A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEMBIOCHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

Accepted Article

Title: Structure and catalytic mechanism of a bacterial Friedel-Crafts acylase

Authors: Tea Pavkov-Keller, Nina G. Schmidt, Anna Żądło-Dobrowolska, Wolfgang Kroutil, and Karl Gruber

To be cited as: *ChemBioChem* 10.1002/cbic.201800462

Link to VoR: http://dx.doi.org/10.1002/cbic.201800462

A Journal of ChemPubSoc Europe

WILEY-VCH

www.chembiochem.org

Structure and catalytic mechanism of a bacterial Friedel-Crafts acylase

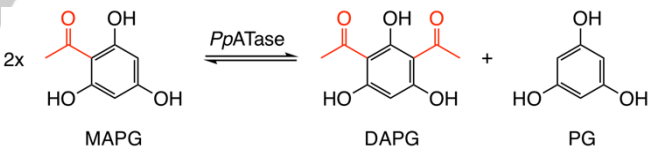
Tea Pavkov-Keller,[a,b] Nina G. Schmidt,[a,c] Anna Żądło-Dobrowolska,[c] Wolfgang Kroutil,*[a,c,d] and Karl Gruber *[a,b,d]

Abstract: C–C bond forming reactions are key transformations to set up the carbon framework of organic molecules. In this context, the Friedel-Crafts acylation is commonly used for the synthesis of aryl ketones, which are common motifs in many fine chemicals and natural products. A bacterial multi-component acyltransferase from *Pseudomonas protegens* (*Pp*ATase) catalyzes such a Friedel-Crafts C-acylation of phenolic substrates in aqueous solution reaching up to >99% conversion without the need of CoA-activated reagents. We determined x-ray crystal structures of the native and ligand-bound complex. This multimeric enzyme consists of three subunits – PhlA, PhlB and PhlC which are arranged in a $Phl(A_2C_2)_2B_4$ composition. The structure of a reaction intermediate obtained from crystals soaked with the natural substrate monoacetylphloroglucinol together with site-directed mutagenesis studies revealed that only residues from the PhlC subunits are involved in the acyl transfer reaction, with Cys88 very likely playing a significant role during catalysis. These structural and mechanistic insights form the basis of further enzyme engineering efforts towards enhancing the substrate scope of this enzyme.

Introduction

Acyltransferases belong to the EC 2.3 groups of enzymes. This subclass contains enzymes that transfer acyl groups from a donor molecule to the hydroxyl, amino or thiol group of an acceptor molecule, forming either esters, amides or thioesters. Different acyltransferase classes have been identified among living organisms that utilize for instance 1-*O*-acylglucosides[1], acylated-acyl carrier proteins[2] or quinic acid ester[3] as high energy activated acyl donors. The most abundant group of acyltransferases requires mostly acyl-CoA derivative as a donor substrate catalyzing various reactions involved in the primary and secondary metabolism[4]. In bacteria, CoA-dependent *O*- and *N*-acylation plays a key role in detoxification of antibiotics, such as chloramphenicol[5], aminoglycosides[6], streptothricin[7] and phosphinothricin[8]. Acyltransferases are also involved in numerous biosynthetic pathways, especially in the biosynthesis of membrane phospholipids[9], wax esters and triacylglycerols[10] and polyketides[11] and also play role in lysozyme resistance[12]. In contrast to the *O*- and *N*- acylation, natural *C*-acylation reactions are scarce. So far, only a couple of acyltransferases, *e.g.* from *Pseudomonas protegens* (*Pp*ATase) and *Pseudomonas fluorescens* have been reported to perform C–C bond formation in the biosynthesis of the antibiotically active polyketide 2,4-diacetylphloroglucinol (DAPG)[13]. The biosynthesis of DAPG is regulated by the phlACBDEFGHI gene cluster, which is divided into regulatory genes phlEFGHI and the biosynthetic operon phlACBD[14]. For many years, biocatalytic applications were limited to gene phlD encoding a type-III polyketide synthase which was employed for the *in vivo* production of the DAPG-precursor PG in either *Pseudomonas* sp. or *E. coli* under controlled conditions in bioreactors[15].

Very recently, the scope of applications was extended to phlACB, which encodes the cofactor-independent acyltransferase *Pp*ATase mentioned above. It catalyzes the reversible disproportionation of monoacetylphloroglucinol (MAPG) into phloroglucinol (PG) and DAPG in a divergent reaction (Scheme 1). *Pp*ATase is a multi-component enzyme and is catalytically active without addition of cofactors like CoA or ATP. A functional enzyme is only obtained upon expression of the entire phlABC operon, whereas mixing and incubation of the individually expressed proteins – PhlA, PhlB, and PhlC – did not lead to the disproportionation of MAPG[13a, 16].



Scheme 1. Natural reaction catalyzed by the *Pseudomonas protegens* acyltransferase (*Pp*ATase) involved in the biosynthesis of 2,4-diacetylphloroglucinol (DAPG).

It was shown that *Pp*ATase accepts various *C*- or *O*-acyl donors such as isopropenyl acetate and transfers an acetyl moiety to a phenolic acceptor in a Friedel-Crafts-type acetylation reaction[16b, 17]. The enzyme complex also shows chemical reaction promiscuity and accepts aniline derivatives as substrates for amide formation[18].

Here, we report on the crystal structure determination of *Pp*ATase in its substrate-free form as well as after soaking the crystals with MAPG. The structural results indicate that only one of the three subunits (PhlC) is involved in catalyzing the acylation reaction. Site-directed mutagenesis studies and activity measurements complement the structural data and enable the proposal of a catalytic mechanism.

[a] Dr. T. Pavkov-Keller, Dr. N. G. Schmidt, Prof. Dr. W. Kroutil, Prof. Dr. K. Gruber
Austrian Centre of Industrial Biotechnology (ACIB)
Petersgasse 14, 8010 Graz, Austria
[b] Dr. T. Pavkov-Keller, Prof. Dr. K. Gruber
Institute of Molecular Biosciences, University of Graz
Humboldtstrasse 50, A-8010 Graz, Austria
[c] Dr. N. G. Schmidt, Dr. A. Żądło-Dobrowolska, Prof. Dr. W. Kroutil
Department of Chemistry, Organic & Bioorganic Chemistry,
University of Graz
Heinrichstrasse 28/2, A-8010 Graz, Austria
[d] Prof. Dr. W. Kroutil, Prof. Dr. K. Gruber
BioTechMed-Graz, Mozartgasse 12/II, 8010 Graz Austria

Supporting information for this article is given via a link at the end of the document.

