Chapter 4:

Polycyclic Ketone Monooxygenase from the Thermophilic Fungus *Thermothelomyces thermophila*: A Structurally Distinct Biocatalyst for Bulky Substrates

Maximilian J.L.J. Fürst,‡a Simone Savino, ‡b Hanna M. Dudek, a J. Rúben Gómez Castellanos, b Cora Gutiérrez de Souza, c Stefano Rovida, b Marco W. Fraaije, *a and Andrea Mattevi. *b

‡These authors contributed equally.

*aMolecular Enzymology Group, University of Groningen, Nijenborgh 4, 9747AG, Groningen, The Netherlands

bDepartment of Biology and Biotechnology, University of Pavia, Via Ferrata 1, 27100, Pavia, Italy

cStratingh Institute for Chemistry, University of Groningen, Nijenborgh 4, 9747 AG, Groningen, The Netherlands

*These authors contributed equally.

*Corresponding authors

Published in:

*Journal of the American Chemical Society, 2017 (139) 627–630*

Reprinted with permission. ©2017 American Chemical Society.


**Abstract**

Regio- and stereoselective Baeyer-Villiger oxidations are difficult to achieve by classical chemical means, particularly when large, functionalized molecules are to be converted. Biocatalysis using flavin-containing Baeyer-Villiger monooxygenases (BVMOs) is a well-established tool to address these challenges, but known BVMOs have shortcomings either in stability or substrate selectivity. We characterized a novel BVMO from the thermophilic fungus *Thermothelomyces thermophila*, determined its three-dimensional structure, and demonstrated its use as a promising biocatalyst. This fungal enzyme displays excellent enantioselectivity, acts on various ketones, and is particularly active on polycyclic molecules. Most notably we observed that the enzyme can perform oxidations on both the A and D ring when converting steroids. These functional properties can be linked to unique structural features, which identify enzymes acting on bulky substrates as a distinct subgroup of the BVMO class.
Introduction

The Baeyer-Villiger oxidation is a widely-used reaction in organic synthesis to break carbon-carbon bonds through the insertion of a single oxygen atom adjacent to a carbonyl moiety, yielding esters or lactones from ketones. Since the inserted oxygen is usually derived from a peroxide or peracid, industrial scale processes raise environmental and safety concerns. Baeyer-Villiger monooxygenases (BVMOs) represent an attractive biocatalytic alternative, which uses molecular oxygen as oxidant and often displays superior chemo-, regio-, and enantioselectivity. The prototype BVMO, cyclohexanone monooxygenase from *Acinetobacter calcoaceticus* NCIMB 9871 (CHMO) was shown to oxidize hundreds of small aromatic, linear, and cyclic ketones with high enantioselectivity. Furthermore, one major hurdle for industrial application of BVMOs, the requirement of stoichiometric amounts of NADPH as reducing cofactor, has been successfully addressed. What remains problematic, however, is the poor solvent tolerance and thermostability of most BVMOs. Even though close to a hundred BVMOs have been recombinantly produced and characterized so far, only two enzymes stand out by their stability: phenylacetone monooxygenase from *Thermobifida fusca* (PAMO) and the recently discovered CHMO from *Thermocrispum municipale*. Yet, these biocatalysts have a restricted substrate scope, being mainly active on small aromatic or cyclic aliphatic ketones. From an application point of view, it would be attractive to have access to a robust BVMO acting on structurally demanding compounds, as it is the case for many pharmaceuticals.

Results and discussion

In an effort to exploit the genome of thermophiles for the discovery of new enzymes with interesting catalytic properties, we considered *Thermothelomyces thermophila* ATCC 42464. This genome-sequenced, thermophilic fungus efficiently degrades cellulose and other compounds derived from plant biomasses and is becoming an attractive organism for large-scale protein production. We identified a gene encoding for a protein containing the typical BVMO fingerprint. The gene is located among a cluster of hypothetical genes with unknown functions. A BLAST search in the UniProt database revealed ten characterized proteins with significant sequence identities (37-44%). Four are monooxygenases involved in the biosynthesis of meroterpenoids in *Aspergillus*. Three are BVMOs from *Streptomyces* taking part in the processing of the antibiotic pentalenolactone. One is a BVMO
forming the carbonate functionality in the mycotoxin cytochalasin. The remaining two are the biocatalytically characterized BVMOs from Dietzia sp. and S. coelicolor. A correlation between sequence similarity and substrate scope has been proposed for BVMOs, and indeed, a feature shared by these enzymes is the activity on bulky, polycyclic ketones. Consistently, a phylogenetic analysis assigns the T. thermophila protein to a cluster containing versatile enzymes active on large substrates, such as cyclododecanone monooxygenase from Rhodococcus ruber (CDMO) and cyclopentadecanone monooxygenase from Pseudomonas sp. HI-70 (CPDMO) (Figure S4.1). Collectively, these data led us to hypothesize that the T. thermophila protein is a BVMO that combines two attractive properties: (i) thermostability and (ii) activity on relatively large molecules.

