

# **Crystal structures and protein engineering of three different penicillin G acylases from Gram-positive bacteria with different thermostability**

Janine Mayer<sup>1</sup>, Jan Pippel<sup>2</sup>, Gabriele Günther<sup>1</sup>, Carolin Müller<sup>1</sup>, Anna Lauermann<sup>1</sup>, Tobias Knuuti<sup>1</sup>, Wulf Blankenfeldt<sup>2,3</sup>, Dieter Jahn<sup>1</sup>, Rebekka Biedendieck<sup>1,\*</sup>

<sup>1</sup>Institute of Microbiology and Braunschweig Integrated Centre of Systems Biology (BRICS), Technische Universität Braunschweig, Rebenring 56, 38106 Braunschweig, Germany;

<sup>2</sup>HZI - Helmholtz Centre for Infection Research, Structure and Function of Proteins, Inhoffenstraße 7, 38124 Braunschweig, Germany

<sup>3</sup>Institute of Biotechnology, Biochemistry and Bioinformatics, Technische Universität Braunschweig, Spielmannstr. 7, 38106 Braunschweig, Germany;

\* Corresponding author

janine.mayer@tu-braunschweig.de; jan.pippel@helmholtz-hzi.de, g.guenther@tu-braunschweig.de, car.mueller@fz-juelich.de, a.lauermann@tu-braunschweig.de; t.knuuti@tu-braunschweig.de; wulf.blankenfeldt@helmholtz-hzi.de, d.jahn@tu-braunschweig.de, r.biedendieck@tu-braunschweig.de.

## Keywords

Penicillin G acylase, Gram-positive, thermostable, crystal structure, *Bacillus megaterium*

## Abstract

Penicillin G acylase (PGA) catalyzes the hydrolysis of penicillin G to 6-aminopenicillanic acid and phenylacetic acid, which provides the precursor for most semisynthetic penicillins. Most applications rely on PGAs from Gram-negative bacteria. Here we describe the first three crystal structures for PGAs from Gram-positive Bacilli and their utilization in protein engineering experiments for the manipulation of their thermostability. PGAs from *Bacillus megaterium* (*Bm*PGA,  $T_m = 56.0$  °C), *Bacillus thermotolerans* (*Bt*PGA,  $T_m = 64.5$  °C) and *Bacillus* sp. FJAT-27231 (FJAT-PGA,  $T_m = 74.3$  °C) were recombinantly produced with *B. megaterium*, secreted, purified to apparent heterogeneity and crystallized. Structures with resolutions of 2.20 Å (*Bm*PGA), 2.27 Å (*Bt*PGA) and 1.36 Å (FJAT-PGA) were obtained. They revealed high overall similarity, reflecting the high identity of up to approx. 75%. Notably, the active center displays a deletion of more than ten residues with respect to PGAs from Gram-negatives. This enlarges the substrate binding site and may indicate a different substrate spectrum. Based on the structures, ten single-chain FJAT-PGAs carrying artificial linkers were produced. However, in all cases complete linker cleavage was observed. While thermostability remained in the wild type range, the enzymatic activity dropped between 30 to 60 %. Furthermore, four hybrid PGAs carrying subunits from two different enzymes were successfully produced. Their thermostabilities mostly lay between the values of the two mother enzymes. For one PGA increased enzyme activity was observed. Overall, the three novel PGA structures combined with initial protein engineering experiments provide the basis for establishment of new PGA-based biotechnological processes.

## Introduction

Penicillin G acylases (PGA; penicillin amidohydrolase; EC 3.5.1.11) are industrially important enzymes for the production of semi-synthetic  $\beta$ -lactam antibiotics. They catalyze the hydrolysis of penicillin G to 6-aminopenicillanic acid (6-APA) and phenylacetic acid for the subsequent synthesis of semisynthetic penicillins (Srirangan et al. 2013). Compared to penicillin G, semisynthetic penicillins are more stable and easier to absorb. They possess less side-effects and might help to overcome microbial antibiotic resistance (Valle et al. 1991). The importance of PGA is underlined by the outstanding position of  $\beta$ -lactam antibiotics at the world antibiotic market with about 65% market share (Marešová et al. 2014). So far, the biological function of PGAs is not understood since the enzyme does not contribute antibiotic resistance in *Escherichia coli* (Merino et al. 1992). However, there are indications that PGA is used to metabolize penicillin G in order to use the phenylacetic group as carbon source (Kim et al. 2004).

PGAs are found in eukaryotes such as fungi and in Gram-positive and Gram-negative bacteria (Valle et al. 1991). Bacterial PGAs are produced as one single-chain precursor protein consisting of an N-terminal signal peptide (SP), followed by the  $\alpha$ -subunit, a linker sequence and the  $\beta$ -subunit (Panbangred et al. 2000). PGAs from Gram-negative bacteria are transported into the periplasm. For *E. coli* PGA a translocation via the Tat pathway in a folded manner was described (Ignatova et al. 2002). PGAs from Gram-positive bacteria are secreted before folding via the Sec pathway into the environment, such as the cultivation broth (Chiang and Bennett 1967). After secretion and cleavage of the signal peptide, the protein folds into its mature heterodimeric conformation composed of the mature  $\alpha$ - and  $\beta$ -subunit. In the pro-protein, the linker peptide blocks the active-site, consequently, cleavage of the linker also activates the enzyme (Ling Lin Fu et al. 2007), (Tjalsma et al. 2004), (Hewitt et al. 2000).

PGAs belong to the superfamily of N-terminal nucleophile (Ntn) hydrolases. The active site catalytic serine residue is located at the N-terminus of the  $\beta$ -subunit. The serine performs a nucleophilic attack at the scissile amide of the substrate. In addition, the nitrogen atom of this catalytically active serine acts as a base, which temporarily accepts protons (Duggleby et al. 1995), (Dodson and Wlodawer 1998), (Srirangan et al. 2013).

PGA crystal structures were elucidated for the enzymes from different Gram-negative bacteria including *E. coli* (Duggleby et al. 1995), *Alcaligenes faecalis* (Varshney et al. 2012), *Kluyvera citrophila* (Chand et al. 2015) and *Providencia rettgeri* (McDonough et al. 1999). PGAs from Gram-negative and Gram-positive bacteria share only around 30% amino acid sequence identity (Rojviriyaya et al. 2011) and until now no crystal structure of a PGA from a Gram-positive bacterium has been reported, albeit crystallization of *Bacillus megaterium* PGA has been described previously (Rojviriyaya et al. 2011). Some mutagenesis studies using *B. megaterium* PGA resulted in enhanced stability in organic solvents and at high temperatures, when potential surface amino acids were exchanged. Amino acid residue exchange next to the active site changed the catalytic behavior (Wang et al. 2007), (Yang et al. 2000).

In order to make further PGAs from Gram-positive bacteria accessible to application, we recombinantly produced and secreted six *Bacillus* PGAs using *B. megaterium* as the expression host. Out of the six produced enzymes we determined the crystal structures of *B. megaterium* PGA (*BmPGA*), *Bacillus* sp. FJAT-27231 PGA (*FJAT-PGA*) and *Bacillus thermotolerans* PGA (*BtPGA*). Their thermostability was tested and modified by enzyme engineering.

## Materials and Methods

### Bacterial strains

For cloning, *E. coli* strain DH10B (Life Technologies, Carlsbad, USA) was used. Protein production was carried out in *B. megaterium* MS941, which is a derivative of the wild-type strain DSM319 and contains a deletion in the *nprM* gene encoding for the major extracellular protease of *B. megaterium* (Wittchen and Meinhardt 1995).