Results and Discussion

Overall structure

We determined x-ray crystal structures of a CoA-independent acyltransferase produced by *Pseudomonas protegens* DSM 19095 (*Pp*ATase) using diffraction data from two different crystal forms at 2.8 and 3.4 Å resolution, respectively (Table 2 and Supporting Information Figure 1). The structures were solved by molecular replacement involving extensive density modification as well as automated and manual rebuilding (see Methods). It had been known before that *Pp*ATase is a multimeric enzyme consisting of three subunits – PhlA, PhlB and PhlC[36]. The final structures showed the hexagonal crystal form (space group $P6_122$) to contain two copies of each of those subunits and the orthorhombic crystal form (space group $P2_12_12_1$) to contain eight copies of each subunit in the asymmetric unit, respectively.

In both structures, an analysis of protein-protein interactions within the crystal using the EBI-Pisa server[19] yielded a hetero-dodecamer with four copies of each subunit as the most likely biologically active oligomer of *Pp*ATase. Thus, the hexagonal crystal form contains half a dodecamer in the asymmetric unit with the other half being generated by a crystallographic diad. The orthogonal crystals, on the other hand, contain two copies of the full dodecamer in the asymmetric unit. Closer inspection of the inter-chain contacts indicated that PhlA and PhlC each form strongly interacting homodimers in the crystal (Supporting Information Figure 2). The composition of the multimeric enzyme complex is thus best described as $PhlI(A_2C_2)_2B_4$, in which the four copies of PhlB mediate the binding of two PhlA and two PhlC dimers (Figure 1b) to form the complete oligomer (Figure 1a and 1b).

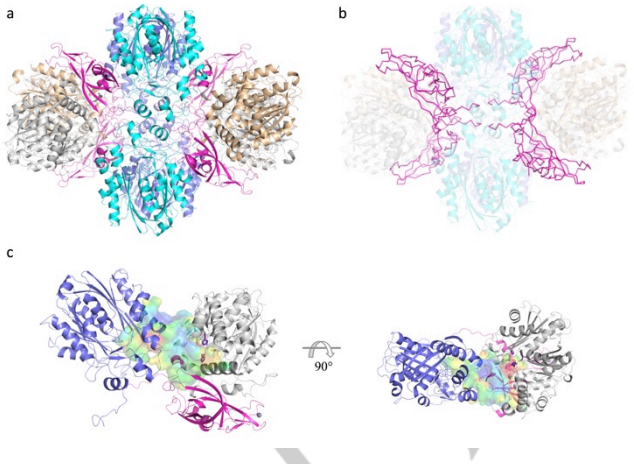


Figure 1. Structural analysis of *Pp*ATase. **(a)** Structure of the hetero-dodecamer $PhlI(A_2C_2)_2B_4$ (PhlA = blue and cyan, PhlB = magenta, PhlC = grey and wheat). **(b)** PhlB mediates the binding of two PhlA and two PhlC dimers. Two N-terminal tails of PhlB (up to Arg28) is involved in tight interactions with adjacent PhlA and PhlC molecules as well as with the N-terminal tail of a neighboring PhlB molecule. PhlB is shown in ribbon representation. **(c)** The continuous cavity between PhlA, PhlB and PhlC subunit trimer. The active site of PhlC (in violet, residues shown in stick representation) is situated adjacent to the cavity of PhlA. The long loop in PhlB (Glu74-Val87, magenta) is closing the cavity from the other side. The cavity is shown in surface representation (red-hydrophobic to blue-hydrophilic). See Methods for detailed descriptions.

A cavity analysis identified four, large, contiguous cavities per dodecamer which are lined by residues from one particular copy of PhlA, PhlB and PhlC. In this hetero-trimeric arrangement the proposed active sites of PhlA and PhlC (see below) are adjacent to each other on one side of the cavity, whereas PhlB (especially through a long loop formed by residues Glu74 to Val87) closes the cavity on the opposite side (Figure 1c).

Comparison with other structures

We used the Dali server[20] as well as PDBeFold[21] to identify structurally similar proteins in the PDB. According to these analyses, the closest structural homologs to PhlA are hydroxymethylglutaryl-CoA synthases (*e.g.* the enzyme MvaS from *Myxococcus xanthus*, PDB-entry: 5hwq)[22] and two β-oxoacyl-(acyl-carrier-protein) synthases: from *Aquifex aeolicus* (PDB-entry: 2ebd, RIKEN Structural Genomics/Proteomics Initiative) and from *E. coli* (PDB-entry: 1hnh)[23]. The sequence identity of PhlA with these homologs is below 24% with root-mean-square-deviations (rmsd) of approximately 2 Å.

Previous sequence analyses had suggested that PhlA could be involved in the first step of diacetyl-phloroglucinol (DAPG) biosynthesis, *i.e.* the formation of acetoacetyl-CoA from acetyl-CoA[24]. Inspections of the cavity observed in our structures in the vicinity of PhlA suggests that there is indeed enough space for acetyl-CoA binding. However, key residues necessary to catalyze this Claisen-type condensation (a cysteine and a glutamic acid) are not present in PhlA. Compared to homologous enzymes, the crucial cysteine corresponds to Gly115 in PhlA and the active site glutamate aligns with Cys83 in PhlA. We aimed at obtaining a complex structure by soaking *Pp*ATase with acetyl-CoA, but the soaked crystal did not diffract sufficiently well for structure determination.

Structural analysis of PhlB suggested a close similarity to a protein of unknown function from *Sulfolobus solfataricus* (PDB-entry: 3irb, rmsd 2.6 Å, seq-id: 18%)[25]. This protein from the DUF35 family (Pfam PF01796) exhibits a two-domain architecture comprising an N-terminal, rubredoxin-like zinc ribbon and a C-terminal oligonucleotide/oligosaccharide-binding (OB) fold domain[26] with an additional N-terminal helical segment possibly involved in protein-protein interactions[25] (Supporting information Figure 3). For members of this protein family, a general role in fatty acid and polyketide biosynthesis as acyl-CoA binding proteins has been predicted[25].

