Applied and Environmental Microbiology

Engineering the Substrate Specificity of a Thermophilic Penicillin Acylase from Thermus thermophilus

Leticia L. Torres, Ángel Cantero, Mercedes del Valle, Anabel Marina, Fernando López-Gallego, José M. Guisán, José Berenguer and Aurelio Hidalgo

Appl. Environ. Microbiol. **2013, 79(5):1555. DOI: 10.1128/AEM.03215-12.**
Published Ahead of Print 21 December 2012.

AEM
Journals.ASM.org

Engineering the Substrate Specificity of a Thermophilic Penicillin Acylase from *Thermus thermophilus*

Leticia L. Torres,[a] Ángel Cantero,[a] Mercedes del Valle,[a] Anabel Marina,[a] Fernando López-Gallego,[b] José M. Guisán,[b] José Berenguer,[a] Aurelio Hidalgo[a]

Center for Molecular Biology Severo Ochoa, Universidad Autónoma de Madrid, Consejo Superior de Investigaciones Científicas (UAM-CSIC), Madrid, Spain[a]; Department of Biocatalysis, Institute for Catalysis and Petrochemistry, Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain[b]

A homologue of the *Escherichia coli* penicillin acylase is encoded in the genomes of several thermophiles, including in different *Thermus thermophilus* strains. Although the natural substrate of this enzyme is not known, this acylase shows a marked preference for penicillin K over penicillin G. Three-dimensional models were created in which the catalytic residues and the substrate binding pocket were identified. Through rational redesign, residues were replaced to mimic the aromatic binding site of the *E. coli* penicillin G acylase. A set of enzyme variants containing between one and four amino acid replacements was generated, with altered catalytic properties in the hydrolyses of penicillins K and G. The introduction of a single phenylalanine residue in position α188, α189, or β24 improved the K_m for penicillin G between 9- and 12-fold, and the catalytic efficiency of these variants for penicillin G was improved up to 6.6-fold. Structural models, as well as docking analyses, can predict the positioning of penicillins G and K for catalysis and can demonstrate how binding in a productive pose is compromised when more than one bulky phenylalanine residue is introduced into the active site.

Penicillin G acylase (PGA) (also known as penicillin amidohydrolase; EC 3.5.1.11) is one of the most relevant enzymes in the pharmaceutical industry. Its main application is in the manufacture of over 10,000 metric tonnes (Tm)/year of 6-aminopenicillanic acid (6-APA). Another application of PGA is in the coupling reaction of 6-APA with esters or amides under kinetic control for the manufacture of antibiotic products, such as ampicillin, amoxicillin, or cephalexin, which have better spectrum and acid stability than does the parental penicillin G (PenG). Furthermore, the low specificity of the amine subsite of PGA allows the acylation of different types of amines, such as 7-amino deoxy cephalosporanic acid (7-ADCA) or aliphatic amines (1). The latter case is more demanding in terms of biocatalyst stability, because the high pK_a value of aliphatic amines requires either high pH or organic solvents in order to obtain a reasonable percentage of reactive nonionized amide. Other possible uses in biocatalysis include kinetic resolutions of amino acids, amines, and secondary alcohols; protection and deprotection of amino groups; and derivatization of sugars (2, 3). Once again, the relaxed specificity of the binding site for the nucleophilic amine allows the use of this enzyme for a wide variety of resolutions, but the use of organic solvents and high pH demands a stable biocatalyst.

Industrially, *Escherichia coli* PGA (EcoPGA) is the enzyme of choice, whether it be recombinant or native. Although the optimal temperature (T_{opt}) of PenG hydrolysis is 50°C, the enzyme loses stability above 30°C and must be used in its immobilized form. Other strategies to increase stability include incubation with the thermosome from *Methanocaldococcus jannaschii* and expression as a chimera in order to extend the half-life of PGA in the presence of water–methanol mixtures (4, 5). However, other PGAs with higher stabilities than EcoPGA have been described that do not require the concourse of chaperones for their functional expression. Among the moderately thermostable PGAs are those from *Alcaligenes faecalis* ($t_{1/2}$ [half-life] of 15 min at 55°C), *Bacillus badius* ($t_{1/2}$ of 20 min at 55°C), and *Achromobacter xylosoxidans* ($t_{1/2}$ of 55 min at 55°C), whose thermostabilities arise from additional disulfide bonds, more salt bridges, and additional buried ionic pairs, respectively (6, 7). However, the most thermostable acylases exhibit a preference for aliphatic (penicillins K and dihydroF), rather than aromatic (penicillins G and V), penicillins. For instance, an acylase from the actinomycete *Actinoplanes utahensis* named aculeacin A acylase, with a T_{opt} of 75°C and $t_{1/2}$ at 65°C of 477 min, preferentially hydrolyzes aliphatic penicillins (8, 9). A common strategy for finding novel thermostable enzymes (thermozymes) is to screen their microbial diversity, especially among extremophiles, because they have been naturally selected to resist hostile environments. Therefore, thermozymes possess metabolic pathways and biological processes catalyzed by enzymes that have adapted to those harsh conditions. In addition to their remarkable resistance to thermal denaturation, thermozymes exhibit an above-average resistance to chemical denaturation, such as that caused by organic solvents, detergents, or changes in pH (10). Thus, thermozymes can be extremely stable catalysts suitable for industrial processes, such as the manufacture of 6-APA. In a previous publication (11), we described a gene that encodes a homologue to EcoPGA, which was found in the genome of the extreme thermophile *Thermus thermophilus* HB27 and NAR1 strains: *T. thermophilus* penicillin acylase [*Tth*PAC]). Because of the nature of this *Tth*PAC and the complex maturation that is crucial to reach its functional heterodimeric final conformation, the overexpression of this enzyme in *E. coli* was carried

Received 18 October 2012 Accepted 17 December 2012
Published ahead of print 21 December 2012
Address correspondence to José Berenguer, jberenguer@cbm.uam.es, or Aurelio Hidalgo, ahidalgo@cbm.uam.es.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.03215-12.

doi:10.1128/AEM.03215-12

out in the presence of recombinantly expressed chaperones and calcium chloride. Furthermore, the thermostability of this enzyme clearly surpassed any previously reported values, with a half-life of 9.2 h at 75°C. The initial characterization of the *Tth*PAC substrate specificity revealed a marked preference for penicillins with an aliphatic side chain, especially penicillin K (PenK).

Given the fact that industrial preparations of PenG obtained as bulk materials through fermentation contain up to 3% aliphatic penicillins (which can be hydrolyzed by PenK acylases, but not by PenG acylases) and represent a significant amount in large-scale operations (12), it would be industrially relevant to broaden the substrate specificity of the *Thermus* penicillin K acylase by improving its activity for penicillin G and by obtaining a biocatalyst with naturally enhanced stability that can simultaneously hydrolyze both types of penicillins. Thus, in this work, we report structural studies leading to the inversion of substrate specificity in *Tth*PAC.