### Construction of plasmids for recombinant PGA production and secretion

After selection of potential PGAs from Gram-positive bacteria, the corresponding coding sequences were synthesized by Life Technologies (Carlsbad, USA) with an optimized codon usage for *B. megaterium* (Grote et al. 2005) and the restriction sites BsrGI (at the 5' end) und BglII (at the 3' end). The various appropriately digested *pga* genes were ligated into the shuttle vector p3STOP1623hp (Stammen et al. 2010) previously also cut with BsrGI und BglII, resulting in the plasmids pRBBm311 carrying the *btpga* gene (GenBank accession number MH142826) coding for PGA from *B. thermotolerans* (*BtPGA*), pRBBm314 carrying the *anpga* gene (MH142827) coding for PGA from *Bacillus niacini* (*BnPGA*), pRBBm315 carrying the *unc-pga* gene (MH142824) coding for PGA from *Bacillus sp.* UNC438CL73TsuS30 (UNC-PGA), pRBBm316 carrying the *fjat-pga* gene (MH142822) coding for PGA from *B. sp.* FJAT-27231 (FJAT-PGA) and pRBBm317 carrying the *bmaspga* gene (MH142823) coding for PGA from *Bacillus massiliogorillae* (*BmasPGA*). All cloning results were confirmed by DNA sequence analyses. Finally, all plasmids (Tab. S1) were individually transferred into *B. megaterium* MS941 by protoplast transformation as described previously (Biedendieck et al. 2011).

### Construction of hybrid and single-chain PGAs

Hybrid PGAs were constructed by merging different subunits of FJAT-PGA, *BtPGA* and *BmPGA*. The subunits were amplified individually by PCR using the constructed plasmids

described above as template. The used primers (Tab. S2) contain an overlap which was used to join the subunits of different PGAs by hybridization. The DNA of the joint fragments was further amplified in a second PCR step and integrated into the shuttle vector p3STOP1623hp (Stammen et al. 2010) previously digested with BsrGI and EagI. The single-chain FJAT-PGAs were also constructed using the overlapping primer strategy. Thus, subunits of interest were individually amplified without signal peptide, but connected by different amino acid linkers (Tab. 2). Corresponding DNA was integrated into the used primers. The subunits were then joined in reverse order, so that the  $\beta$ -subunit was followed by the  $\alpha$ -subunit according the single-chain *E. coli* PGA of Flores et al. (Flores et al. 2004). The signal peptide of lipase A for SEC-dependent protein export was added by integration into the vector pSPLipA-hp (Stammen et al. 2010) previously digested with KasI and EagI. Between the signal peptide and the  $\beta$ -subunit an additional alanine was integrated by the used primer for better processing of the signal peptide.

### **Recombinant protein production and purification**

All recombinant *B. megaterium* MS941 strains were stored at -80 °C in 20 % (v/v) glycerol solution in LB medium. For recombinant gene expression, 200 mL LB medium (10 g L<sup>-1</sup> tryptone, 5 g L<sup>-1</sup> yeast extract, 5 g L<sup>-1</sup> NaCl) supplemented with 10 mg L<sup>-1</sup> tetracycline and 2.5 mM CaCl<sub>2</sub> (Yang et al. 2006) in baffled flasks were inoculated with 2 mL of an overnight pre-culture and incubated at 37 °C with shaking of 200 rpm. At an optical density at 578 nm (OD<sub>578</sub>) of 0.3 to 0.4, recombinant gene expression was induced with the addition of 0.5 % (w/v) xylose and the culture was further incubated for another 20 to 24 h. The cell-free supernatant containing the exported recombinant protein of interest was obtained by centrifugation (3,300 x g, 20 min, 4 °C) and sterile filtration (0.22  $\mu$ m). For SDS-PAGE analysis, extracellular proteins secreted by *B. megaterium* strains were precipitated from the growth supernatant by the addition of ammonium sulfate. For this purpose, 1.5 mL cell-free culture supernatant were incubated with 660 mg (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> for 2 h at 4 °C and 1000 rpm. After centrifugation (21,000 x g, 30 min, 4 °C), precipitated proteins were suspended in 10  $\mu$ L

of ddH<sub>2</sub>O and analyzed by SDS-PAGE analysis (12 % (w/v) of acrylamide) as described before (Righetti et al. 1990). PGA enzymes were individually purified directly from the cell-free supernatant of the culture by ion exchange chromatography. *Bm*PGA and FJAT-PGA were purified from 200 to 800 mL cell-free supernatant by cation exchange chromatography on SP Sepharose High Performance (GE Healthcare, Freiburg, Germany) (Rojviriyaya et al. 2011). For this purpose, the SP Sepharose was equilibrated with 50 mM sodium phosphate buffer (pH 5.5). Prior loading, the pH value of the cell-free supernatant was also adjusted to 5.5. After loading of the supernatant (200 to 300 mL supernatant per mL SP Sepharose), the column was washed with 50 mM sodium phosphate buffer (pH 5.5) and PGA was eluted stepwise with the identical buffer containing 100, 200, 300, 400, 500, 600 and 700 mM NaCl, respectively. Ten  $\mu$ L of each elution fraction were analyzed by SDS-PAGE as described above (Righetti et al. 1990). For determination of protein concentration, the Bradford Assay was performed (Bradford 1976). *Bm*PGA eluted at 400 and 500 mM NaCl, while FJAT-PGA was found in 400 to 600 mM NaCl fractions. *Bt*PGA from 200 mL of cell-free supernatant was purified by anion exchange chromatography on Q Sepharose High Performance (GE Healthcare). Twenty mM Tris-HCl (pH 8.0) buffer was used for column equilibration and washing. The pH value of the cell-free supernatant was adjusted to 8.0 prior loading. The column was loaded with 200 to 300 mL supernatant per mL Q Sepharose. For elution, 20 mM Tris-HCl (pH 8.0) containing 100, 200, 300, 400, 500 mM NaCl, respectively, was used. *Bt*PGA eluted at 200 to 400 mM NaCl. For the final purification, preparative gel permeation chromatography (GPC) was performed. For this purpose, the PGA solutions containing *Bm*PGA, FJAT-PGA or *Bt*PGA obtained by ion exchange chromatography were concentrated using Amicon centrifugal filters MWCO 10,000 (Merck Millipore, Darmstadt, Germany) to approximately 10 mg mL<sup>-1</sup> and loaded onto a HiLoad 26/600 Superdex 75 PG column (GE Healthcare, 62.5  $\mu$ g protein mL<sup>-1</sup> column volume). *Bm*PGA and FJAT-PGA were chromatographed using 50 mM sodium phosphate buffer (pH 5.5) containing 150 mM NaCl. *Bt*PGA was purified using 20 mM Tris-HCl (pH 8) with 150 mM NaCl. Elution was monitored at 280 and 450 nm. Two mL fractions were collected and 10  $\mu$ L of each elution fraction were

analyzed by SDS-PAGE as described above (Righetti et al. 1990). PGA containing fractions were combined, concentrated to maximal concentrations of 20 mg protein per mL and dialyzed against 10 mM Hepes (pH 7.0). To verify the integrity and identity of the purified proteins, the N-terminal amino acids sequence of the  $\alpha$ - and  $\beta$ -subunit were analyzed by automatic Edman degradation (Edman and Begg 1967).