Although a similar domain organization is observed in PhlB and the zinc ribbon with its Cys_4 metal ion binding site is conserved, severe differences are evident in the OB-domain and at the N-terminus. Instead of two N-terminal helices, PhlB exhibits an elongated N-terminal tail and a short, kinked α-helix. In the dodecameric arrangement of *Pp*ATase, this region of PhlB (up to Arg28) is not accessible to the solvent but buried within the enzyme complex structure, where it is involved in tight interactions with adjacent PhlA and PhlC molecules as well as with the N-terminal tail of a neighboring PhlB molecule (Figure 1b). The same is true for the two long loop regions within the OB domain (residue 73 to 89 and residues 127-136).

Finally, the sequence of PhlC shows that it belongs to the thiolase superfamily[27], although sequence identities to

members of this enzyme family are below 45%. The structure of this subunit exhibits an α/β-hydrolase-type fold and is most similar to structures of the thiolase-like protein ST0096 from *Sulfolobus tokodaii* (PDB-entry: 4yzo, rmsd 1.6 Å, seq-id: 26%) and the SCP2 thiolase from *Trypanosoma brucei* (PDB-entry: 4bi9, rmsd 1.7 Å, seq-id: 24%)[28]. In thiolases, a cysteine and a histidine residue (corresponding to Cys88 and His347 in PhlC) are highly conserved and were found to be important for the enzymatic reaction. Beside those two residues, the active site cavity of PhlC is lined by amino acid residues His56, Asn87, His144, Trp211, Tyr298 and Ser349, which may play a role in substrate binding or catalysis (Supporting Information Figure 7). Even in the untreated (hexagonal) crystal, residual density was observed at the sidechain of Cys88 indicating that this residue is at least partially acetylated.

Recently, the structure of an archaeal acetoacetyl-CoA thiolase/HMG-CoA synthase (HMGCS) complex from *Methanothermococcus thermolithotrophicus* was reported[29]. The arrangement of the individual protein subunits within this complex is very similar to that in *Pp*ATase: the HMGCS subunit corresponds to PhlA in *Pp*ATase (sequence identity 38.5%), whereas the thiolase subunit corresponds to PhlC (sequence identity 28.9%). The complex also contains a protein from the DUF35 family that is related to PhlB in *Pp*ATase (sequence identity 26.7%). Sequence alignments of the three separate subunits are shown in Supporting Information Figures 4-6. In contrast to the *PpATase*, however, this complex utilizes acetyl-CoA as an acyl donor and is involved in the mevalonate pathway. The HMGCS subunit is involved in the exergonic condensation of acetoacetyl-CoA and acetyl-CoA to 3-hydroxy-3-methylglutaryl-CoA. The key Cys- and Glu-residues necessary to catalyze this Claisen-type condensation are indeed present in HMGCS (Cys114 and Glu82), whereas they are missing in PhlA as mentioned above (Supporting Information Figure 8). The presence of the active Cys in the thiolase subunit is preserved in both complexes (Cys88 for PhlC and Cys85 in the thiolase subunit). Soaking of the thiolase/HMGCS complex with acetyl-CoA (PDB-entry: 6esq) revealed the binding of CoA at the subunit interface comprising residues of all three proteins [29]. This region is significantly different when compared with the structure of *Pp*ATase (Supporting Information Figure 9) showing altered relative positions of individual secondary structure elements, the presence of additional residues in the PhlB region of *Pp*ATase, two additional short β-strands in HMGCS as well as a lack of CoA-binding residues in *Pp*ATase. The overall fold of PhlB is very similar to the protein from the DUF35 family in the thiolase/HMGCS complex (Supporting Information Figure 3). The elongated N-terminus of PhlB that is buried within the *Pp*ATase complex structure, however, is completely missing in the other structure and notable conformational changes are observed in the opposite loop region.

Structure of the complex with MAPG

In order to identify the subunit responsible for the observed transferase activity, we performed crystal-soaking experiments. Orthorhombic crystals of *Pp*ATase were soaked with the native acetyl donor/acceptor MAPG and the structure of the complex was solved at 3.4 Å resolution. Residual density was observed in the vicinity of residue Cys88 in four out of the eight crystallographically independent copies of the PhlC subunit (Supporting Information Figure 10). This density was compatible with an acetylated Cys88 and with a (deacetylated) phloroglucinol molecule bound at this site (Figure 2). The resulting structure revealed that only residues from PhlC interact with the bound phloroglucinol. Specifically, the sidechains of His56 as well as Tyr124 and His347 form hydrogen bonds with the hydroxyl groups at positions 1 and 3 of the aromatic ring, respectively (Figure 2). In contrast to that, the hydroxyl group at position 5 is apparently not involved in any direct interactions with PhlC. Instead, it points towards a small cavity lined by mostly hydrophobic residues (such as Phe148, Leu209 and Leu300).

The carbonyl oxygen of the Cys88-bound acetyl group is positioned in the "oxyanion hole" of the α/β-hydrolase-type fold forming hydrogen bonds with the mainchain amide groups of Cys88 itself and of Gly385. Additional polar stabilization may be provided by the helix dipole oriented favorably towards Cys88 (the "nucleophile elbow" present in this type of fold[30]). The acetyl group is oriented more or less parallel to the aromatic ring and its carbonyl carbon atom is appropriately positioned for an electrophilic attack on the C-6 atom of the substrate with distances between the two atoms ranging from 3 to 3.5 Å (Figure 2). Its methyl group points towards Phe148.

