosporin C shown in cpk color scheme. Bro1 residues are superimposed and shown in green (common atoms retain cpk color scheme). This model is based on the sequence alignment with wild-type *P. rettgeri* PA in Figure 7, the Bro1 PA structure, and *E. coli* PA structure complexed with phenylacetic acid PDB entry 1pnl (Duggleby et al., 1995).

1995). Residues Pheα144, Tyrβ52, and Trpβ154 in Bro1 PA (Valα167, Valβ52, and Leuβ154, respectively, in CA) remain hydrophobic but are reduced in size. Residue Argα143 in Bro1 is reduced in size to Gly in CA. The reduction in residue size in CA could allow for the larger β-lactam R-group of cephalosporin C. For reasons discussed earlier, Bro1 PA α140 was modeled as Leu. The fact that Bro1 PA α140 aligns with CA Leuα163 is consistent with the Bro1 substrate specificity being more CA-like. Modeling the CA differences into the Bro1 acylase structure and placing cephalosporin C into the specificity pocket indicate that these changes could account for the difference in substrate preference between the PAs and CA.

Materials and methods

Data collection and processing

A native data set was collected from a 0.3 mm hexagonal bipyramidal crystal grown from the P1 isoform of Bro1 PA. An MSC R-axisIIC imaging plate area detector system was used with CuKα radiation from a Rigaku RU-200 rotating anode X-ray generator operating at 50 kV and l00 mA and equipped with a graphite monochromator, 0.3 mm collimator, and 0.3 × 0.3 mm focus. The crystal was flash cooled to 100 K in an Oxford Cryosystems nitrogen stream. The crystal storage solution of 60% saturated ammonium sulfate, 15% v/v glycerol, 0.02% w/v sodium azide, and 50 mM potassium phosphate at pH 7.0–7.5, was also an appropriate cryo-solvent. The exposure time and oscillation range per image were 15 min and 0.5°, respectively. Data were processed, scaled, and merged with the HKL program suite (Table 1) (Otwinowski & Minor, 1997).

Structure determination and refinement

The phase problem was solved by the method of MR (Rossmann, 1972) as implemented in the program AMoRe (Navaza, 1994) from the CCP4 crystallographic software suite (CCP4, 1994). The search model used was based on the *E. coli* PA structure (Duggleby et al., 1995). Manual alignment of the wild-type *P. rettgeri* PA sequence with the *E. coli* PA sequence bracketed the location of the four residue deletion in the *P. rettgeri* sequence. In agreement with other published alignments (Ljubijankic et al., 1992), four residues (β490–493) were deleted from the *E. coli* PA structure and the polypeptide chain was rejoined with the program CHAIN (Sack, 1988) to form a shortened loop. The search model was further modified with the program MUTATE by removing side-chain atoms known to be absent from the Bro1 PA structure based on the above alignments. After the rotation and translation solutions from AMoRe were applied to the search model, six cycles of refinement/fitting were done with the programs X-PLOR and CHAIN. The initial *B*-factors, taken directly from the search model, were not refined in the first cycle. Isotropic individual atomic *B*-factors were refined in all subsequent cycles. In the last four refinement cycles, solvent molecules were placed at peaks of $(F_o - F_c)$ density above 4σ (3σ in the final cycle), which exhibited good hydrogen bonding geometry with protein atoms. Anomalous difference Fourier maps were used to identify the location of the calcium ion, and its valence coordination was analyzed with the program VALE. Sulfate ions were also located from anomalous difference Fourier maps and were confirmed by inspection of their local environment. A cycle of X-PLOR refinement was done with Friedel mates unmerged to test if detectable native anomalous signals from the calcium ion, sulfate ions, and $S\delta$ of the Met residues could be used to improve the model. Of the 70,369 measured reflections, the 68,668 reflections between 8.0–2.5 Å resolution with $I/\sigma(I)$ greater than 2.0 were used for the refinement with 1,626 (2.4%) of the 68,668 reflections reserved for the *R*-free calculation. Care was taken to use the same set of test reflections as when Friedel mates were merged only, now, both F^+ and F^- for acentric reflections were placed in the test set. The quality of the electron density maps and statistics were essentially unchanged. No attempt was made to model and refine alternate conformations due to the 2.5 Å resolution despite evidence of discrete disorder for some residues, such as Argβ557, in omit maps. Structure validation was performed with the programs PROCHECK and WHATCHECK. A three-dimensional structural homology search against a database of 872 protein chains was done with the program DALI. Coordinates of the Bro1 PA have been deposited in the Protein Data Bank (PDB) as entry 1cp9.

Acknowledgments

We especially thank Gaston Daumy of Pfizer Central Research for supplying the Bro1 mutant of *P. rettgeri* and for many useful comments. We appreciate the assistance of Gayle Schulte of Pfizer Central Research for her help with the collection of the flash-cooled data. We thank Peter Moody and Guy Dodson of the University of York who kindly provided the coordinates of the *E. coli* PA prior to their release from the PDB. We would like to thank Steven Sheriff of Bristol-Myers Squibb for the use of MUTATE. This work was supported by the State of Connecticut Critical Technologies Program in Drug Design.

References

Bhat TN, Cohen GH. 1984. OMITMAP: An electron density map suitable for the examination of errors in a macromolecular model. *J Appl Cryst 17*:244–248.