### **PGA activity assay using 2-nitro-5-[(phenylacetyl)amino]-benzoic acid (NIPAB)**

For PGA activity measurement the alternative substrate 2-nitro-5-[(phenylacetyl)amino]-benzoic acid (NIPAB, Sigma Aldrich (St. Louis, USA)) (Zhang et al. 1986) was used. 90  $\mu$ L of substrate solution (600 mg L<sup>-1</sup> NIPAB in 9.40 mM NaH<sub>2</sub>PO<sub>4</sub>, 40.60 mM Na<sub>2</sub>HPO<sub>4</sub>, pH 7.5) were added per well of a 96-well-plate with flat bottom (Kisker Biotech, Steinfurt, Germany). To start the reaction, 10  $\mu$ L of cell-free supernatant or 0.5 to 1  $\mu$ g of purified enzyme were added per well, mixed and the plate was immediately incubated at 37 °C. Absorption (A) at 405 nm was measured every six seconds for three minutes using a microplate reader (Tecan, Männedorf, Switzerland). The enzymatic activity EA [U mL<sup>-1</sup>] was calculated using the following equation with the reaction volume  $V_R$  [cm<sup>3</sup>], the sample volume  $V_E$  [cm<sup>3</sup>], the layer thickness  $d$  [mm] and the extinction coefficient  $\epsilon$  of 8.98 cm<sup>2</sup>  $\mu$ mol<sup>-1</sup>. A unit is defined as the amount of PGA which hydrolyzes one micromole of NIPAB per minute.

$$EA = \frac{\frac{dA}{dt} \cdot V_R}{\epsilon \cdot d \cdot V_E}$$

### **PGA activity assay using penicillin G (PenG)**

Penicillin G (potassium salt) as alternative substrate for the various recombinant wild-type PGA enzymes and phenylacetic acid were purchased from Sigma Aldrich (St. Louis, USA). Two  $\mu$ g of pure enzyme (FJAT-PGA, *Bt*PGA and *Bm*PGA) were incubated with 10 mM of penicillin G in 50 mM sodium phosphate buffer pH 7.5 at 37 °C with agitation of 500 rpm. Samples were taken after 5, 10, 20, 30 and 40 min of incubation time. The enzymatic

reaction was stopped by adding one volume of methanol (Daumy et al. 1985). For analysis via high-performance liquid chromatography (HPLC) 10  $\mu\text{L}$  of the samples were injected into a Chromolith® HighResolution RP-18e HPLC column, 4.6  $\times$  100 mm (Merck, Darmstadt, Germany) and separated using a Merck-Hitachi LaChrom HPLC system with pump L-7100, autosampler L-7200, Diode Array Detector L-7450 and LaChrom Elite column oven L-2350 (Darmstadt, Germany, Tokyo, Japan). The mobile phase contained 70 % of 50 mM  $\text{KH}_2\text{PO}_4$  pH 3 and 30 % of 90 % acetonitrile in water. The oven temperature was 40  $^\circ\text{C}$  and the flow rate 1 ml  $\text{min}^{-1}$ . Penicillin G and phenylacetic acid were detected at 220 nm. Enzymatic activities (EA) were determined in U  $\text{mg}^{-1}$  by depletion of penicillin G over time.

### **Thermal shift assay**

To determine the thermostability of a protein under appropriate conditions, different buffers were tested in thermal shift assays varying in their pH (pH 4 - 9). To identify the best buffer for each wild-type enzyme, 5  $\mu\text{L}$  of protein solution (2 mg  $\text{mL}^{-1}$ ), 40  $\mu\text{L}$  buffer and 5  $\mu\text{L}$  SYPRO™ Orange protein gel stain (50 x) (Life Technologies) were mixed for each test in individual wells of transparent 96-well plates (Microseal®, Bio-Rad, USA). The fluorescence was measured over a temperature gradient from 10 to 90  $^\circ\text{C}$  using the Thermocycler CFX96 (Bio-Rad, USA). The melting point can be estimated as the minimum of the first negative derivative of the melting curve determined by fluorescence measurement using the software CFX Manager (Bio-Rad, USA). The best buffer conditions for the wildtype PGAs were: *Bm*PGA 20 mM Tris HCl (pH 8), *Bt*PGA 40 mM Tris-HCl (pH 8) with 150 mM NaCl, and FJAT-PGA 40 mM Hepes (pH 7) (Tab. S3). For the hybrid PGAs, thermostability was tested in the buffers used for the wild-type PGAs (Tab. S3). Single-chain PGAs were assayed using 40 mM Hepes (pH 7).

### **Protein crystallization**

PGA crystallization experiments were carried out at 17  $^\circ\text{C}$  using the sitting-drop vapor diffusion method mixing 1  $\mu\text{L}$  of protein (5 to 10 mg  $\text{mL}^{-1}$ ) with 1  $\mu\text{L}$  of reservoir buffer from a

60  $\mu$ L reservoir in Intelli-Plates 96-3 (Art Robbins Instruments, Sunnyvale, USA). Reservoir buffers were from NeXtal JCSG I, JCSG II and Cryos Suite (Qiagen, Venlo, Netherlands). FJAT-PGA crystals were obtained in a solution of 85 mM HEPES (pH 7.5), 8.5% (w/v) PEG 8000 and 15% (v/v) glycerol. *Bt*PGA crystals grew in a solution of 18 mM calcium chloride, 90 mM sodium acetate (pH 4.6), 27% (v/v) MPD (2-methyl-2,4-pentanediol) and 10% (v/v) glycerol while *Bm*PGA crystals arose from a precipitant containing 200 mM  $MgCl_2$  with 20% (w/v) PEG 3350.

### **Data collection, structure determination and refinement**

Whereas no additional cryoprotection was required for crystals of FJAT-PGA and *Bt*PGA, crystals of *Bm*PGA were briefly washed in mother liquor supplemented with 10% (v/v) (2*R*,3*R*)-(-)-2,3-butanediol prior to flash cooling in liquid nitrogen. Diffraction data for FJAT-PGA and *Bm*PGA were collected on beamline PXIII of the Swiss Light Source (SLS, Paul Scherrer Institute, Villigen, Switzerland), and the data set for *Bt*PGA has been obtained at beamline P11 of the PETRAIII synchrotron (DESY, Hamburg, Germany) (Burkhardt et al. 2016). Data for FJAT-PGA and *Bt*PGA were indexed and integrated with XDS (Kabsch 2010) and then scaled with AIMLESS (Evans and Murshudov 2013), but the non-isotropic character of the diffraction pattern of *Bm*PGA crystals required anisotropic processing with the STARANISO-server (Tickle et al. 2018) after indexing and integration with XDS.

Initial phases for FJAT-PGA were determined by molecular replacement with PHASER (McCoy et al. 2007) from the ccp4 suite (Winn et al. 2011), using the crystal structure of PGA from *Providencia rettgeri* (PDB entry 1CP9) (McDonough et al. 1999) as a search model. Subsequent refinement involved alternating rounds of manual adjustments in Coot (Emsley et al. 2010) and computational optimization in phenix.refine (Afonine et al. 2012) of the Phenix software suite (Adams et al. 2010), using automatically assigned fragments and the implemented routine for refining TLS parameters (Schomaker and Trueblood 1968). Once

the refined structure of FJAT-PGA was available, it was used as a molecular replacement model for *Bt*PGA and *Bm*PGA, which were then refined in a similar manner.

Full data collection and refinement statistics can be found in Tab. 1. Coordinates and structure factor amplitudes have been deposited in the Protein Data Bank (Berman et al. 2000) with entry codes 6NVX (FJAT-PGA), 6NVY (*Bt*PGA) and 6NVW (*Bm*PGA).

## Results

### Identification and production of PGA from Gram-positive bacteria

The overarching aim of this study was the production, structural and functional characterization of novel PGAs with different thermostabilities from Gram-positive bacteria. For this purpose, a BLAST search (NCBI) was carried out employing the amino acid sequence of the PGA from *B. megaterium* ATCC14945. The new *pga* genes of interest should encode a signal peptide for the Sec-dependent protein secretion into the growth medium to facilitate the easy purification of the proteins. For the corresponding analyses, the tool SignalP (Petersen et al. 2011) was used. The codon adaptation index (CAI) of the identified genes was analyzed and, where necessary, adapted using the program JCat (Grote et al. 2005). To enable PGA production and secretion in *B. megaterium* a threshold of 0.3 for the CAI was chosen (Yang et al. 2007). Five hypothetical PGA proteins from *B. sp.* FJAT-27231, *B. thermotolerans*, *Bacillus niacini*, *Bacillus sp.* UNC438CL73TsuS30 and *Bacillus massiliogorillae* were identified. They revealed an amino acid identity between 52 and 72 % with the PGA of *B. megaterium*. The corresponding genes were completely synthesized with optimized codon usage and cloned into the *B. megaterium* shuttle vector p3Stop1623hp (Stammen et al. 2010). The general structure of the resulting plasmid for the production of various PGAs is shown in Fig. S1 (Biedendieck et al. 2011).