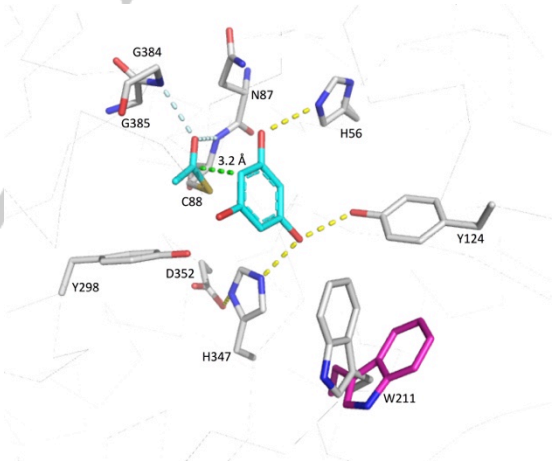


Figure 2. Active site structure of PhlC after soaking with MAPG. This depicts the intermediate formed upon acetyl transfer from the substrate to Cys88 of the enzyme (the transforming co-product PG and the transferred acetyl group are depicted in cyan). Trp211 can act as a lid covering the entrance to the active site (closed: grey; open: magenta). Hydrogen bonds are shown as yellow dashed lines, the contact between the carbonyl carbon of the acetyl group and C-6 of the substrate aromatic ring is indicated as a green dashed line. The carbonyl oxygen of the Cys88-bound acetyl group is positioned in the "oxyanion hole" of the α/β-hydrolase-type fold forming hydrogen bonds (indicated in the light blue dashed lines) with the main chain amide groups.

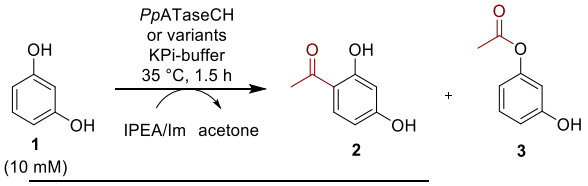
Whereas the conformation of most residues forming the active site of PhlC is similar in both structures, especially Trp211

stands out as it adopts an open conformation in the untreated crystal and closes the cavity upon substrate binding in the soaked crystal (Figure 2).

Mutagenesis of PhlC active site residues

The analysis of the PhlC active site identified Cys88, His144, Asn87, His56, Ser349, Tyr124 Tyr298, Asp352, His347 and Trp211 as a group of mostly polar or charged residues lining the cavity (Supporting Information Figure 7). Those residues were selected for site-directed mutagenesis and we analyzed the activity of the corresponding enzyme variants in the acetylation of resorcinol **1** using isopropenyl acetate (IPEA) as acetyl donor and imidazole (see Methods). The results are shown in Table 1 as relative amounts of the educt resorcinol **1**, the C-acetylation product 4-acetyl-resorcinol **2** and the O-acetylation product 3-hydroxyphenyl acetate **3**.

Table 1. Activity of *Pp*ATase variants for the Friedel-Crafts acetylation of benzene-1,3-diol (resorcinol, **1**) and specific activity for the natural reaction



Entry	*Pp*ATase variant	Acetylation using IPEA/Im[a] 1 [%]	2 [%]	3 [%]	Specific activity [mU/mg][b]
1	w/o enzyme	68	n.d.	32	n.d.
2	wt enzyme	41	44	15	44
3	H56A	81	n.d.	19	n.d.
4	H56S	81	n.d.	19	n.d.
5	N87A	84	6	10	4
6	C88A	74	n.d.	26	n.d.
7	C88S	69	n.d.	31	n.d.
8	H144A	79	5	16	n.d.
9	H144S	81	n.d.	19	n.d.
10	W211A	31	66	3	16
11	W211F	48	40	12	24
12	Y298A	79	3	18	1
13	Y298V	82	n.d.	18	n.d.
14	Y298F	82	3	15	5
15	H347F	77	1	22	<1
16	S349A	73	n.d.	27	n.d.
17	D352V	73	1	26	<1

[a]Conditions: cell-free extract of *Pp*ATase or variant (30 mg protein), resorcinol (**1**, 10 mM), potassium phosphate buffer (50 mM, pH 7.5, total volume 1 mL), IPEA (100 mM) and imidazole (100 mM added from a 1 M stock solution prepared in the reaction buffer), 35 °C, 1.5 h and 750 rpm. The relative amounts of **1-3** were determined by HPLC according to standard curves with authentic samples. n.d. not detected.[b] Specific activity was determined spectrophotometrically following the disproportionation of MAPG into DAPG and PG

As already shown previously[16b], the reaction using the wildtype enzyme produces predominantly the C-acetylation product **2** (entry #2), while the non-enzymatic reaction (entry #1) yields only the O-acetylation product **3** under these conditions. Replacing Cys88 by alanine or serine completely abolishes ATase activity (entries #6 and #7). The same is true for replacements of His56 (entries #3 and #4), whereas exchanges of Asn87 (entry #5), His144 (entries #8 and #9), Tyr298 (entries #12-#14), His347 (entry #15), Ser349 (entry #16) and Asp352 (entry #17) yielded variants which retained minute activities in some cases. As anticipated, the exchange of Trp211 by alanine or phenylalanine did not abolish the enzymatic activity of *Pp*ATase (entries #10 and #11). The variant W211A, however, significantly shifted the product spectrum in favor of the C-acetylation product.

Proposed mechanism of acyl-transfer

*Pp*ATase catalyzes the reversible acetylation of monoacetylphloroglucinol (MAPG) into phloroglucinol (PG) and 2,4-diacetylphloroglucinol (DAPG)[13]. It has been shown that only a multi-component complex consisting of PhlA, PhlB and PhlC subunits performs a disproportionation of MAPG[13a, 16]. The crystal structures presented here revealed that only residues from PhlC interact with the bound PG, strongly suggesting that PhlA and PhlB are not directly involved in the actual acyl transfer step. Instead, these subunits are very likely required for the formation of a properly folded and functional dodecameric enzyme complex. This finding also supports previous suggestions regarding their possible involvement in preceding steps of DAPG biosynthesis[24].

Combining the structural results with activity data obtained for enzyme variants (Table 1), a plausible catalytic mechanism can be proposed (Figure 3). The observation of an acetylated Cys88 in both crystal structures and the lack of activity measured for the cysteine to alanine variant indicates that Cys88 very likely plays a significant role during catalysis. We suggest that thiol group of Cys88 attacks the acyl moiety of the donor and subsequently forms a covalent intermediate similar to acyl-enzyme intermediates formed by serine or cysteine hydrolases[46]. As in these hydrolases, the nucleophilic attack is facilitated by stabilizing the ensuing negative charge at the acyl-oxygen atom by polar interactions within the “oxyanion hole”. There is no clear indication of a base in the vicinity of Cys88, which could activate the thiol by deprotonation. It is reasonable, however, that the sidechain is already deprotonated to some extent at neutral or slightly basic pH-values. Additional stabilization of the thiolate is possible through polar interactions with the OH-group of Tyr298 and the imidazole moiety of His347, although the distances to both groups are larger (>3.6 Å) than observed for typical hydrogen bonds. The proposed formation of an acyl-enzyme intermediate is also consistent with our previous finding that conversions with DAPG/MAPG as acetyl donor did not yield any phenyl acetate derivatives and that the enzyme also catalyzes an intermolecular Fries-rearrangement[16b].