MATERIALS AND METHODS

Materials. All reagents used were of analytical grade and were purchased from Sigma (St. Louis, MO) or Merck (Darmstadt, Germany). PenK was kindly provided by Antibióticos S.A. (León, Spain). PenG potassium salt was purchased from Sigma.

Strains. *E. coli* strains DH5α[*supE44 lacU169*(80*lacZ*M15) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*] and BL21(DE3) [F^- *ompT* $hsdS_B(r_B^-\ m_B^-)$ *gal dcm*(DE3)] were used routinely for cloning purposes and protein production, respectively.

Heterologous expression and purification of *Tth*PAC. The *Tthpac* gene (GenBank accession no. TTC1972) was expressed as a His-tagged product devoid of its signal peptide (Δsp*Tth*PAC) in *E. coli* BL21(DE3) cells as reported previously (11). *E. coli* BL21(DE3) strain was transformed with plasmid pET28a, carrying the Δsp*Tthpac* gene (pET28a-Δsp*Tth*PAC), and with plasmid pGro7 (TaKaRa Bio, Inc.), carrying the genes for the GroES/EL chaperone system. Transformants were plated onto LB agar containing kanamycin (30 $\mu g \cdot ml^{-1}$) and chloramphenicol (20 $\mu g \cdot ml^{-1}$).

The inoculum for batch protein production was prepared by an overnight cultivation of the selected clone in 100-ml shake flasks with 20 ml of LB medium at 37°C. For protein production, 1.5 ml of the corresponding inoculum culture was transferred to 150 ml of fresh LB medium supplemented with 10 mM $CaCl_2$ and containing 0.5 $g \cdot liter^{-1}$ of L-arabinose for chaperone induction. Cells were cultivated at 37°C and 200 rpm until they reached a cell density corresponding to an optical density at 600 nm (OD_{600}) of ≈0.6 ± 0.05. Induction was performed by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM, and cultivation was continued for 17 h at 22°C and 200 rpm. Cells were harvested by centrifugation (4,000 × *g* for 15 min) at 4°C. Cell pellets were resuspended in 10 ml of lysis buffer (50 mM sodium phosphate [pH 7.0], 300 mM NaCl) and disrupted by sonication in an ice bath (3 rounds of 10 min, 0.6 Hz, and 40% power), using a Labsonic U sonicator (B. Braun). Soluble and insoluble protein fractions were separated by centrifugation at 12,000 × *g* for 15 min at 4°C. The supernatant containing the His-tagged protein was partially purified from thermolabile *E. coli* proteins by a heat shock treatment of 20 min at 65°C. Thermostable proteins were recovered from the supernatant after centrifugation (12,000 × *g* for 30 min) at 4°C and subjected to immobilized metal ion affinity chromatography (IMAC) using a Talon Cell-through resin (BD Biosciences) that had been equilibrated with lysis buffer. The column was washed once with 10 volumes of the lysis buffer. The His-tagged protein was eluted with 150 mM imidazole; diafiltrated; concentrated with Amicon Ultra-15 10-kDa centrifugal devices (Millipore, Cork, Ireland) using 50 mM sodium phosphate (pH 7.0), 5 mM NaCl, and 0.5 mM $CaCl_2$; and stored at −20°C until use. The use of stabilizing additives, like polyols, was avoided because we observed their interference with the activity assay. Protein content was determined using the Bio-Rad protein assay dye reagent concentrate (Bio-Rad) with bovine serum albumin as the standard.

Enzyme assays. The deacylating activity of *Tth*PAC against penicillins was measured using an endpoint assay based on the reaction of the primary amine group in the resulting 6-aminopenicillanic acid (6-APA) with fluorescamine. Enzymatic reactions were carried out at 65°C using 0.8 to 6.0 μg of purified enzyme. As the optimum pH of *Tth*PAC for the hydrolysis of PenK differs from that of PenG, incubations in the presence of PenK (0.1 to 20 mM) were made in 20 mM MES [2-(*N*-morpholino)ethanesulfonic acid] (pH 5.5), and incubations with PenG (1 to 100 mM) were performed in 20 mM sodium phosphate (pH 7.8). Forty-microliter aliquots of the reaction mixture were taken at regular intervals over 1 h and were immediately frozen in dry ice to stop the reactions. Samples were then supplemented with 140 μl of 200 mM acetate buffer (pH 4.5) and 20 μl of 1 $mg \cdot ml^{-1}$ fluorescamine in acetone. After a 60-min incubation at room temperature, fluorescence was determined (excitation wavelength of 380 nm, emission wavelength of 530 nm) in a FLUOStar Optima plate reader (BMG LabTech, Germany). All experiments were performed in duplicate, and the effect of nonenzymatic hydrolysis of the substrates was adequately subtracted. Hydrolysis rates were determined by using a linear regression of samples taken at different time intervals, and kinetic parameters were calculated by nonlinear fitting of rates to the Michaelis-Menten equation using GraphPad Prism. The results were statistically evaluated by an unpaired two-tailed Student *t* test with a 95% confidence interval ($P < 0.05$) using GraphPad Prism.

Site-directed mutagenesis. The mutations αL188F, αL188R, αS189F, βL24F, and βI57F were introduced on the pET28a-Δsp*Tth*PAC construct using the QuikChange site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions and were confirmed by sequencing. Additional combination mutations were introduced in subsequent rounds of PCR. The primer sequences are given in Table S1 in the supplemental material.

Molecular modeling. Homology modeling was carried out using the automated procedure in Yasara Structure (Yasara Biosciences) running in a computer cluster with 8 parallel processors. Of the templates selected automatically, the structure corresponding to the inactive EcoPGA precursor was not considered (Protein Data Bank [PDB] code 1E3A) because the goal of homology modeling was to identify residues involved in substrate specificity and catalysis; thus, their correct position would only be found in the active form of the enzyme. Homology modeling was finally carried out exclusively with PGA templates (PDB codes 1KEC, 1CP9, 3K3W, 1H2G, and 1GM7). Special care was taken to consider the three structures from different microorganisms available so far, in order not to bias the result. For each of the templates listed above, five alternative models were built. Loops were modeled using an optimized version of the SCWRL program (13, 14). The quality of each model was denoted by the Z-score. The overall Z-scores for all models have been calculated as the weighted averages of the individual Z-scores using the following formula: overall = (0.145 × dihedrals) + (0.390 × packing1D) + (0.465 × packing3D). Finally, the best parts of the models were combined to obtain a hybrid model, with the hope of increasing the accuracy beyond each of the contributors, and a final round of simulated annealing minimization in explicit solvent was performed to generate the final model. Images were created with PyMol v.0.99 (DeLano Scientific).