*B. megaterium* strain MS941 was individually transformed with the newly constructed plasmids, cultivated and the recombinant protein production was induced. Samples were taken 20 hours after induction of the gene expression and the proteins of the cell-free supernatant were precipitated with ammonium sulfate to analyze recombinant PGA production and secretion (Fig. 1). Besides the recombinant PGA of *B. megaterium*, recombinant hypothetical PGAs from *B. sp.* FJAT-27231, *B. thermotolerans*, *B. niacini*, *B. sp.* UNC438CL73TsuS30 and *B. massiliogorillae* were found produced and secreted. Proteins with relative molecular masses of around 25.000 and 62.000 were observed, matching the calculated molecular weight of the PGA  $\alpha$ - and  $\beta$ -subunits. For all recombinant

PGAs, the correct N-terminus of  $\alpha$ - and  $\beta$ -subunit was verified using protein sequencing. The efficient secretion of five novel PGAs clearly indicated that the signal peptides of the other *Bacillus* species were recognized by the *B. megaterium* secretion machinery. Next, the enzyme activity for all PGAs was determined using the cell-free supernatant and a colorimetric assay with 2-nitro-5-[(phenylacetyl)amino]-benzoic acid (NIPAB) as substrate (Fig. 1). Here, the three PGAs from *B. sp.* FJAT-27231 (FJAT-PGA), *B. thermotolerans* (*Bt*PGA) and *B. megaterium* (*Bm*PGA) revealed significant enzyme activities between 0.22 and 0.55 U mL<sup>-1</sup>. The PGA of *B. massiliogorillae* exhibited low enzyme activity while the PGAs from *B. niacini* and *B. sp.* UNC438CL73TsuS30 failed to display the corresponding activity at all. The FJAT-PGA exhibited even a higher secretion level and enzyme activity than the homologous *Bm*PGA. In summary, the database search resulted in two new efficiently produced and secreted PGAs from Gram-positive bacteria with considerable enzymatic activity.

Between *Bm*PGA and the novel PGAs from *B. sp.* FJAT-27231 and *B. thermotolerans* an amino acid sequence identity of 71 % was observed when the signal peptide,  $\alpha$ - and  $\beta$ -subunit and the linker region were included. FJAT- and *Bt*PGA showed an amino acid sequence identity of 75 %. As indicated in Fig. S2, PGA regions with the major differences in the amino acid sequences were in the peptide termini of the two subunits (identity of around 50 %).

#### **Purification of recombinant PGA enzymes directly from the cell-free supernatant**

*B. megaterium* secreted recombinant PGA target proteins with only little amounts of contaminating host proteins (Fig. 1). This allowed for the chromatographic purification of the three PGA variants from *B. sp.* FJAT-27231, *B. thermotolerans* and *B. megaterium* directly from the growth medium (Rojviriyaya et al. 2011). Depending on the net charge (isoelectric point pI) of the protein surface calculated by the ExPASy compute pI/MW tool, anion or cation exchange chromatography was employed. The calculated pI for the FJAT-PGA was 8.28, so it was purified by cation exchange chromatography. *Bt*PGA with a pI of 4.72 was

purified by anion exchange chromatography (Fig. S3). For *Bm*PGA with a pI of 6.08 both methods were tested. The cation exchange chromatography yielded higher amounts of *Bm*PGA with less contaminating proteins (Fig. S3). The highest yield was achieved for the FJAT-PGA, whereas the yield of *Bt*PGA and *Bm*PGA showed around 50 and 70 % of FJAT-PGA yield, respectively (Tab. 2). The yield of *Bt*PGA was already enhanced 5.7-fold by adding 2.5 mM CaCl<sub>2</sub> to the cultivation broth during recombinant protein secretion (Yang et al. 2006). Under the given conditions, the highest specific activity with NIPAB was determined for FJAT-PGA, followed by *Bm*PGA and *Bt*PGA (Tab. 2). Furthermore, the enzymatic activities of purified recombinant PGAs were assayed with the natural substrate penicillin G. FJAT-PGA showed the highest activity (176.5 U mg<sup>-1</sup>), followed by *Bt*PGA (122.3 U mg<sup>-1</sup>) and *Bm*PGA (88.1 U mg<sup>-1</sup>)."

Finally, N-terminal protein sequencing was used to verify the predicted N-terminal sequences after the C-terminal cleavage of the signal peptide from the  $\alpha$ -subunit and the C-terminal cleavage of the linker region from the  $\beta$ -subunit. N-terminal amino acid sequence determination of all our recombinantly produced PGAs did not detect any heterogeneity at the corresponding N-termini.

### **Thermal stability of the different PGAs**

Next, the thermostability of the three purified PGAs, specified by the temperature at which the enzyme is heat denatured, was determined by a thermal shift assay. *Bm*PGA from the mesophilic *B. megaterium* exhibited an unexpectedly high melting point of  $56.0 \pm 0.5$  °C, and the other two novel PGAs were even more thermostable with  $64.5 \pm 0.4$  °C (*Bt*PGA) and  $74.3 \pm 0.2$  °C (FJAT-PGA). Furthermore, all three PGA variants showed a high storage stability at 20 °C. After storage for 28 days, the three PGA variants retained more than 75 % of the initial activity. Incubation of the three wild-type PGAs at the cultivation temperature of 37 °C for 24 h did not result in activity decrease for FJAT-PGA and only up to 5 % loss of activity for *Bm*PGA and *Bt*PGA.

## Crystal structures

No crystal structure of a PGA from Gram-positive bacteria has been elucidated until now. For purification of all PGAs to the homogeneity required for crystallization, gel permeation chromatography (GPC) was employed (Fig. S4). Subsequently, crystallization screening experiments were carried out for the three purified PGAs. Despite the high sequence identity of around 75%, the crystallization conditions were found highly different and also led to different crystal forms, which in retrospect is no surprise given the clearly different surface charge profiles of the three PGAs that also reflects their pIs (Fig. 2). FJAT-PGA crystals were obtained from a solution of 85 mM HEPES (pH 7.5), 8.5% (w/v) PEG 8000 and 15% (v/v) glycerol, giving crystals that belong to space group  $P2_12_12_1$  with one PGA heterodimer in the asymmetric unit (ASU). *Bt*PGA crystallized in a solution of 18 mM calcium chloride, 90 mM sodium acetate (pH 4.6), 27% (v/v) MPD (2-methyl-2,4-pentanediol) and 10% (v/v) glycerol, leading to crystals with H3 symmetry and two *Bt*PGA molecules in the ASU. *Bm*PGA crystallized in a solution of 200 mM  $MgCl_2$  and 20% (w/v) PEG 3350, yielding the same  $P2_1$  unit cell with one PGA heterodimer in the ASU that has been described in a previous report (Rojviriyaya et al. 2011). Crystals of FJAT-PGA diffracted to a very high resolution of 1.36 Å, whereas the diffraction pattern obtained for *Bm*PGA was highly anisotropic, extending to approx. 2.2 Å in the best direction but fading at 3 Å in the worst. This behavior has been described previously (Rojviriyaya et al. 2011), and the diffraction data were therefore reduced with the STARANISO web server (Tickle et al. 2018), which determines an ellipsoidal data cut-off to forward only significant diffraction data to the subsequent steps of crystallographic structure determination.