The second step of the reaction involves the transfer of the acyl moiety from Cys88 to the aromatic ring of an acceptor molecule. As mentioned above, the carbonyl carbon of the cysteine bound

acetyl group is appropriately positioned to attack the C-6 atom of the bound phloroglucinol in the complex structure (Figure 3). Electronic activation of the aromatic ring most likely involves deprotonation of the phenolic OH group either at C-1 (by His56) and/or at C-3 (by the diad His347-Asp352). We have previously shown that – in addition to phloroglucinol – resorcinol but not phenol can act as an acyl acceptor[16b] indicating that at least two OH-groups (at positions 1 and 3) are necessary for substrate binding and/or activity. In the structure, the OH-group at C-5 of phloroglucinol does not form any polar interactions with the enzyme and it can indeed be replaced by other (alkyl) substituents in *Pp*ATase substrates[16b]. The rearomatization step by an intramolecular proton transfer does very likely not involve any enzyme intervention.

PhlC of *Pp*ATase differs from other, more common thiolases by the presence of the tryptophan residue Trp211. This residue appears to have a lid function, because its conformation significantly changes upon binding of the substrate (Figure 3). This conformational flexibility could very well influence the activity and above all the substrate specificity of the enzyme. Regarding substitutions at the aromatic ring, the active site cavity appears to provide more space for attachments at C-4 as compared to C-5. This is in good agreement with the observed preference for substituents at C-4 of the resorcinol core structure over the C-5 position[16b].

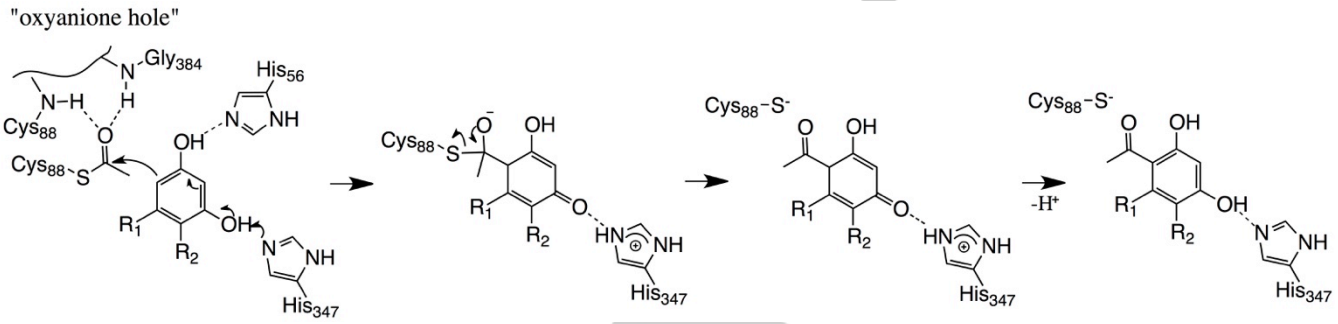


Figure 3. Proposed mechanism of the acyl-transfer.

Conclusions

The presented structural characterization of the multi-component acytransferase *Pp*ATase reveals a close interaction of all three enzymes PhlA, PhlB and PhlC. However, the C-C bond formation without the utilization of CoA is performed only by the PhlC subunit. This structural information provides a basis for developing libraries of catalysts tailored for specific chemical substrates. The extension of the substrate scope will be achieved using established protein engineering techniques.

Experimental Section

Expression and purification

The CoA-independent acyltransferase from *Pseudomonas protegens* DSM 19095 (*Pp*ATase) was expressed in *E. coli* as described previously[16b] from a plasmid containing codon optimized ORFs coding for the three subunits of the enzyme: PhlA, PhlB and PhlC. The nucleotide sequence of the expression plasmid is available from GenBank using the accession number KY173355. Purification of the enzyme was achieved using size-exclusion-chromatography as described previously[16b].

Crystallization and soaking

Screening for crystallization conditions was performed with an Oryx8 crystallization robot (Douglas Instruments). Initial trials were set up employing the sitting drop vapor-diffusion method in 96-well plates using Index HT (Hampton Research), JCSG+ and Morpheus (Molecular Dimensions) screens. A stock solution of *Pp*ATase (12 mg mL^{-1}, in 50mM potassium phosphate buffer pH 7.5) was used for all crystallization experiments. Both in initial screens and subsequent optimizations, drops of 1 µL were set with a 1:1 ratio of protein and precipitant solution. The crystallization plates were incubated at 289 K.

Crystal clusters of *Pp*ATase were readily obtained in several conditions. Microseed-matrix-screening experiments[31] were set up using initial crystals obtained in conditions Index #2 and #81 as seeding stock (1:1000 dilution, 0.1 µL added to the drop). Using this technique, single crystals were obtained in conditions Index #91 (0.15 M DL-malic acid pH 7.0, 20% w/v polyethylene glycol 3,350) and #55 (0.05 M magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 30% v/v polyethylene glycol monomethyl ether 550) after about one week. Soaking experiments were performed by adding small amounts of pure monoacetylphloroglucinol (MAPG) dissolved in DMSO directly to the crystallization drop with a cryo-loop. Soaking times varied between 20 to 60 sec. Untreated and soaked crystals were flash-cooled in liquid nitrogen using 15% v/v glycerol as cryoprotectant. Numerous crystals from different crystallization conditions were tested for diffraction.