Docking. Models of the site-directed mutants were built on the basis of the best homology model of *Tth*PAC. Amino acid changes were introduced using Yasara Structure (Yasara Biosciences) (15) running in a computer cluster. The overall fold of the model was energy-minimized using the Yamber3 forcefield. The three-dimensional structures of the substrates were either extracted from the PDB when available or were modeled with Corina from the corresponding Simplified Molecular Input Line Entry System (SMILES) (16). Molecular docking was carried out using the Autodock-based docking tool in Yasara Structure (17). After 25 runs, solutions were clusterized if their root mean square deviation (RMSD)

was >5.0 Å. Binding energies were calculated with Yasara (Yasara Biosciences). Images were created with Pymol v.0.99 (DeLano Scientific).

Matrix-assisted laser desorption–ionization time-of-flight mass spectrometry analysis. The protein of interest was divided into two aliquots. After denaturation of the protein with 8 M urea, one of the aliquots was reduced and alkylated: disulfide bonds from cysteinyl residues were reduced with 10 mM dithiothreitol (DTT) for 1 h at 37°C, and then thiol groups were alkylated with 50 mM iodoacetamide for 1 h at room temperature in the dark. Both aliquots were diluted to reduce urea concentrations to <1.4 M and were digested using sequencing-grade trypsin (Promega, Madison, WI) overnight at 37°C at a 1:5 (wt/wt) trypsin/protein ratio (18). Digestion was stopped by the addition of 1% trifluoroacetic acid (TFA). Whole supernatants were dried down and then desalted until the mass spectrometric analysis. Samples were desalted on ZipTip C_{18} tips (Millipore) using an elution solution of 50% acetonitrile (ACN) in 0.1% TFA.

The protein digest was dried, resuspended in 11 μl of 0.1% formic acid, and analyzed by reverse-phase liquid-chromatography tandem mass spectrometry (RP-LC-MS/MS) in an Easy-nLC II system coupled to an ion trap LTQ Orbitrap Velos Pro mass spectrometer (Thermo Fisher Scientific, Waltham, MA). The peptides were concentrated (online) by reverse-phase chromatography using a 0.1 mm by 20 mm BioBasic C_{18} RP precolumn (Thermo Fisher Scientific) and then separated using a 0.075 mm by 100 mm BioBasic C_{18} RP column (Thermo Fisher Scientific) operating at 0.3 μl/min. Peptides were eluted using a 60-min gradient from 5 to 40% solvent B (solvent A, 0.1% formic acid in water; solvent B, 0.1% formic acid, 80% acetonitrile in water). Electrospray ionization (ESI) was done using a Nanobore stainless steel emitter (inside diameter [ID], 30 μm; Proxeon) interface. The Orbitrap resolution was set at 30,000.

Peptides from the carbamidomethylated sample were detected in survey scans from 400 to 1,600 atomic mass units (amu) (1 μm scans) followed by 10 data-dependent MS/MS scans, using an isolation width of 2 mass-to-charge ratio units, normalized collision energy of 35%, and dynamic exclusion, applied for 30 s. Peptide identification from raw data was carried out using Proteome Discoverer 1.3 (Thermo Fisher Scientific).

For the sample that was not carbamidomethylated, the mass spectrometer was operated in the selected MS/MS ion monitoring mode (SMIM mode). In this mode, the LTQ Orbitrap Velos Pro detector was programmed to perform, along the same entire gradient, a continuous sequential operation in the MS/MS mode on the doubly or triply charged ions corresponding to the peptide(s) selected previously from the theoretical prediction. The MS/MS spectra from the peptide were analyzed by assigning the fragments to the candidate sequence, after a calculation of the series of theoretical fragmentations, according to the nomenclature of the series as described previously (19).

RESULTS

Structural model of *Tth*PAC. In order to build a molecular model of *Tth*PAC, alignments and secondary structure predictions were generated as detailed in Materials and Methods using the FASTA sequence of *Tth*PAC (NCBI accession no. YP_005941) as the input. Using the multiple alignments and the templates detailed in Materials and Methods, we generated a model with a Z-score of −2.082. Due to the low percent identity of the *Tth*PAC sequence with the templates (between 25.7 and 30.5%), an additional model of *Tth*PAC was generated by protein threading using the Phyre server. This threading model was based on an automatically selected template, corresponding to the *Providencia rettgeri* PGA (PrePGA) (PDB: 1CP9). Although the model showed discontinuous regions due to gaps in the alignment, the active site was modeled with a good score. Both models were superposed with templates 1GM7 and 1GM9 in order to study which residues from *Tth*PAC corresponded to those described as having an effect on substrate specificity for EcoPGA (Fig. 1A and 1B, respectively). Residues αPhe146, βPhe24, and βPhe57 have been described to determine the acyl specificity in EcoPGA (20). However, as shown in Fig. 1A and 1B, these positions are occupied by the nonaromatic amino acids αSer189, βLeu24, and βIle57, both in the homology and threading models of *Tth*PAC. Consequently, in order to mimic the EcoPGA aromatic acyl binding pocket, residues αLeu188, αSer189, βLeu24, and βIle57 were chosen to be replaced by phenylalanine. Additionally, it has been reported that upon the binding of PenG, EcoPGA undergoes a conformational change that enables the enzyme to accommodate the substrate by moving away the side chain of residues αArg145 and αPhe146 (21). In order to enhance binding of PenG to the *Tth*PAC enzyme, αLeu188 (the residue structurally aligned with the EcoPGA αArg145) (Fig. 1A and 1B) was selected to be replaced by arginine and combined with the αS189F substitution.

The calcium interaction in the *Tth*PAC model was also investigated because the catalytic activity of *Tth*PAC was enhanced by calcium supplementation of the culture medium at the time of enzyme overproduction (11). The side chain of residues βAsp73 and βAsp76, as well as the carbonyl groups of the main chain of βGlu75 and βGly206, are responsible for calcium coordination in the *Tth*PAC model (Fig. 1C). In comparison to EcoPGA, the first two residues are strictly conserved, whereas the last two are replaced. However, no equivalent was found to βAsp252, which is the sixth coordinated residue in EcoPGA.