Brannigan JA, Dodson G, Duggleby HJ, Moody PCE, Smith J, Tomchick DR, Murzin AG. 1995. A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. *Nature 378*:416–419

Brünger AT. 1992a. *X-PLOR version 3.1 manual.* New Haven, Connecticut: Yale University Press.

Brünger AT. 1992b. The free *R* value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature 355*:472–474.

CCP4 (Collaborative Computational Project, Number 4). 1994. The CCP4 suite: Programs for protein crystallography. *Acta Cryst D 50*:760–763.

Daumy GO, Danley D, McColl AS. 1985a. Role of protein subunits in *Proteus rettgeri* penicillin G acylase. *J Bacteriol 163*:1279–1281.

Daumy GO, Danley D, McColl AS, Apostolakos D, Vinick FJ. 1985b. Experimental evolution of penicillin G acylases from *Escherichia coli* and *Proteus rettgeri*. *J Bacteriol 163*:925–932.

Duggleby HJ, Tolley SP, Hill CP, Dodson EJ, Dodson G, Moody PCE. 1995. Penicillin acylase has a single amino acid catalytic center. *Nature 373*:264–268.

Francetic O, Deretic V, Marjanovic N, Glišin V. 1988. Penicillin acylases: Their function, structures, genes, and commercial applications. *BTF Biotech Forum 5*:90–94.

Holm L, Sander C. 1995. DALI: A network tool for protein structure comparison. *Trends Biochem Sci 20*:478–480.

Huang HT, Seto TA, Shull GM. 1963. Distribution and substrate specificity of benzylpenicillin acylase. *Appl Microbiol 11*:1–6.

Klei HE, Daumy GO, Kelly JA. 1995. Purification and preliminary crystallographic studies of penicillin G acylase from *P. rettgeri. Protein Sci 4*:433–441.

Konstantinovic M, Marjanovic N, Ljubijankic G, Glišin V. 1994. The penicillin amidase of *Arthrobacter viscosus* ATCC 15294. *Gene 143*:79–83.

Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr 24*:946–950.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr 26*:283–291.

Ljubijankic G, Konstantinovic M, Glišin V. 1992. The primary structure of *P. rettgeri* penicillin G amidase gene and its relationship to other gram negative amidases. *J DNA Sequencing and Mapping 3*:195–200.

Martín J, Prieto I, Barbero JL, Peréz-Gil J, Mancheño JM, Arche R. 1990. Thermodynamic profiles of penicillin G hydrolysis catalyzed by wild-type and Met-Ala168 mutant penicillin acylases from *Kluyvera citrophila*. *Biochim Biophys Acta 1037*:133–139.

Matsuda A, Toma K, Komatsu K. 1987. Nucleotide sequences of the genes for two distinct cephalosporin acylases from a *Pseudomonas* strain. *J Bacteriol 169*:5821–5826.

Moret C, Briley M. 1988. The "forgotten" amino acid pyroglutamate. *Trends Pharmacol Sci 9*:278–279.

Murzin AG. 1996. Structural classification of proteins: New superfamilies. *Curr Opin Struct Biol 6*:386–394.

Murzin AG, Brenner SE, Hubbard T, Chothia C. 1995. SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol 247*:536–540.

Navaza J. 1994. AMoRe: An automated procedure for molecular replacement. *Acta Crystallogr A 50*:157–163.

Nayal M, Di Cera E. 1994. Predicting Ca^{2+} binding sites in proteins. *Proc Natl Acad Sci USA 91*:817–821.

Otwinowski Z, Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol 276*:307–326.

Roa A, Garcia JL, Salto F, Cortez E. 1994. Changing the substrate specificity of penicillin G acylase from *Kluyvera citrophila* through selective pressure. *Biochem J 303*:869–876.

Rossmann MG, ed. 1972. *The molecular replacement method*. New York: Gordon and Breach.

Sack JS. 1988. CHAIN: A crystallographic modeling program. *J Mol Graphics 6*:224–225.

Schumacher G, Sizmann D, Haug H, Buckel P, Böck A. 1986. Penicillin acylase from *E. coli*: Unique gene-protein relation. *Nucleic Acids Res 14*:5713–5727.

Sousa R, Lafer EM. 1990. The use of glycerol in crystallization of T7 RNA polymerase: Implications for the use of cosolvents in crystallizing flexible proteins. *Methods: A companion to Methods Enzymol 1*:50–56.

Thöny-Meyer L, Böck A, Hennecke H. 1992. Prokaryotic polyprotein precursors. *FEBS Lett 307*:62–65.

Vandamme EJ, Voets JP. 1974. Microbial penicillin acylases. *Adv Appl Microbiol 17*:311–369.

Vriend G. 1990. WHATIF: A molecular modeling and drug design program. *J Mol Graphics 8*:52–56.

Williams JA, Zuzel TJ. 1985. Penicillin G acylase E.C.3.4.1.11 substrate specificity modification by in vitro mutagenesis. *J Cell Biochem 9B*:99.

Yuuki T, Hidetoshi T, Yabuuchi S. 1989. Purification and some properties of two enzymes from a β-glucanase hyperproducing strain, *Bacillus subtilis* HL-25. *Agric Biol Chem 53*:2341–2346.