Data collection, processing, structure determination and analysis

Data were collected at 100 K on synchrotron beamlines ID23-1 and ID30B (ESRF, Grenoble, France)[32] from an untreated crystal (hexagonal, space group $P6_122$) and a crystal soaked with MAPG (orthorhombic, space group $P2_12_12_1$) to crystallographic resolutions of approximately 2.8 and 3.4 Å, respectively. Diffraction data were processed and scaled using the XDS package[33]. Initial automated molecular replacement attempts using Balbes[34] and the better resolved hexagonal dataset indicated the dimer of 3-hydroxy-3-methylglutaryl-coenzyme A synthase

from *Staphylococcus aureus* (PDB-entry: 1tvz, 23% sequence identity)[35] as a suitable search template for PhlA. A homology model of PhlC was generated using the Phyre2 server[36] with the structure of SCP2 thiolase from *Leishmania mexicana* (PDB-entry: 3zbg, 27% sequence identity)[28] as template. Molecular replacement was continued within the CCP4 suite[37] using the program Phaser[38]. Fixing the previously positioned 1tvz-dimer (template for PhlA), two copies of the PhlC homology model could be placed in the asymmetric unit. Density modification based on phases from this partial model using the program Resolve[39] yielded well-defined electron density for the whole *Pp*ATase-complex, including density for two missing PhlB molecules. Manual rebuilding of PhlA and PhlC was continued using the program Coot[40] and the improved model was then submitted to automated rebuilding using the program Buccaneer[41]. The resulting model containing two copies of PhlA, PhlB and PhlC each was completed manually in Coot and refined using the PHENIX software suite[42].

The structure of the *Pp*ATase soaked with MAPG was solved by molecular replacement using a part of the previously determined hexagonal structure (one copy of PhlA, PhlB and PhlC each) as search template. Structure solution resulted in eight copies of this trimeric arrangement in the asymmetric unit of the orthorhombic unit cell. Structure refinement continued in the same manner as described above using the programs Coot and PHENIX. Clear difference electron density was observed in all eight chains of PhlC in the vicinity of residue Cys88. In four of those chains we interpreted this density as a molecule of phloroglucinol. Additional density around Cys88 was interpreted as an acetyl group covalently attached to the Sγ-atom of this amino acid.

Table 2. Crystal structure of *Pp*ATase, data collection and refinement statistics

	*Pp*ATase hexagonal[a]	*Pp*ATase orthorhombic (soaked)[a]
Wavelength	0.8726	0.95
Resolution range	48.74 - 2.83 (2.94 - 2.83)	49.72 - 3.44 (3.56 - 3.44)
Space group	*P* 6_1 2 2	*P* 2_1 2_1 2_1
Unit cell dimensions	222.63 222.63 237.10 90 90 120	104.79 229.78 311.13 90 90 90
Total reflections	695209 (67477)	602513 (51680)
Unique reflections	82036 (7914)	99724 (9200)
Multiplicity	8.5 (8.5)	6.0 (5.6)
Completeness (%)	1.00 (0.98)	0.99 (0.93)
Mean I/σ(I)	7.87 (2.23)	8.10 (2.57)
Wilson B-factor	36.41	53.96
R-merge	0.255 (0.891)	0.268 (0.788)
R-meas	0.271 (0.948)	0.294 (0.868)
$CC_{1/2}$	0.98 (0.67)	0.97 (0.81)
CC*	0.99 (0.90)	0.99 (0.95)
Reflections used in refinement	82033 (7914)	99695 (9198)
Reflections used for R-free	4102 (395)	4986 (460)
R-work	0.163 (0.251)	0.166 (0.219)
R-free	0.204 (0.316)	0.219 (0.292)
CC(work)	0.96 (0.83)	0.95 (0.90)
CC(free)	0.94 (0.74)	0.91 (0.82)
Number of non-H atoms	14084	54563
Macromolecules	13594	54398
Ligands	20	116
Protein residues	1800	7196
RMS(bonds)	0.011	0.005
RMS(angles)	0.91	0.72
Ramachandran favored (%)	96	97
Ramachandran allowed (%)	3.7	3.1
Ramachandran outliers (%)	0.11	0.18
Rotamer outliers (%)	3.2	1.5
Clashscore	4.29	10.12
Average B-factor	21.32	43.64
Macromolecules	21.24	43.63
Ligands	25.53	52.78
Solvent	23.51	30.28
PDB code	5M3K	5MG5

[a]Statistics for the highest-resolution shell are shown in parentheses.

For both structures, validation was performed using the program MolProbity[43]. Data collection and refinement statistics are summarized in Table 2.

All structure-related figures were generated using PyMOL (http://www.pymol.org). Structures were superimposed using the SSM Superposition tool[21] as implemented in Coot. Cavities were identified in the final structures using the LIGSITE algorithm[44] as implemented in the CaSoX plugin for PyMOL. The analysis of the hydrophobicity of these cavities utilized the corresponding function of the program VASCo[45].

Site directed mutagenesis and activity measurements

Variants of *Pp*ATase were prepared in order to investigate the role of selected amino acid residues in the enzymatic reaction. Gene mutations were introduced with the QuickChange II site-directed mutagenesis kit (Agilent Genomics) according to the standard procedure provided by the supplier without modifications. All primer sequences and plasmids used in this study are collated in Supporting Information Table 1. The following variants carrying a single amino acid exchange within *phlC* were generated: H56A, H56S, N87A, C88A, C88S, H144A, H144S, W211A, W211F, Y298A, Y298V, Y298F, H347F, S349A and D352V. All variants were expressed as described for the wildtype enzyme[16b] and were used as cell free extracts in the activity measurements (Table 1).

The activity of wildtype *Pp*ATase and of its variants was tested using the acetylation of benzene-1,3-diol (resorcinol) under the following conditions: Resorcinol (10 mM final concentration in the reaction mixture) was dissolved in potassium phosphate buffer (50 mM, pH 7.5) and preheated to 35 °C for 10 minutes. Cell-free extracts of the recombinant enzymes (30 mg protein) were subsequently added to the preheated mixture. The bioacetylation was started by addition of imidazole (100 mM, added from a 1 M stock solution prepared in the reaction buffer causing the pH to increase to 8.30) followed by the addition of isopropenyl

acetate (IPEA, 100 mM). To ensure proper suspension of the donor in the mixture, the vessel was manually shaken thoroughly right after starting the reaction. The reaction mixture (1 mL total volume) was horizontally shaken for 1.5 h at 35 °C and 750 rpm in an orbital shaker. Reactions were aborted by addition of acetonitrile (1 mL). The precipitated protein was removed by centrifugation (10 min, 14,000 rpm) and the supernatant (900 µL) was transferred to an Eppendorf tube and left standing for another 40 minutes. Any residual precipitated protein was once again removed by centrifugation and the supernatant was directly subjected to HPLC for determination of conversions. The relative amounts of resorcinol, the C-acetylation product 4-acetyl-resorcinol and the O-acetylation product 3-hydroxyphenyl acetate were determined by HPLC according to standard curves with authentic samples. Each reaction was performed as a duplicate.