Both *Tth*PAC models were also analyzed for structural features that may explain the thermostability of *Tth*PAC. Some characteristics were revealed when we compared the models with the structure of PGAs from mesophiles, such as 1GM9 and 1CP9. First, both at the overall protein level and particularly in the α-helices, the amounts of glutamate, arginine, and lysine residues were approximately 40% higher, suggesting the presence of more salt bridges, which are clear contributors to thermotolerance. Second, more stable α-helices were found in *Tth*PAC, as the frequency of β-branched residues (valine, isoleucine, and threonine) known to be α-helix destabilizers (22) was 67% lower. In addition, *Tth*PAC probably presents a more rigid structure since the content of proline residues was 32 to 42% superior to those of its mesophilic counterparts. Finally, the presence of two closely located cysteines (βCys372 and βCys382) was investigated. These residues were found to be conserved in some sequences of *Tth*PAC homologues in different thermophiles, such as *Thermus aquaticus* or *Meiothermus ruber*. Due to the fact that the low stability of the thiol group at high temperatures conditions the presence of cysteine residues in thermophilic proteins, unless they have a regulatory or structural role, their involvement in a putative disulfide bond was investigated. To that extent, the protein was subjected to differential treatment with DTT and iodoacetamide in order to reduce the putative disulfide bond and to alkylate the resulting cysteine residues prior to tryptic digestion and matrix-assisted laser desorption-ionization time-of-flight (MALDI-TOF) analysis of the resulting fragments. As evidenced in Fig. S1 in the supplemental material, a mass corresponding to a peptide featuring the intrapeptide disulfide bond was found when the protein sample was not reduced, while the corresponding carbamidomethylated peptides were found in the opposite case.

Site-directed mutagenesis and characterization of variants. In order to improve the binding of PenG by *Tth*PAC, the aromatic acyl pocket of EcoPGA was emulated. The amino acid replacements suggested by *Tth*PAC modeling were introduced into the

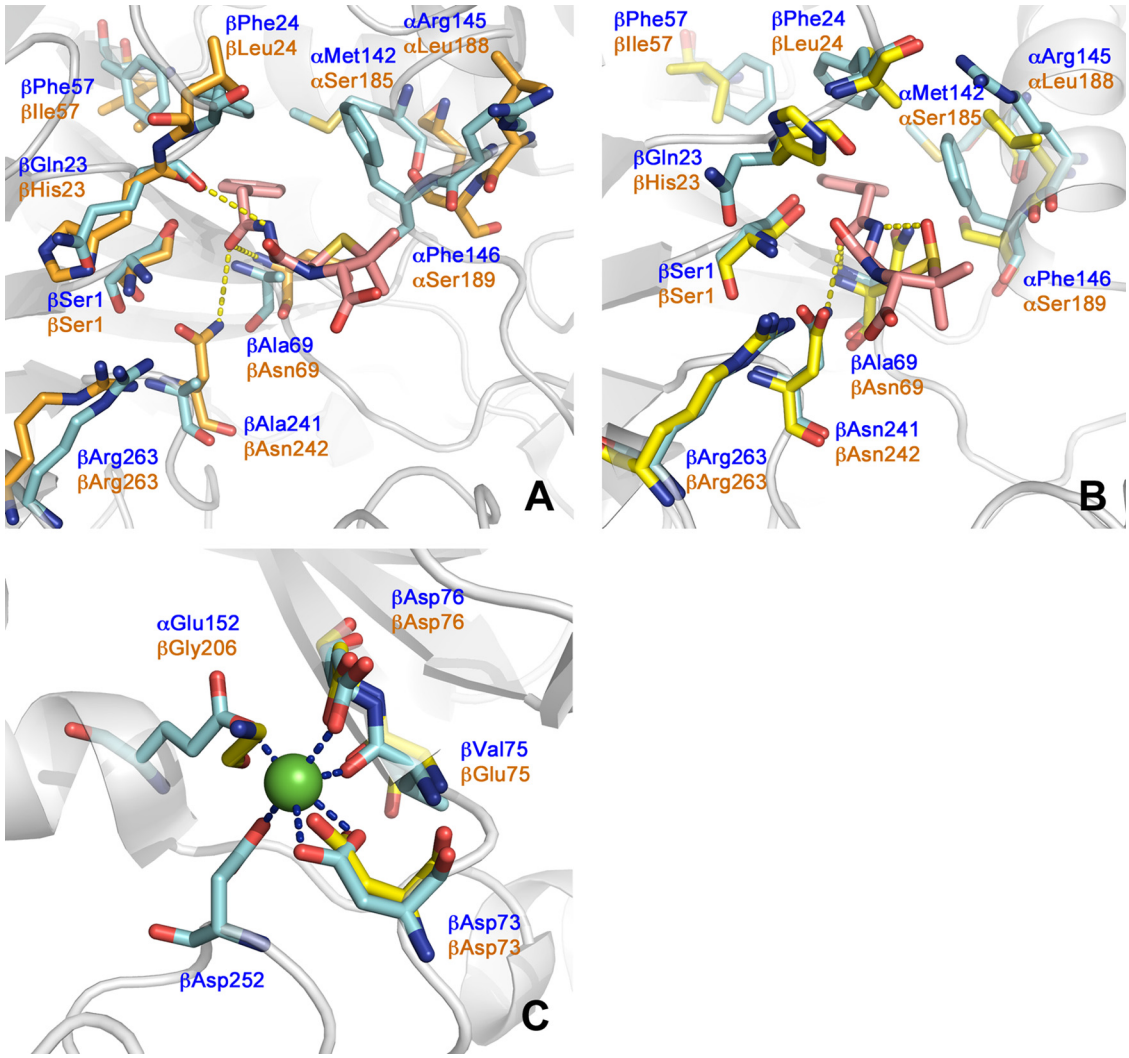

FIG 1 Structural features of *Tth*PAC relevant to binding and catalysis. (A) Residues in the binding pocket of the *Tth*PAC homology model (orange) superimposed on the EcoPGA mutant 1GM7 structure; protein, cyan; cocrystallized PenG, pink. (B) Residues in the binding pocket of the threading model (yellow) superimposed on the wild-type EcoPGA 1GM9 structure; protein, cyan; cocrystallized PenG sulfoxide, pink. (C) *Tth*PAC residues involved in calcium coordination, according to the homology model (orange) and superposition with the equivalent residues in the EcoPGA structure 1GM9 (cyan).