Initial phasing of the diffraction data obtained from FJAT-PGA crystals by molecular replacement was straightforward. The crystal structure of PGA from *Providencia rettgeri* (PDB entry 1CP9) (McDonough et al. 1999) was chosen as a search model since this protein structure displayed the highest sequence identity to FJAT-PGA (approx. 34% for the complete heterodimer). Alternate rounds of manual refinement in Coot (Emsley et al. 2010) and maximum likelihood refinement at 1.36 Å in phenix.refine (Afonine et al. 2012) led to a

model of excellent geometry and R-factors of  $R_{\text{work}} = 14.8\%$  and  $R_{\text{free}} = 16.7\%$ . The electron density allowed tracing of residues 2 – 204 of the  $\alpha$ - and of residues 1 – 527 of the  $\beta$ -subunit. This suggests that approx. eight C-terminal residues of both chains remained invisible in the final structure, however the exact identity of the C-termini in FJAT-PGA is not known at present. Residual electron density in the active site suggests the presence of an unidentified ligand with dimensions similar to the crystallization buffer HEPES (Fig. S5). However, fitting of a HEPES molecule led to unfavorable interactions, and the respective position was therefore left unoccupied in the final model.

The structural model of *Bt*PGA could be built to a similar extent, using FJAT-PGA as a search model for molecular replacement. In addition to two structural  $\text{Ca}^{2+}$  ions that bind in a similar fashion as in structures of PGAs from Gram-negative bacteria, the crystal form obtained here contained several additional  $\text{Ca}^{2+}$ -binding sites that also participated in the formation of crystal contacts, explaining the requirement for  $\text{CaCl}_2$  in the crystallization buffer. Again, the active site of both copies in the ASU indicated the presence of a ligand that could not be modeled and was hence not included in the final model.

Finally, the model for *Bm*PGA could be refined to satisfactory parameters despite the anisotropic character of the diffraction pattern. Here, the active site was found unoccupied, which may, however, reflect the inferior resolution of these diffraction data. Nevertheless, it was possible to clearly position a significant number of water molecules, which may indicate the value of the anisotropic treatment by STARANISO. Complete data collection and refinement parameters are shown in Tab. 1.

### **Single-chain and hybrid PGAs**

Next, we tried to use the obtained structures to understand basics of the thermostability of the PGAs. Two approaches were used for this purpose. First, since *in silico* models of interchanged  $\alpha$ - and  $\beta$ -subunits revealed no significant clashes, we constructed hybrid PGAs to understand the contribution of the two single subunits of a certain enzyme to its overall thermostability to manipulate thermostability in future protein design experiments. Secondly,

we tried to create single-chain FJAT-PGAs by covalent joining of the  $\alpha$ -subunit to the C-terminus of the  $\beta$ -subunit as shown for *E. coli* PGA before (Flores et al. 2004) to analyze the potential contribution of a single-chain enzyme to thermostability. The expectation was that the monomeric enzyme might be more stable than its corresponding dimeric form.

The plasmids for the production of the hybrid PGAs FJAT- $\alpha$ /Bt- $\beta$ , Bt- $\alpha$ /FJAT- $\beta$ , Bm- $\alpha$ /FJAT- $\beta$ , and FJAT- $\alpha$ /Bm- $\beta$  were constructed by PCR amplification of the single PGA subunit genes, their joining and cloning into the *B. megaterium* shuttle vector p3Stop1623hp (Stammen et al. 2010). *B. megaterium* strain MS941 was transformed with the resulting plasmids, and the corresponding hybrid PGAs were produced and purified (Fig. 3). For all hybrid PGAs the specific activities and thermostabilities were determined (Tab. 2). The specific activities (13.3 to 15.9 U/mg) of the hybrid PGAs were mostly in the same range as the wild-type enzymes (Tab. 2). Interestingly, Bm- $\alpha$ /FJAT- $\beta$  revealed with 15.9 U/mg the highest activity of all tested proteins. In contrast, the FJAT- $\alpha$ /Bm- $\beta$  showed with 5.6 U/mg one of the lowest PGA activities.

For testing the thermostability, the melting points of the hybrid PGAs were analyzed in the three buffers used for wild-type PGAs (Tab. S3). For all hybrid PGAs tested the melting points in the assay were found to be between those of the wild-type PGAs independently of the buffer used. For PGAs FJAT- $\alpha$ /Bt- $\beta$  (70.8 °C) and Bt- $\alpha$ /FJAT- $\beta$  (71.4 °C) they were between that of FJAT-PGA (74.3 °C) and BtPGA (64.5 °C). This was also the case for FJAT- $\alpha$ /Bm- $\beta$  (63.3 °C) where BmPGA (56 °C) contributes the  $\beta$ -subunit. The only exception was Bm- $\alpha$ /FJAT- $\beta$  (57.1 °C) which behaved rather like BmPGA (56 °C). Obviously, both subunits equally contribute to the thermostability and to a degree exhibited by the parental enzyme.

For the construction of a single-chain FJAT-PGA the  $\alpha$ - and  $\beta$ -subunit were amplified by PCR and joined in reverse order ( $\beta$ -linker- $\alpha$ ). Due to the close proximity of the C-terminus of the  $\beta$ -subunit and the N-terminus of the  $\alpha$ -subunit, a short amino acid linker (GARD) according to (Flores et al. 2004) was integrated via a PCR primer. The linker region was further varied as detailed in Tab. 2 to obtain a stable single-chain FJAT-PGA. Additional amino acids were

introduced stepwise (GARDG, GARDGA, GARDGARD) for minimization of potential tension. Furthermore, amino acid residues K1A, D2A, Q3A, K4A of FJAT- $\alpha$ , N535B, A536B, R537B, K538B of FJAT- $\beta$  were omitted in given combinations (Tab. 2) to ensure closer proximity of both termini of the single-chain proteins.

In between the signal peptide, which is responsible for the translocation of the protein through the membrane via the SEC system, and the  $\beta$ -subunit an additional alanine was inserted to enhance the cleavage of the signal peptide after the translocation. This was predicted by the online tool SignalP (Petersen et al. 2011). The various single-chain PGA genes were integrated into the vector pSPLipAhp (Stammen et al. 2010) already containing the signal peptide of lipase A. *B. megaterium* strain MS941 was transformed with these plasmids, the corresponding single-chain PGAs were produced and purified. All single-chain PGAs were produced as monomer with a relative molecular mass of 90,000 instead of the original PGA heterodimer (Fig. 4). However, only portion of the PGA proteins kept the monomeric structure in the elution buffer containing 400 mM NaCl after ion exchange chromatography, and processing into the  $\alpha$ - and  $\beta$ -subunits was visible (Fig. 4). In the assay buffer, all of them were completely processed and revealed almost identical thermostability and a reduced enzyme activity (4.3 to 9.9 U/mg, with 14.4 U/mg for the wild type). Even if we failed to produce stable single-chain PGA, we showed that highly active protein can also be produced, exported and processed with the expression of the  $\alpha$ - and  $\beta$ -subunit in a reversed order ( $\beta$ -linker- $\alpha$ ) and with a highly modified linker.

## Discussion

PGAs represent one of the commercially important classes of enzymes. Until now mostly enzymes from Gram-negative bacteria were investigated and employed for biotechnological purposes. Only a few investigation on the physical properties and their optimization of PGAs from Gram-positive bacteria are found in the literature (Wang et al. 2007), (Yang et al. 2000), (Xu et al. 2018), (Rajendran et al. 2014).