Specific activities were measured on a Thermo Scientific Genesys 10 UV Scanning UV/Vis spectrophotometer according to a modified procedure from literature[13a]. When following the disproportionation of MAPG into DAPG and PG spectrophotometrically, an increase of absorption is recorded due to the formation of DAPG (ε = 20 mM^{-1} cm^{-1}, λ = 370 nm). One unit of activity was defined as the µmol of product formed by an enzyme in 1 min per 1 milligram of protein under the following conditions: potassium phosphate buffer (960 µL, 100 mM, pH 7.5) and MAPG (1.2 µmol, 30 µL of a 40 mM stock solution prepared in DMSO) were added to a cuvette and preheated to 35 °C. The reaction (1 mL total volume, 3 vol% DMSO) was started by the addition of the enzyme-containing cell-free extract (10 µL). The reaction was followed for 1 minute. All reactions were performed as a duplicate. The protein concentration was measured using Bradford reagent [(ε = 0.083 mL mg^{-1} cm^{-1}, λ = 595 nm)] and specific activities were determined as units per mg protein.

Acknowledgements

This study was financed by the Austrian FFG, BMWFJ, BMVIT, SFG, Standortagentur Tirol and ZIT through the Austrian FFG-COMET- Funding Program. COST Action CM1303 “Systems Biocatalysis” is acknowledged. Additional support was provided by the Austrian Science Fund (FWF) through a Lise-Meitner fellowship (M2172-B21) to AZD and through project P29432 to TPK. We acknowledge the ESRF (ID23-1 and ID30B, Grenoble, France) and DESY (P11, PETRAIII, Hamburg, Germany) for provision of synchrotron-radiation facilities and support during data collection.

Keywords: acyltransferase • crystal structure • Friedel-Crafts acylation • multi-component enzyme

[1] a) Y. Liu, L. Gao, L. Liu, Q. Yang, Z. Lu, Z. Nie, Y. Wang, T. Xia, *J. Biol. Chem.* **2012**, *287*, 44406-44417; b) R. Niemetz, G. G. Gross, *Phytochemistry* **2001**, *58*, 657-661.

[2] D. M. Byers, H. Gong, *Biochem. Cell Biol.* **2007**, *85*, 649-662.

[3] L. Hoffmann, S. Maury, F. Martz, P. Geoffroy, M. Legrand, *J. Biol. Chem.* **2003**, *278*, 95-103.

[4] M. C. Hunt, M. I. Siponen, S. E. H. Alexson, *BBA - Mol. Bas. Dis.* **2012**, *1822*, 1397-1410.

[5] W. Zhang, J. Huffman, S. Li, Y. Shen, L. Du, *BMC Biotechnol.* **2017**, *17*.

[6] L. Favrot, J. S. Blanchard, O. Vergnolle, *Biochemistry* **2016**, *55*, 989-1002.

[7] S. R. Partridge, R. M. Hall, *J. Clin. Microbiol.* **2005**, *43*, 4298-4300.

[8] D. M. Stalker, K. E. McBride, L. D. Malyj, *Science* **1988**, *242*, 419-423.

[9] a) L. Paoletti, Y. J. Lu, G. E. Schujman, D. De Mendoza, C. O. Rock, *J. Bacteriol.* **2007**, *189*, 5816-5824; b) Y. M. Zhang, C. O. Rock, *J. Lipid Res.* **2008**, *49*, 1867-1874.

[10] R. Kalscheuer, A. Steinbüchel, *J. Biol. Chem.* **2003**, *278*, 8075-8082.

[11] Y. Q. Cheng, G. L. Tang, B. Shen, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3149-3154.

[12] A. Bera, S. Herbert, A. Jakob, W. Vollmer, F. Götz, *Mol. Microbiol.* **2005**, *55*, 778-787.

[13] a) A. Hayashi, H. Saitou, T. Mori, I. Matano, H. Sugisaki, K. Maruyama, *Biosci. Biotechnol. Biochem.* **2012**, *76*, 559-566; b) F. Yang, Y. Cao, *Appl. Microbiol. Biotechnol.* **2012**, *93*, 487-495.

[14] a) C. Keel, D. M. Weller, A. Natsch, G. Defago, R. J. Cook, L. S. Thomashow, *Appl. Environ. Microbiol.* **1996**, *62*, 552-563; b) J. A. Moynihan, J. P. Morrissey, E. R. Coppoolse, W. J. Stiekema, F. O'Gara, E. F. Boyd, *Appl. Environ. Microbiol.* **2009**, *75*, 2122-2131.

[15] a) Y. Cao, X. Jiang, R. Zhang, M. Xian, *Appl. Microbiol. Biotechnol.* **2011**, *91*, 1545-1552; b) Y. Cao, M. Xian, *Biotechnol. Lett* **2011**, *33*, 1853-1858; c) G. Rao, J. K. Lee, H. Zhao, *Appl. Microbiol. Biotechnol.* **2013**, *97*, 5861-5867; d) F. Yang, Y. Cao, *Appl. Microbiol. Biotechnol.* **2012**, *93*, 487-495; e) W. Zha, S. B. Rubin-Pitel, H. Zhao, *J. Biol. Chem.* **2006**, *281*, 32036-32047.

[16] a) J. Achkar, M. Xian, H. Zhao, J. W. Frost, *J. Am. Chem. Soc.* **2005**, *127*, 5332-5333; b) N. G. Schmidt, T. Pavkov-Keller, N. Richter, B. Wiltschi, K. Gruber, W. Kroutil, *Angew. Chem. Int. Ed.* **2017**, *56*, 7615-7619.

[17] N. G. Schmidt, W. Kroutil, *Eur. J. Org. Chem.* **2017**, *2017*, 5865-5871.