wild-type enzyme, yielding 8 different variants (Table 1). The eight *Tth*PAC variants were expressed recombinantly and were purified as described in Materials and Methods. In Western blot analyses, all purified mutants displayed the bands corresponding to the mature α- and β-subunits, with the exception of the quadruple mutant (M8), which showed only the band corresponding to the unprocessed pro-*Tth*PAC (data not shown). For this reason, the M9 mutant was not kinetically characterized. For the rest of the variants, kinetic analyses of the hydrolysis of PenG and PenK were performed (Table 1). The βL24F mutation (variant M3) had the largest effect on PenG binding, improving K_m 12-fold, while the αL188F and αS189F replacements (variants M1 and M2) improved K_m 10-fold and 9-fold, respectively. The catalytic efficiencies of these three mutants were 6.6-fold for M3 and 3.7-fold for M1 and M2, which are also greater than the efficiency determined for the wild-type enzyme. Replacement βI57F (variant M4) also reduced K_m 4.5-fold, while the catalytic efficiency remained at wild-type values. No improvement was observed with the introduction of αL188R in the M2 variant (variant M5, αL188R-αS189F), and no additive effect was obtained when single replacements were combined, as in variants M6 (βL24F-βI57F), M7 (αL188F-βL24F-βI57F), and M8 (αS189F-βL24F-βI57F). In fact, K_m for PenG was reduced to a small extent and k_{cat} dropped markedly in comparison to the kinetic parameters of the single mutants. Interestingly, the catalytic efficiencies for PenK of the M3 (βL24F), M1 (αL188F), and M2 (αS189F) variants (the best PenG binding variants) were reduced up to 6.7-fold, primarily originated by a 10-fold decrease in k_{cat}.

Docking studies. To elucidate the way in which *Tth*PAC interacted with its substrates, docking studies were carried out using the homology model and all the penicillins that were experimentally tested for deacylation in our previous work (11). A cubic grid of 25 Å was defined considering the position of the PenG ligand in structure 1GM7, after alignment with the *Tth*PAC homology model. The structures of the penicillins were either extracted from the PDB if available or modeled from the corresponding SMILES

TABLE 1 Kinetic constants of the rationally engineered *Tth*PAC variants in the hydrolysis of penicillins K and G

Enzyme variant	Amino acid replacement	Penicillin K[d]			Penicillin G		
		K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)
WT		0.32	5.16	16.1	164	1.56	0.009
M1	αL188F	0.10[a]	0.43[b]	4.3	15.5[c]	0.54[c]	0.034
M2	αS189F	0.10[a]	0.36[b]	3.6	17.8[c]	0.65[c]	0.036
M3	βL24F	0.12[a]	0.32[b]	2.4	13.4[c]	0.85[a]	0.063
M4	βI57F	0.27[a]	4.91[a]	18.2	36.5[b]	0.44[c]	0.012
M5	αL188R-S189F	ND	ND	ND	43.3[b]	0.02[c]	5×10^{-4}
M6	βL24F-I57F	0.05[a]	0.75[b]	15.0	37.6[b]	0.30[c]	0.008
M7	αL189F-βL24F-I57F	0.15[a]	1.43[b]	9.5	44.4[b]	0.13[c]	0.003
M8	αL188R-S189F-βL24F-I57F	ND	ND	ND	ND	ND	ND

[a] Not significantly different from the WT value.
[b] Significantly different from the WT value ($P < 0.05$).
[c] Significantly different from the WT value ($P < 0.001$).
[d] ND, not detected.

using Corina. As a control, PenG was docked into 1GM7, and the resulting best pose was found to be identical to the conformation obtained experimentally (21). For each docking assay, the distance between the attacking oxygen atom of *Tth*PAC βSer1 and the electrophilic carbon of each of the tested penicillins, as well as the binding energy for this interaction, is detailed in Table 2. Of the 4 solutions that contained the octanoyl side chain of PenK located in the acyl binding site, only one was productive in the sense that the ligand was positioned in agreement with a nucleophilic attack by βSer1 (see Fig. S2A in the supplemental material). Furthermore, the PenK side chain carbonyl group was properly oriented toward the putative oxyanion hole. As shown in Fig. S2A in the supplemental material, the stabilizing N atoms forming the oxyanion hole are in the main chain of βAsn325 and the side chain of βAsn498, at a distance of 2.1 to 3.0 Å. From the other two aliphatic penicillins, penicillin F was properly positioned in six solutions that were also productive, while penicillin DHF was productive in only 2 solutions. When PenG was used as a ligand, only 1 solution containing the phenylacetyl moiety docked in the acyl binding site. However, this pose also had the side chain carbonyl group properly oriented toward the putative oxyanion hole, and the distance between the nucleophilic oxygen of βSer1 and the electrophilic carbon of PenG was 2.2 Å (see Fig. S2B in the supplemental material). The acyl moiety of penicillin V also docked in the acyl binding site, and the energy of binding observed was the best of all the dockings performed (8.39 kcal/mol).

Finally, the amino acid replacements corresponding to M1 to M8 were introduced into the *Tth*PAC homology model, and the resulting variants were energy-minimized and docked with PenG.

TABLE 2 Docking parameters of *Tth*PAC interaction with different penicillins

Penicillin ligand	Acyl chain	Binding energy (kcal/mol)	Distance (Å)[a]
V	Phenoxyacetyl	8.39	2.9
G	Phenylacetyl	7.81	2.3
F	Hexenoyl	7.75	2.1
K	Octanoyl	7.53	3.4
DHF	Hexanoyl	6.48	2.4

[a] Distance is between the *Tth*PAC βSer1 oxygen atom and the electrophilic carbon of the corresponding penicillin.

The binding energies of the M1 (αL188F), M2 (αL189F), M3 (βL24F), and M4 (βI57F) models were equal to or larger than that observed for the wild-type (WT) model docked to PenG (Table 3). In each case, a productive solution was obtained where the introduced phenylalanine residues also interacted with the ligand. However, as more replacements were performed in the wild-type *Tth*PAC, such as in M5 to M8, the acyl binding pocket became too hindered to fit the ligand and the binding energy was reduced (Table 3).

DISCUSSION

Due to the industrial interest in PGA, mutational studies have been performed for over 20 years with the aim of altering its thermostability (23, 24) and substrate specificity (25–31). To date, crystal structures of *E. coli* (32), *P. rettgeri* (28), and *A. faecalis* (33) PGAs are available, which have contributed either to rational design or to the explanation of results obtained by random mutagenesis techniques. In this work, taking advantage of the available structures, we have modeled *Tth*PAC and have described some characteristics that may explain its thermal stability as well as the structural features that account for its substrate selectivity.

Thermostability. The thermostability of EcoPGA has been improved by immobilization techniques through the introduction of additional superficial lysines (23), but also at the genetic level by consensus mutagenesis coupled with structural constraints (24).

TABLE 3 Docking parameters of the interaction of *Tth*PAC wild-type and mutant models with penicillin G

*Tth*PAC model	Binding energy (kcal/mol)	Binding residue(s)	Distance (Å)[a]
WT	7.81	Leu188, Ser189, Leu24, Ile57	2.27
M1	7.89	Phe188	2.3
M2	7.95	Phe189	2.4
M3	7.81	Phe24	2.9
M4	8.09	Phe57	2.3
M5	7.68	Arg188	3.3
M6	7.66	Phe24	>5
M7	7.12	Phe24	>5

[a] Distance is between the *Tth*PAC βSer1 oxygen atom and the electrophilic carbon of penicillin G.