Here, we present the first crystal structures of PGAs from three different Gram-positive bacteria. They are highly similar to each other, superimposing with rmsd values between 0.62 and 0.84 Å over the C- $\alpha$  positions of the complete structures, which clearly reflects the high sequence similarity of the three proteins (approx. 75% identity). The two copies in the ASU of the *Bt*PGA crystals are nearly identical (rmsd of 0.20 Å), however, larger differences exist to PGAs from Gram-negative bacteria: the overall rmsd between the initial search model used for molecular replacement (PGA from *Providencia rettgeri*, PDB entry 1CP9) (McDonough et al. 1999) and FJAT-PGA is 1.84 Å, the well-studied *E. coli* PGA superimposes with an rmsd of 1.87 Å. The three PGAs investigated here display a typical PGA structure, which resembles a triangle with similar edges of approx. 60 to 80 Å length (Fig. 5). The active site of these proteins sits in a deep trough in the center of this triangle and is predominantly positively charged (Fig. 2). This is similar to PGAs from Gram-negative bacteria, and residues that have previously been identified as being required for enzymatic activity such as Ser1B and N245B are also conserved in the structures described here (Fig. 6). Interestingly, however, despite the fact that the overall fold of FJAT-PGA, *Bt*PGA and *Bm*PGA is nearly identical to that of e.g. *E. coli* PGA, the  $\beta$ -subunit of the PGAs from Gram-negatives contains an insertion of more than 10 residues in their active sites (Fig. S2). Whereas the respective region in the three Gram-positive PGAs investigated here folds into two  $\alpha$ -helices connected by a short loop (FJAT-PGA: helix S371B-T278B, loop D379B-A384B, helix T385B-S392B), the Gram-negative PGAs display a much larger loop (*E. coli* PGA:  $3_{10}$ -helix F367B-Y371B, loop S372B-S390B, helix T391B-V400B) that protrudes into

the active site and therefore reduces its size (Fig. 7). It is therefore conceivable that the Gram-positive PGAs investigated here have a different substrate spectrum than previously characterized PGAs from Gram-negative bacteria. This is also corroborated by the observation that modeling of a complex of the PGAs from Gram-positive bacteria using published structure of *E. coli* PGA in complex with  $\beta$ -lactams such as penicillin G (PDB entry 1GM7) (McVey et al. 2001) leads to several unacceptably short distances between protein and ligand (Fig. 7). This is on the one hand due to differences surrounding the phenylacetyl binding site, making it much narrower than in *E. coli* PGA, and on the other a consequence of alterations resulting from the inserted loop in PGAs from Gram-negatives. The structural differences in this region also lead to the presence of an arginine residue in PGAs from Gram-positive bacteria (R381B). This residue could interact with a negatively charged moiety of a substrate molecule, but  $\beta$ -lactams would have to bind in a slightly different orientation with respect to the position observed in *E. coli* PGA. Interestingly, such an interaction seems to exist between the arginine and the unidentified ligand, as indicated by the location of its electron density (Fig. S5). Together, these observations justify further investigation of the substrate spectrum of PGAs from Gram-positive bacteria.

Two of the crystalized enzymes revealed high thermostability, even one of the corresponding Bacilli hosts was not all known as thermophile. Most likely due to the extracellular nature of PGAs thermostability might provide an advantage in the different habitats of the host strains. It is difficult to identify certain amino acid residues responsible for the observed thermostability via amino acid sequence comparisons with PGAs from mesophilic Bacilli. However, we observed a clear cut difference in the surface charge between the meso- and thermophilic strains (Fig. 2). Both thermophilic PGAs were either more negatively or positively charged compared to the mesophilic enzyme.

Interestingly, hybrid PGA daughter enzymes mostly behaved in their thermostability and activity like the mean between the two mother enzymes. Similar observations have been made for chimeric endoglucanases (Zheng et al. 2018), Superoxide Dismutase (Retnoningrum et al. 2016), L-aspartase (Sheng et al. 2005), cellobiohydrolases (Voutilainen

et al. 2014),  $\alpha$ -amylase (Parashar and Satyanarayana 2016) and alkaline phosphatases (Sasajima et al. 2014). Finally, we tried to establish active single-chain PGAs as outlined before for *E. coli* in order to further improve the thermostability (Flores et al. 2004). We failed to obtain stable *Bacillus* single-chain PGAs. The *E. coli* PGA investigation was based on a screen using a random linker sequence fused to the  $\beta$ - and  $\alpha$ -protein domains. Obviously, the different linkers tested in this study did not sustain stable single-chain formation. Future experiments might use a similar selection approach as outlined for *E. coli* PGA.

## **Authors' contributions**

J.M., W.B., D.J. and R.B. conceived of the study and wrote the paper. J.M., G.G., C.M., A.L., T.K. and R.B. carried out experiments and analyzed the results. J.P. and W.B. conducted crystal x-ray measurements and solved the structure of the three PGAs. All authors read and approved the final version of the manuscript.

## **Acknowledgments**

We would like to thank Beate Jaschok-Kentner (HZI, Braunschweig, Germany) for N-terminal protein sequencing. The Swiss Light Source (SLS, Paul Scherrer Institute, Villigen, Switzerland) and the PETRAIII synchrotron (DESY, Hamburg, Germany) are acknowledged for providing access to their facilities.

## **Funding**

This work was funded by the German Research Foundation (DFG), within the priority program, SPP1934 “Dispersitäts-, Struktur- und Phasenänderungen von Proteinen und biologischen Agglomeraten in biotechnologischen Prozessen.”

## **Compliance with Ethical Standards**

**Conflict of Interest:** The authors declare that they have no conflict of interest.

**Ethical approval:** This article does not contain any studies with human participants or animals performed by any of the authors.

## References

- Adams PD, Afonine P V., Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D Biol Crystallogr* 66:213–221. doi: 10.1107/S0907444909052925
- Afonine P V., Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr Sect D Biol Crystallogr* 68:352–367. doi: 10.1107/S0907444912001308
- Balci H, Ozturk MT, Pijning T, Ozturk SI, Gumusel F (2014) Improved activity and pH stability of *E. coli* ATCC 11105 penicillin acylase by error-prone PCR. *Appl Microbiol Biotechnol* 98:4467–4477. doi: 10.1007/s00253-013-5476-7
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Res* 28:235–242. doi: 10.1093/nar/28.1.235
- Biedendieck R, Borgmeier C, Bunk B, Stammen S, Scherling C, Meinhardt F, Wittmann C, Jahn D (2011) Systems biology of recombinant protein production using *Bacillus megaterium*, 1st edn. Elsevier Inc.
- Bradford MM (1976) A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye-binding. *Anal Biochem* 72:248–54. doi: 10.1016/0003-2697(76)90527-3
- Burkhardt A, Pakendorf T, Reime B, Meyer J, Fischer P, Stübe N, Panneerselvam S, Lorbeer O, Stachnik K, Warmer M, Rödig P, Göries D, Meents A (2016) Status of the crystallography beamlines at PETRA III. *Eur Phys J Plus* 131:0–8. doi: 10.1140/epjp/i2016-16056-0
- Chand D, Varshney N, Ramasamy S, Panigrahi P, Brannigan JA, Wilkinson AJ, Suresh CG (2015) Structure mediation in substrate binding and post-translational processing of penicillin acylases: information from mutant structures of *Kluyvera citrophila* penicillin G acylase. *Protein Sci* 24:1660–1670. doi: 10.1002/pro.2761
- Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr Sect D Biol Crystallogr* 66:12–21. doi: 10.1107/S0907444909042073
- Chiang C, Bennett RE (1967) Purification and properties of penicillin amidase from *Bacillus megaterium*. *J Bacteriol* 93:302–308
- Daumy GO, Danley D, McColl AS, Apostolakos D, Vinick FJ (1985) Experimental evolution of penicillin G acylases from *Escherichia coli* and *Proteus rettgeri*. *J Bacteriol* 163:925–932
- Dodson G, Wlodawer A (1998) Catalytic triads and their relatives. *Trends Biochem Sci* 23:347–352. doi: 10.1016/S0968-0004(98)01254-7
- Duggleby H, Tolley S, Hill C, Dodson E (1995) Penicillin acylase has a single-amino-acid catalytic centre. *Nature* 373:264–68
- Edman P, Begg G (1967) A protein sequenator. *Eur J Biochem* 1:80–91. doi: 10.1285/i20705948v4n2p203
- Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr Sect D Biol Crystallogr* 66:486–501. doi: 10.1107/S0907444910007493
- Evans PR, Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr Sect D Biol Crystallogr* 69:1204–1214. doi: 10.1107/S0907444913000061
- Flores G, Soberón X, Osuna J (2004) Production of a fully functional, permuted single-chain penicillin G acylase. *Protein Sci* 13:1677–1683. doi: 10.1110/ps.03436604
- Grote A, Hiller K, Scheer M, Münch R, Nörtemann B, Hempel DC, Jahn D (2005) JCat: A novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res* 33:526–531. doi: 10.1093/nar/gki376
- Hewitt L, Kasche V, Lummer K, Lewis RJ, Murshudov GN, Verma CS, Dodson GG, Wilson