[18] A. Zadło-Dobrowolska, N. G. Schmidt, W. Kroutil, *Chem. Commun.* **2018**, *54*, 3387-3390.

[19] E. Krissinel, K. Henrick, *J. Mol. Biol.* **2007**, *372*, 774-797.

[20] L. Holm, S. Kaariainen, P. Rosenstrom, A. Schenkel, *Bioinformatics* **2008**, *24*, 2780-2781.

[21] E. Krissinel, K. Henrick, *Acta Cryst. D* **2004**, *60*, 2256-2268.

[22] T. Bock, J. Kasten, R. Müller, W. Blankenfeldt, *ChemBioChem* **2016**, 1257-1262.

[23] X. Y. Qiu, C. A. Janson, W. W. Smith, M. Head, J. Lonsdale, A. K. Konstantinidis, *J. Mol. Biol.* **2001**, *307*, 341-356.

[24] M. G. Bangera, L. S. Thomashow, *J. Bacteriol.* **1999**, *181*, 3155-3163.

[25] S. S. Krishna, L. Aravind, C. Bakolitsa, J. Caruthers, D. Carlton, M. D. Miller, P. Abdubek, T. Astakhova, H. L. Axelrod, H. J. Chiu, T. Clayton, M. C. Deller, L. A. Duan, J. Feuerhelm, J. C. Grant, G. W. Han, L. Jaroszewski, K. K. Jin, H. E. Klock, M. W. Knuth, A. Kumar, D. Marciano, D. McMullan, A. T. Morse, E. Nigoghossian, L. Okach, R. Reyes, C. L. Rife, H. van den Bedem, D. Weekes, Q. P. Xu, K. O. Hodgson, J. Wooley, M. A. Elsliger, A. M. Deacon, A. Godzik, S. A. Lesley, I. A. Wilson, *Acta Cryst. F* **2010**, *66*, 1160-1166.

[26] R. L. Flynn, L. Zou, *Crit. Rev. Biochem. Mol. Biol.* **2010**, *45*, 266-275.

[27] A. M. Haapalainen, G. Meriläinen, R. K. Wierenga, *Trends Biochem. Sci* **2006**, *31*, 64-71.

[28] R. K. Harijan, T. R. Kiema, M. P. Karjalainen, N. Janardan, M. R. Murthy, M. S. Weiss, P. A. Michels, R. K. Wierenga, *Biochem. J.* **2013**, *455*, 119-130.

[29] B. Vögeli, S. Engilberge, E. Girard, F. Riobé, O. Maury, T. J. Erb, S. Shima, T. Wagner, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 3380-3385.

[30] L. David, E. Cheah, M. Cygler, B. Dijkstra, F. Frolow, M. Sybille, M. Harel, S. James Remington, I. Silman, J. Schrag, L. Joel, H. G. V. Koen, A. Goldman, *Protein Eng. Des. Sel.* **1992**, *5*, 197-211.

[31] A. D'Arcy, F. Villard, M. Marsh, *Acta Cryst. D* **2007**, *63*, 550-554.

[32] P. Theveneau, R. Baker, R. Barrett, A. Beteva, M. W. Bowler, P. Carpentier, H. Caserotto, D. de Sanctis, F. Dobias, D. Flot, M. Guijarro, T. Giraud, M. Lentini, G. A. Leonard, M. Mattenet, A. A. McCarthy, S. M. McSweeney, C. Morawe, M. Nanao, D. Nurizzo, S. Ohlsson, P. Pernot, A. N. Popov, A. Round, A. Royant, W. Schmid, A. Snigirev, J. Surr, C. Mueller-Dieckmann, *11th International Conference on Synchrotron Radiation Instrumentation (Sri 2012)* **2013**, *425*.

[33] W. Kabsch, *Acta Cryst. D* **2010**, *66*, 125-132.

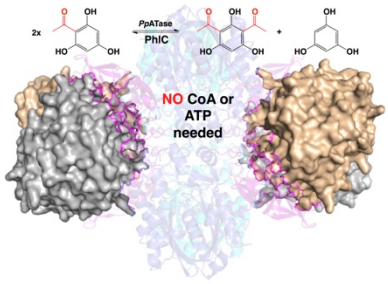
[34] F. Long, A. A. Vagin, P. Young, G. N. Murshudov, *Acta Cryst. D* **2008**, *64*, 125-132.
[35] N. Campobasso, M. Patel, I. E. Wilding, H. Kallender, M. Rosenberg, M. N. Gwynn, *J. Biol. Chem.* **2004**, *279*, 44883-44888.
[36] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. Sternberg, *Nat. Protoc.* **2015**, *10*, 845-858.
[37] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta Cryst. D* **2011**, *67*, 235-242.
[38] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, 658-674.
[39] T. C. Terwilliger, *Acta Cryst. D* **2000**, *56*, 965-972.
[40] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Cryst. D* **2010**, *66*, 486-501.
[41] K. Cowtan, *Acta Cryst. D* **2006**, *62*, 1002-1011.
[42] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Cryst. D* **2010**, *66*, 213-221.
[43] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Cryst. D* **2010**, *66*, 12-21.
[44] M. Hendlich, F. Rippmann, G. Barnickel, *J. Mol. Graph. Model.* **1997**, *15*, 359-363.
[45] G. Steinkellner, R. Rader, G. G. Thallinger, C. Kratky, K. Gruber, *BMC Bioinformatics* **2009**, *10*, 1-11.
[46] L. Hedstrom, *Chem. Rev.* **2002**, *102*, 4501-4523.

Entry for the Table of Contents

Layout 1:

FULL PAPER

Friedel-Crafts acylase: This multimeric enzyme consists of three subunits – PhlA, PhlB and PhlC which are arranged in a $Phl(A_2C_2)_2B_4$ composition. The subunit PhlC catalyzes a Friedel-Crafts C-acylation of phenolic substrates in aqueous solution without the need of CoA-activated reagents.



Tea Pavkov-Keller, [a,b] *Nina G. Schmidt,* [a,c] *Anna Żądło-Dobrowolska,* [c] *Wolfgang Kroutil,* *[a,c,d] *and Karl Gruber* *[a,b,d]

Page No. – Page No.

Structure and catalytic mechanism of a bacterial Friedel-Crafts acylase