Of the 10 amino acid replacements described by Polizzi et al. (24) that resulted in greater EcoPGA thermal stability, only two of them are naturally occurring in *Tth*PAC (βT311P and βV400L). This is not surprising, since the consensus mutagenesis approach was performed with the information obtained from sequence alignments of PGAs phylogenetically unrelated to *Tth*PAC. The thermal stability of *Tth*PAC seems to arise from a combination of different structural features, such as a more rigid structure, additional salt bridges, more stable α-helices, and the existence of a disulfide bond within the *Tth*PAC structure. The thermostabilizing effect of such a bond has already been described for the PGA of *A. faecalis* (34). Consistent with this, *Tth*PAC is an extracytoplasmic protein, and the oxidizing environment of the periplasmic space is suitable for the efficient formation of a disulfide bond. Another contribution to its thermostability is the fact that *Tth*PAC is not a soluble enzyme but has its alpha N terminus anchored to the outer face of the cytoplasmic membrane (11), therefore reducing its unfolding entropy (22).

Substrate specificity. PACs may accept different penicillins as substrates. Among them, PGAs are the most abundant enzymes (e.g., from *E. coli* [35], *Kluyvera cryocescens*, *Kluyvera citrophila* [36], *Arthrobacter viscosus* [37], and *A. faecalis* [34]). These enzymes exhibit a narrow substrate specificity and are unable to hydrolyze aliphatic penicillins due to the aromatic nature of their acyl binding sites (20). The penicillin V acylases from *Bacillus sphaericus* (38), *Bacillus subtilis*, and *Fusarium* sp. (39) are able to hydrolyze penicillin V but, again, not aliphatic penicillins. However, the acylases from *A. utahensis* and *Streptomyces lavendulae* hydrolyze all types of penicillins and show a marked preference for PenK, with K_m values 100- to 400-fold smaller than those for PenG (8, 12). According to the kinetic data available, the PAC from *T. thermophilus* would fit the latter category (the category with the acylases from *A. utahensis* and *S. lavendulae*) (11).

The particular substrate selectivity of *Tth*PAC for alkyl-substituted penicillins can be explained by a structural analysis of its acyl binding pocket. In the present work, we built a set of homology and threading models that helped us understand the substrate selectivity of *Tth*PAC. When performing docking analyses, we experienced some bias with the use of models, as some substrates that experimentally proved to be partially reactive or not reactive at all fit perfectly in the binding pocket of *Tth*PAC. For instance, PenV showed the best binding energy to the wild-type model, while experimentally, it is not the preferred substrate. In addition, PenG docked as effectively as did aliphatic penicillins, and the chromogenic substrate 2-nitro-5-(phenylacetyl amino)benzoic acid (NIPAB) was also found in productive solutions, while *in vitro*, it is not reactive whatsoever (data not shown). The reason for this lies in the structures that were used as templates at the time of modeling. As the best models were built upon PGA structures, the modeled *Tth*PAC catalytic pocket may be partially biased toward having a higher likeness to the pocket of a PGA than it really has. However, despite these three particular biases, the analysis of the acyl binding site was satisfactory overall and was in agreement with experimental data.

In EcoPGA, the acyl specificity is determined by residues αPhe146, βPhe24, and βPhe57, while in *Tth*PAC, the equivalent positions are occupied by residues of an aliphatic nature (Fig. 1). Further examples that link those particular residues to acyl specificity are found throughout PACs. For instance, the PGA from *A. viscosus* shows substitutions of aromatic residues for aliphatic residues in the equivalent positions to EcoPGA αPhe146 and βPhe24, and it also shows K_m increases that are almost 10-fold in the hydrolysis of PenG with respect to EcoPGA (34). In the *Bacillus megaterium* PGA, the position equivalent to βPhe24 is occupied by valine, and the K_m for PenG is increased 100-fold with respect to that of EcoPGA (7). In fact, mutational studies performed on the EcoPGA, in which positions βPhe24 and βPhe57 were saturated, showed that when these residues are replaced by leucine, K_m increased 10-fold and 3-fold, respectively, relative to the wild-type enzyme. In agreement with this, we observed that when either *Tth*PAC αLeu188 or αSer189 (positions equivalent to EcoPGA αPhe146) was replaced with phenylalanine, the K_m for PenG was reduced 10-fold and the catalytic efficiency of the enzyme increased 3.8-fold (Table 1). Also, when the position equivalent to EcoPGA βPhe24 (*Tth*PAC βLeu24) was replaced with phenylalanine, K_m improved significantly (12-fold), while the substitution βI57F reduced K_m 4.5-fold. Docking analyses with M1 (αL188F), M2 (αS189F), M3 (βL24F), or M4 (βI57F) models and PenG showed results that were congruent with the kinetic data. A productive solution was obtained in each case, and the binding energies were equal to or greater than that obtained with the wild-type model (Table 3).

In EcoPGA, and upon binding of PenG, a conformational change occurs in which residues αArg145 and αPhe146, located in the acyl binding pocket, shift their positions, enabling the enzyme to accommodate the substrate. Because of their role in the binding of the ligand, both residues have been repeatedly targeted for mutagenesis in EcoPGA. The replacement of αArg145 with leucine (the amino acid found in *Tth*PAC in this position) has led to a 3-fold increase of the K_m for PenG and an 8-fold reduction in catalytic efficiency (40). The EcoPGA αArg145-Phe146 pair was emulated in the *Tth*PAC M5 variant (αL188R-S189F). In this case, K_m was successfully reduced 4-fold, but the catalytic efficiency for the hydrolysis of PenG dropped more than 10-fold. The corresponding docking assay supported the experimental data, as it showed that the M5 model accommodated PenG in the catalytic pocket in such a way that its electrophilic carbon was positioned away from the βSer1 nucleophilic oxygen.

The residues from EcoPGA described here that are involved in the binding of PenG reside on different chains of the enzyme. However, they are structurally linked through a calcium ion that holds the two subunits together and helps to define the structure of the binding pocket (21). From the six calcium-coordinating residues described in EcoPGA (32), *Tth*PAC only conserves three of them (*Tth*PAC αGlu195, βAsp73, and βAsp76), while βGlu75 and βGly206 are replacements for βVal75 and βPro205, respectively, in EcoPGA. The last two would not necessarily have to be conserved, as they only make main-chain contacts with the calcium. Finally, residue βAsp252 in EcoPGA has no *Tth*PAC counterpart. Crystal structures with a calcium coordination number of five have been described, although they are not common (41). More plausibly, a water molecule took the place of the missing residue in *Tth*PAC, resulting in the usual hexahedral calcium coordination.