- KS (2000) Structure of a slow processing precursor penicillin acylase from *Escherichia coli* reveals the linker peptide blocking the active-site cleft. *J Mol Biol* 302:887–98. doi: 10.1006/jmbi.2000.4105
- Ignatova Z, Hörnle C, Kasche V, Nurk A (2002) Unusual signal peptide directs penicillin amidase from *Escherichia coli* to the tat translocation machinery. *Biochem Biophys Res Commun* 291:146–149 . doi: 10.1006/bbrc.2002.6420
- Jurrus E, Engel D, Star K, Monson K, Brandi J, Felberg LE, Brookes DH, Wilson L, Chen J, Liles K, Chun M, Li P, Gohara DW, Dolinsky T, Konecny R, Koes DR, Nielsen JE, Head-Gordon T, Geng W, Krasny R, Wei GW, Holst MJ, McCammon JA, Baker NA (2018) Improvements to the APBS biomolecular solvation software suite. *Protein Sci* 27:112–128. doi: 10.1002/pro.3280
- Kabsch W (2010) XDS. *Acta Crystallogr Sect D Biol Crystallogr* 66:125–132. doi: 10.1107/S0907444909047337
- Kim HS, Kang TS, Hyun JS, Kang HS (2004) Regulation of penicillin G acylase gene expression in *Escherichia coli* by repressor PaaX and the cAMP-cAMP receptor protein complex. *J Biol Chem* 279:33253–33262. doi: 10.1074/jbc.M404348200
- Kleywegt GJ (1997) Validation of protein models from C( $\alpha$ ) coordinates alone. *J Mol Biol* 273:371–376. doi: 10.1006/jmbi.1997.1309
- Lakowitz A, Krull R, Biedendieck R (2017) Recombinant production of the antibody fragment D1.3 scFv with different *Bacillus* strains. *Microb Cell Fact* 16:1–18. doi: 10.1186/s12934-017-0625-9
- Ling Lin Fu, Zi Rong Xu, Wei Fen Li, Jiang Bing Shuai, Ping Lu, Chun Xia Hu (2007) Protein secretion pathways in *Bacillus subtilis*: Implication for optimization of heterologous protein secretion. *Biotechnol Adv* 25:1–12. doi: 10.1016/j.biotechadv.2006.08.002
- Marešová H, Plačková M, Grulich M, Kyslík P (2014) Current state and perspectives of penicillin G acylase-based biocatalyses. *Appl Microbiol Biotechnol* 98:2867–2879. doi: 10.1007/s00253-013-5492-7
- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) *Phaser* crystallographic software. *J Appl Crystallogr* 40:658–674 . doi: 10.1107/S0021889807021206
- McDonough M a, Klei HE, Kelly J a (1999) Crystal structure of penicillin G acylase from the Bro1 mutant strain of *Providencia rettgeri*. *Protein Sci* 8:1971–81. doi: 10.1110/ps.8.10.1971
- McVey CE, Walsh M a, Dodson GG, Wilson KS, Brannigan J a (2001) Crystal structures of penicillin acylase enzyme-substrate complexes: structural insights into the catalytic mechanism. *J Mol Biol* 313:139–150. doi: 10.1006/jmbi.2001.5043
- Merino E, Balbás P, Recillas F, Becerril B, Valle F, Bolivar F (1992) Carbon regulation and the role in nature of the *Escherichia coli* penicillin acylase (*pac*) gene. *Mol Microbiol* 6:2175–2182. doi: 10.1111/j.1365-2958.1992.tb01391.x
- Panbangred W, Weeradechapon K, Udomvaraphant S, Fujiyama K, Meevootisom V (2000) High expression of the penicillin G acylase gene (*pac*) from *Bacillus megaterium* UN1 in its own *pac* minus mutant. *J Appl Microbiol* 89:152–157. doi: 10.1046/j.1365-2672.2000.01093.x
- Pantoliano MW, Petrella EC, Kwasnoski JD, Lobanov VS, Myslik J, Graf E, Carver T, Asel E, Springer BA, Lane P, Salemme FR (2001) High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J. Biomol. Screen.* 6:429–440
- Parashar D, Satyanarayana T (2016) A chimeric  $\alpha$ -amylase engineered from *Bacillus acidicola* and *Geobacillus thermoleovorans* with improved thermostability and catalytic efficiency. *J Ind Microbiol Biotechnol* 43:473–484. doi: 10.1007/s10295-015-1721-7
- Pei J, Grishin N V, Road FP (2014) PROMALS3D: Multiple Sequence Alignment Methods enhanced with evolutionary and 3-dimensional structural information. *Methods Mol Biol* 1079:263–271. doi: 10.1007/978-1-62703-646-7
- Petersen TN, Brunak S, Von Heijne G, Nielsen H (2011) SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat Methods* 8:785–786. doi: 10.1038/nmeth.1701
- Rajendran K, Sekar S, Mahadevan S, Kumar Shanmugam B, Jeyaprakash R, Paramasamy

- G, Mandal AB (2014) Biological real-time reaction calorimeter studies for the production of penicillin G acylase from *Bacillus badius*. *Appl Biochem Biotechnol* 172:3736–3747. doi: 10.1007/s12010-014-0800-y
- Retnoningrum DS, Rahayu AP, Mulyanti D, Dita A, Valerius O, Ismaya WT (2016) Unique Characteristics of Recombinant Hybrid Manganese Superoxide Dismutase from *Staphylococcus equorum* and *S. saprophyticus*. *Protein J* 35:136–144. doi: 10.1007/s10930-016-9650-5
- Righetti P, Gianazza E, Gelfi C, Chairi M (1990) Gel electrophoresis of proteins: a practical approach. In: Hames B, Rickwood D (eds) 2nd editio. IRL Press - Oxford University Press, New York, pp 149–214
- Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42:320–324. doi: 10.1093/nar/gku316
- Rojviriyaya C, Pratumrat T, Saper MA, Yuvaniyama J (2011) Improved X-ray diffraction from *Bacillus megaterium* penicillin G acylase crystals through long cryosoaking dehydration. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 67:1570–1574. doi: 10.1107/S1744309111040462
- Sasajima Y, Kohama Y, Kojima-Misaizu M, Kurokawa N, Hara Y, Dong J, Ihara M, Ueda H (2014) Simultaneous Retention of Thermostability and Specific Activity in Chimeric Human Alkaline Phosphatases. *Mol Biotechnol* 56:953–961. doi: 10.1007/s12033-014-9774-9
- Schomaker V, Trueblood KN (1968) On the rigid-body motion of molecules in crystals. *Acta Crystallogr Sect B Struct Crystallogr Cryst Chem* 24:63–76. doi: 10.1107/S0567740868001718
- Sheng Y, Li S, Gou X, Kong X, Wang X, Sun Y, Zhang J (2005) The hybrid enzymes from  $\alpha$ -aspartyl dipeptidase and l-aspartase. *Biochem Biophys Res Commun* 331:107–112. doi: 10.1016/J.BBRC.2005.03.140
- Srirangan K, Orr V, Akawi L, Westbrook A, Moo-Young M, Chou CP (2013) Biotechnological advances on Penicillin G acylase: Pharmaceutical implications, unique expression mechanism and production strategies. *Biotechnol Adv* 31:1319–1332. doi: 10.1016/j.biotechadv.2013.05.006
- Stammen S, Müller BK, Korneli C, Biedendieck R, Gamer M, Franco-Lara E, Jahn D (2010) High-yield intra- and extracellular protein production using *Bacillus megaterium*. *Appl Environ Microbiol* 76:4037–4046. doi: 10.1128/AEM.00431-10
- Tickle IJ, Flensburg C, Keller P, Paciorek W, Sharff A, Vonrhein C, Bricogne G (2018) STARANISO
- Tjalsma H, Antelmann H, Jongbloed JDH, Braun PG, Darmon E, Dorenbos R, Dubois F, Westers H, Zanen G, Quax WJ, Kuipers OP, Bron S, Hecker M, Dubois JF, Dijk JM Van (2004) Proteomics of protein secretion by *Bacillus subtilis*: separating the "secrets" of the secretome. *Microbiol Mol Biol Rev* 68:207–233. doi: 10.1128/MMBR.68.2.207
- Valle F, Balbás P, Merino E, Bollvar F (1991) The role of penicillin amidases in nature and in industry. *Trends Biochem Sci* 16:36–40
- Varshney NK, Suresh Kumar R, 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. doi: 10.1107/S1744309111053930
- Voutilainen SP, Nurmi-Rantala S, Penttilä M, Koivula A (2014) Engineering chimeric thermostable GH7 cellobiohydrolases in *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 98:2991–3001. doi: 10.1007/s00253-013-5177-2
- Wang J, Zhang Q, Huang H, Yuan Z, Ding D, Yang S, Jiang W (2007) Increasing synthetic performance of penicillin G acylase from *Bacillus megaterium* by site-directed mutagenesis. *Appl Microbiol Biotechnol* 74:1023–1030. doi: 10.1007/s00253-006-0752-4
- Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr Sect D Biol Crystallogr* 67:235–242. doi:

10.1107/S0907444910045749

- Wittchen KD, Meinhardt F (1995) Inactivation of the major extracellular protease from *Bacillus megaterium* DSM319 by gene replacement. *Appl Microbiol Biotechnol* 42:871–877. doi: 10.1007/BF00191184
- Xu G, Zhao Q, Huang B, Zhou J, Cao F (2018) Directed evolution of a penicillin V acylase from *Bacillus sphaericus* to improve its catalytic efficiency for 6-APA production. *Enzyme Microb Technol* 119:65–70. doi: 10.1016/j.enzmictec.2018.08.006
- Yang S, Huang H, Li S, Ye Y, Wan L, Zhang F, Yuan Z (2000) Enhancing penicillin G acylase stability by site-directed mutagenesis. *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)* 32.6:581–585
- Yang Y, Biedendieck R, Wang W, Gamer M, Malten M, Jahn D, Deckwer W-D (2006) High yield recombinant penicillin G amidase production and export into the growth medium using *Bacillus megaterium*. *Microb Cell Fact* 5:36. doi: 10.1186/1475-2859-5-36
- Yang Y, Malten M, Grote A, Jahn D, Deckwer W-D (2007) Codon Optimized *Thermobifida fusca* Hydrolase Secreted by *Bacillus megaterium*. *Biotechnol Bioeng* 96:780–794. doi: 10.1002/bit
- Zhang Q, Zhang L, Han H, Zhang Y (1986) A method for screening penicillin G acylase-producing bacteria by means of 2-nitro-5-phenylacetaminobenzoic acid test paper. *Anal Biochem* 156:413–416. doi: 10.1016/0003-2697(86)90274-5
- Zheng F, Vermaas J V., Zheng J, Wang Y, Tu T, Wang X, Xie X, Yao B, Beckham GT, Luo H (2018) Characterizing Activity and Thermostability of GH5 Endoglucanase Chimeras from Mesophilic and Thermophilic Parents. *Appl Environ Microbiol*. doi: 10.1128/AEM.02079-18

## Figure legends

**Figure 1: Extracellular proteins of recombinant *B. megaterium* after the production of various PGAs from Gram-positive bacteria and the corresponding PGA activity.** a) SDS-PAGE analysis of the extracellular proteins secreted by *B. megaterium* strains carrying plasmids for the production of PGA from *B. thermotolerans* (Bt), *Bacillus niacini* (Bn), *Bacillus sp.* UNC438CL73TsuS30 (UNC), *Bacillus sp.* FJAT-27231 (FJAT), *Bacillus massiliogorillae* (Bmas), *B. megaterium* (Bm) and an empty vector control (EV). The secreted proteins were precipitated after 20 h of cultivation and separated by 10 % SDS-PAGE, Coomassie Brilliant Blue stained in comparison to a protein standard (M). b) PGA enzyme activity (EA) of extracellular proteins secreted by *B. megaterium* strains shown in a).

**Figure 2: Front and backside view of the electrostatic potential mapped to the molecular surfaces of FJAT-PGA, BtPGA and BmPGA,** displayed at  $\pm 5 \text{ k}_b\text{Te}^{-1}$  (blue: positive, red: negative). The electrostatic potential was calculated with APBS (Jurrus et al. 2018). This figure and all other molecular representations have been prepared with PyMol (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

**Figure 3: SDS-PAGE of hybrid PGAs compared to the wild-type enzymes.** *B. megaterium* expressing the different PGA variants was cultivated in 200 mL of LB medium at 37 °C, cell-free extract obtained after centrifugation and filtration was purified by anion exchange chromatography using Q Sepharose for BtPGA or cation exchange chromatography using SP Sepharose for the other variants. HP7 = FJAT- $\alpha$ /Bt- $\beta$ , HP10 = Bt- $\alpha$ /FJAT- $\beta$ , HP22 = Bm- $\alpha$ /FJAT- $\beta$  and HP23 = FJAT- $\alpha$ /Bm- $\beta$ .

**Figure 4: SDS-PAGE of single-chain FJAT-PGAs compared to the wild-type FJAT-PGA in buffers of different salinity.** Five  $\mu\text{g}$  of each protein in 400 mM NaCl + 50 mM sodium phosphate buffer (pH 5.5) [odd numbers] and 10 mM Hepes buffer (pH 7) [even numbers] are applied on SDS gel. 1/2: FJAT-PGA, 3/4: SC1, 5/6: SC25, 7/8: SC26, 9/10: SC27, 11/12: SC35, 13/14: SC37, 15/16: SC38, 17/18: SC4, 19/20: SC28, 21/22: SC30.

**Figure 5: Front and back view of FJAT-PGA.** The  $\alpha$ - and  $\beta$ -subunits are shown in yellow and magenta, respectively. N and C mark the position of N- and C-termini. The active center is located at the N-terminus of the  $\beta$ -subunit.

**Figure 6: Comparison of the active sites of FJAT (top) and *E. coli* PGA (bottom; cross-eyed stereo plot).** Residues of the  $\alpha$ - and  $\beta$ -subunits are shown in yellow or cyan and magenta or green, respectively. Note that the *E. coli* protein is an inactive variant (N241B has been replaced by alanine) in complex with the substrate penicillin G (PDB entry 1GM7) (McVey et al. 2001).

**Figure 7: Comparison of the molecular surfaces of FJAT (top) and *E. coli* PGA (bottom), colored according to their electrostatic potential at  $\pm 5 \text{ k}^b\text{Te}^{-1}$ .** The inserts show a magnification of the active sites. Penicillin G, present in the *E. coli* protein (PDB entry 1GM7) (McVey et al. 2001), is shown in stick representation and has been modeled into FJAT-PGA to indicate that the active site of this enzyme is not ideally suited to accommodate this substrate. The green loop shows an insertion in in the  $\beta$ -subunit PGAs from Gram-negative bacteria that reduces the active site pocket of these enzymes.