Regarding the enzymatic penicillin deacylation by *Tth*PAC, we propose a homologous mechanism to that of EcoPGA. Initially, the substrate is properly positioned and the catalytic serine attacks the scissile amide with the involvement of a hydrogen-bonded water molecule that acts as a base (32). As a result, an oxyanion tetrahedral intermediate is formed and stabilized by basic residues

(βAla69 and βAsn241 in EcoPGA) similarly to serine proteases (21). In *Tth*PAC, it is the backbone amide of βAsn69 and the Nδ atom of βAsn242 that are responsible for the stabilization of the transient negative oxygen ion, while the Nδ lone pair of residue βHis23 stabilizes the charged state of the α-amino group of Serβ1, which occurs during nucleophilic activation (32). The next step in the mechanism is the formation of a recently described "helper" hydrogen bond between the scissile amine group and a hydrogen bond acceptor, which serves to promote pyramidal inversion of the N atom and to facilitate release of the cleaved amine (42). This function is carried out by the main chain carbonyl of βGln23 in EcoPGA or βHis23 in *Tth*PAC (Fig. 1).

The last step of the reaction is the deacylation of the enzyme. It proceeds via a water molecule or any other nucleophile, which may give rise to condensation reactions. Of the two steps in the catalytic mechanism, acylation of the enzyme is the rate-limiting step, thus controlling the rate of the global reaction. In *Tth*PAC, every substitution that was performed reduced the k_{cat} for PenG, while all replacements but βI57F decreased the k_{cat} for PenK. Residue βPhe57 is at the bottom of the substrate binding pocket, but αPhe188, αPhe189, and βPhe24 are located near the catalytic serine. The introduction of a phenylalanine residue in the center of the catalytic pocket that impairs both PenG and PenK hydrolysis suggests that there is probably a space limitation to the accommodation of the ligands, so that the nucleophilic attack is compromised. When the phenylalanine residue was placed at the bottom of the binding site (βPhe57), only PenG hydrolysis was affected. Since residue βPhe57 is far from the catalytic center, something other than a steric issue must be responsible for the reduced k_{cat} for PenG. The introduced phenylalanine probably modifies the orientation of the PenG substrate in a way that impairs the nucleophilic attack or destabilizes the tetrahedral intermediate.

In the present study, we have demonstrated that the selectivity of *Tth*PAC is a consequence of the aliphatic nature of its binding pocket that can indeed be modified with a single mutation. The introduction of phenylalanine residues transforms the *Tth*PAC catalytic site into a proper environment for PenG, although due to the restricted volume of the binding pocket, side effects on its turnover are observed. The fact that such effects are less dramatic when the substitution is introduced at the back of the binding pocket and when the substrate is relatively flexible, such as in the case of PenK, suggests that in order to improve both turnover and binding, an iterative site-directed mutagenesis strategy (43) is required to readjust simultaneously the aromatic nature of the acyl binding pocket and the residues involved in binding of the β-lactam ring and those involved in the catalytic mechanism itself.

ACKNOWLEDGMENTS

This work was supported by the Spanish Ministry of Science (grants CIT no. 010000-2009-29, RyC2006-02441, and BIO2010-18875) and an institutional grant from Fundación Ramón Areces to CBMSO. A generous allocation of computing time at the Scientific Computation Center of the UAM (CCC-UAM) is also acknowledged.

A.H. and J.B. thank M. J. de Soto and E. Sánchez for their assistance.

REFERENCES

1. **Alkema WB, de Vries E, Floris R, Janssen DB.** 2003. Kinetics of enzyme acylation and deacylation in the penicillin acylase-catalyzed synthesis of beta-lactam antibiotics. Eur. J. Biochem. **270**:3675–3683.
2. **Arroyo M, de la Mata I, Acebal C, Castillón MP.** 2003. Biotechnological applications of penicillin acylases: state-of-the-art. Appl. Microbiol. Biotechnol. **60**:507–514.
3. **Waldmann H, Braun P, Kunz H.** 1991. New enzymatic protecting group techniques for the construction of peptides and glycopeptides. Biomed. Biochim. Acta **50**(10–11):S243–S248.
4. **Bergeron LM, Gomez L, Whitehead TA, Clark DS.** 2009. Self-renaturing enzymes: design of an enzyme-chaperone chimera as a new approach to enzyme stabilization. Biotechnol. Bioeng. **102**:1316–1322.
5. **Bergeron LM, Tokatlian T, Gomez L, Clark DS.** 2009. Redirecting the inactivation pathway of penicillin amidase and increasing amoxicillin production via a thermophilic molecular chaperone. Biotechnol. Bioeng. **102**:417–424.
6. **Cai G, Zhu S, Yang S, Zhao G, Jiang W.** 2004. Cloning, overexpression, and characterization of a novel thermostable penicillin G acylase from *Achromobacter xylosoxidans*: probing the molecular basis for its high thermostability. Appl. Environ. Microbiol. **70**:2764–2770.
7. **Rajendhran J, Gunasekaran P.** 2007. Molecular cloning and characterization of thermostable beta-lactam acylase with broad substrate specificity from *Bacillus badius*. J. Biosci. Bioeng. **103**:457–463.
8. **Torres-Bacete J, Hormigo D, Stuart M, Arroyo M, Torres P, Castillón MP, Acebal C, García JL, de la Mata I.** 2007. Newly discovered penicillin acylase activity of aculeacin A acylase from *Actinoplanes utahensis*. Appl. Environ. Microbiol. **73**:5378–5381.
9. **Hormigo D, de la Mata I, Acebal C, Arroyo M.** 2010. Immobilized aculeacin A acylase from *Actinoplanes utahensis*: characterization of a novel biocatalyst. Bioresour. Technol. **101**:4261–4268.
10. **Vieille C, Burdette DS, Zeikus JG.** 1996. Thermozymes. Biotechnol. Annu. Rev. **2**:1–83.
11. **Torres LL, Ferreras ER, Cantero A, Hidalgo A, Berenguer J.** 2012. Functional expression of a penicillin acylase from the extreme thermophile *Thermus thermophilus* HB27 in *Escherichia coli*. Microb. Cell Fact. **11**:105.
12. **Torres-Guzmán R, de la Mata I, Torres-Bacete J, Arroyo M, Castillón MP, Acebal C.** 2002. Substrate specificity of penicillin acylase from *Streptomyces lavendulae*. Biochem. Biophys. Res. Commun. **291**:593–597.
13. **Canutescu AA, Shelenkov AA, Dunbrack RL, Jr.** 2003. A graph-theory algorithm for rapid protein side-chain prediction. Protein Sci. **12**:2001–2014.
14. **Canutescu AA, Dunbrack RL, Jr.** 2003. Cyclic coordinate descent: a robotics algorithm for protein loop closure. Protein Sci. **12**:963–972.
15. **Jones TA, Zou JY, Cowan SW, Kjeldgaard M.** 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A **47**(Pt 2):110–119.
16. **Sadowski J, Gasteiger J, Klebe G.** 1994. Comparison of automatic three-dimensional model builders using 639 x-ray structures. J. Chem. Inf. Comput. Sci. **34**:1000–1008.
17. **Morris GM, Huey R, Olson AJ.** 2008. Using AutoDock for ligand-receptor docking. Curr. Protoc. Bioinformatics **24**:8.14.1–8.14.40.
18. **Villar M, Ortega-Pérez I, Were F, Cano E, Redondo JM, Vázquez J.** 2006. Systematic characterization of phosphorylation sites in NFATc2 by linear ion trap mass spectrometry. Proteomics **6**(Suppl 1):S16–S27.
19. **Roepstorff P, Fohlman J.** 1984. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. Biomed. Mass Spectrom. **11**:601.
20. **Alkema WB, Dijkhuis AJ, de Vries E, Janssen DB.** 2002. The role of hydrophobic active-site residues in substrate specificity and acyl transfer activity of penicillin acylase. Eur. J. Biochem. **269**:2093–2100.
21. **McVey CE, Walsh MA, Dodson GG, Wilson KS, Brannigan JA.** 2001. Crystal structures of penicillin acylase enzyme-substrate complexes: structural insights into the catalytic mechanism. J. Mol. Biol. **313**:139–150.
22. **Li WF, Zhou XX, Lu P.** 2005. Structural features of thermozymes. Biotechnol. Adv. **23**:271–281.
23. **Abian O, Grazú V, Hermoso JA, González R, García JL, Fernández-Lafuente R, Guisán JM.** 2004. Stabilization of penicillin G acylase from *Escherichia coli*: site-directed mutagenesis of the protein surface to increase multipoint covalent attachment. Appl. Environ. Microbiol. **70**:1249–1251.
24. **Polizzi KM, Chaparro-Riggers JF, Vazquez-Figueroa E, Bommarius AS.** 2006. Structure-guided consensus approach to create a more thermostable penicillin G acylase. Biotechnol. J. **1**:531–536.
25. **Forney LJ, Wong DC.** 1989. Alteration of the catalytic efficiency of penicillin amidase from *Escherichia coli*. Appl. Environ. Microbiol. **55**:2556–2560.
26. **Forney LJ, Wong DC, Ferber DM.** 1989. Selection of amidases with novel

substrate specificities from penicillin amidase of *Escherichia coli*. Appl. Environ. Microbiol. **55**:2550–2555.
27. **Martin J, Prieto I, Barbero JL, Pérez-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.
28. **McDonough MA, Klei HE, Kelly JA.** 1999. Crystal structure of penicillin G acylase from the Bro1 mutant strain of *Providencia rettgeri*. Protein Sci. **8**:1971–1981.
29. **Morillas M, McVey CE, Brannigan JA, Ladurner AG, Forney LJ, Virden R.** 2003. Mutations of penicillin acylase residue B71 extend substrate specificity by decreasing steric constraints for substrate binding. Biochem. J. **371**:143–150.
30. **Roa A, Garcia JL, Salto F, Cortes E.** 1994. Changing the substrate specificity of penicillin G acylase from *Kluyvera citrophila* through selective pressure. Biochem. J. **303**:869–875.
31. **Flores G, Soberón X, Osuna J.** 2004. Production of a fully functional, permuted single-chain penicillin G acylase. Protein Sci. **13**:1677–1683.
32. **Duggleby HJ, Tolley SP, Hill CP, Dodson EJ, Dodson G, Moody PC.** 1995. Penicillin acylase has a single-amino-acid catalytic centre. Nature **373**:264–268.
33. **Varshney NK, Kumar RS, Ignatova Z, Prabhune A, Pundle A, Dodson E, Suresh CG.** 2012. Crystallization and X-ray structure analysis of a thermostable penicillin G acylase from *Alcaligenes faecalis*. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. **68**:273–277.
34. **Verhaert RM, Riemens AM, van der Laan JM, van Duin J, Quax WJ.** 1997. Molecular cloning and analysis of the gene encoding the thermostable penicillin G acylase from *Alcaligenes faecalis*. Appl. Environ. Microbiol. **63**:3412–3418.
35. **Cole M.** 1969. Hydrolysis of penicillins and related compounds by the cell-bound penicillin acylase of *Escherichia coli*. Biochem. J. **115**:733–739.
36. **Alvaro G, Fernandez-Lafuente R, Rosell CM, Blanco RM, Garcia-Lopez JL, Guisan JM.** 1992. Penicillin G acylase from *Kluyvera citrophila*: new choice as industrial enzyme. Biotechnol. Lett. **14**:285–290.
37. **Ohashi H, Katsuta Y, Nagashima M, Kamei T, Yano M.** 1989. Expression of the *Arthrobacter viscosus* penicillin G acylase gene in *Escherichia coli* and *Bacillus subtilis*. Appl. Environ. Microbiol. **55**:1351–1356.
38. **Olsson A, Hagström T, Nilsson B, Uhlén M, Gatenbeck S.** 1985. Molecular cloning of *Bacillus sphaericus* penicillin V amidase gene and its expression in *Escherichia coli* and *Bacillus subtilis*. Appl. Environ. Microbiol. **49**:1084–1089.
39. **Sudhakaran VK, Shewale JG.** 1993. Biosynthesis of penicillin V acylase by *Fusarium* sp.: effect of culture conditions. World J. Microbiol. Biotechnol. **9**:233–239.
40. **Alkema WB, Prins AK, de Vries E, Janssen DB.** 2002. Role of alphaArg145 and betaArg263 in the active site of penicillin acylase of *Escherichia coli*. Biochem. J. **365**:303–309.
41. **Katz AK, Glusker JP, Beebe SA, Bock CW.** 1996. Calcium ion coordination: a comparison with that of beryllium, magnesium, and zinc. J. Am. Chem. Soc. **118**:5752–5763.
42. **Syrén PO, Hult K.** 2011. Amidases have a hydrogen bond that facilitates nitrogen inversion, but esterases have not. ChemCatChem. **3**:853–860.
43. **Reetz MT, Wang LW, Bocola M.** 2006. Directed evolution of enantioselective enzymes: iterative cycles of CASTing for probing protein-sequence space. Angew. Chem. Int. Ed. Engl. **45**:1258–1263.