An Escherichia coli codon-optimized copy of the gene was cloned in two translational fusion variants: a SUMO tagged construct and a cofactor-recycling phosphite dehydrogenase (PTDH) fusion. Both constructs yielded very high expression levels in E. coli and the proteins were easily obtained in high purity and yield using standard metal affinity chromatography by exploiting the N-terminal 6xHis tag (Figure S4.2A). The purified yellow protein exhibited a typical flavoprotein spectrum with absorbance peaks at 371 and 454 nm (Figure S4.2B). SDS treatment revealed non-covalently bound FAD as cofactor and the extinction coefficient $\varepsilon_{454}$ of the holoenzyme was estimated to be 10.2 mM$^{-1}$ cm$^{-1}$. Upon mixing the enzyme with 100 µM NADPH and the BVMO model substrate bicyclo[3.2.0]hept-2-en-6-one (13), rapid consumption of NADPH was detected (Figure S4.3). By monitoring NADPH consumption in the absence of substrate, we determined an uncoupling rate of only 0.02 s$^{-1}$ (0.9%). We also compared the catalytic behavior of the two recombinant protein variants and determined the activity of the enzyme fused to PTDH and without any tag (obtained by cleaving SUMO). As previously observed, the effect of the dehydrogenase fusion was very modest and, in fact, resulted in a slightly higher activity (Figure S4.3) and unchanged uncoupling. The resulting self-sufficient biocatalyst regenerates NADPH with turnovers in the order of several thousands, as long as phosphite is present in the reaction. Therefore, all subsequent experiments were performed with the PTDH-fused enzyme. The preliminary analysis was completed by determining basic enzyme characteristics for optimal catalytic conditions. The pH profile highlighted good stability in a range of 6.5 - 9 with a slight preference for moderately alkaline conditions (Figure 1A).
We next sought to probe the enzyme’s thermostability. We found the activity to be highest at 50 °C, where it was four times faster than at room temperature (Figure 1B). Using the ThermoFAD method, we further observed that both native as well as PTDH-fused enzyme have an apparent melting temperature ($T_m$) of 47 °C, the second highest reported for a BVMO after PAMO. This $T_m$ is sustained in a pH range of 7-9 and is increased a further 2 °C upon exposing the protein to 100 µM NADP+, consistent with a tight binding of this dinucleotide cofactor (Figure S4.4). In parallel, we evaluated the effect of organic solvents which was probed by adding DMSO, 1,4-dioxane, ethanol, acetone or 1-propanol. The resulting reductions in $T_m$ ranged from 1 - 5 °C to 2 - 8 °C with 5% and 10% solvent, respectively (Figure S4.4). The enzyme seemed to exhibit a better tolerance towards polar solvents, with DMSO having the least and 1-propanol the highest impact on $T_m$. Because DMSO can act as a substrate, and the effect of the remaining solvents was comparable, we chose dioxane as preferred cosolvent. We monitored the enzymatic activity as a function of time and observed a satisfactory half-life of approximately 24 hours when incubating the enzyme in buffer at 35 °C, or at room temperature with 10% dioxane (Figure 1C). These experiments indicated that the BVMO features a considerable degree of thermostability and only moderate sensitivity to organic solvents.

We next investigated the enzyme’s substrate scope and selectivity (fully listed in Table S4.1). The closely related CPDMO and CDMO are highly active on large cyclic ketones (Figure S4.1). Consistently, we found conversion of 10 and 11 (Figure S4.5). The fifteen membered cyclopentadecanone (8) was fully converted and kinetic parameters were determined: $k_{cat} = 0.1$ s$^{-1}$ and $K_M = 144$ µM (Figure S4.6). Of special interest, we observed efficient conversions of structurally demanding compounds, including several steroids. In particular, we tested stanolone (17), which has only one carbonyl group positioned at C3.
of the A-ring and observed complete conversion after 24 h (Figure S4.5). As a reference, we performed the conversion with CPDMO, for which the 3-keto-4-oxa-4a-homo isomer was reported as the only product.22 The MS spectrum of the resulting products with the expected mass shift of +16 Da were the same for both enzymes (Figure S4.5) and we confirmed the structure by NMR (Figure S4.7). With pregnenolone (18), a steroid harboring a carbonyl group only at the C17 sidechain, we found close to full conversion after 24 h (Figures 2 and S4.5). We again observed the +16 Da mass shift and the product's regiochemistry was determined via NMR to be the corresponding acetate (Figure S4.8). When the enzyme was presented with androstenedione or androstanediene (15 and 16), steroids with keto groups in both the A and D rings, it selectively oxidized the D-ring to yield the pharmaceutically relevant testo(lo)lactone (Figure S4.5). Again, complete conversion was observed after 24 hours.

Table 1. Substrate scope of PockeMO

<table>
<thead>
<tr>
<th>Category</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear ketones</td>
<td>1: R¹ = CH₃(CH₂)ₙ, R² = CH₃</td>
</tr>
<tr>
<td></td>
<td>2: R¹ = Cyclohexane, R² = H</td>
</tr>
<tr>
<td></td>
<td>3: R¹ = CH₂-Phe, R² = H</td>
</tr>
<tr>
<td>Cyclic ketones</td>
<td>4: n = 1</td>
</tr>
<tr>
<td></td>
<td>5: n = 3</td>
</tr>
<tr>
<td></td>
<td>6: n = 6</td>
</tr>
<tr>
<td>Substituted cyclic ketones</td>
<td>7: n = 7</td>
</tr>
<tr>
<td></td>
<td>8: n = 10</td>
</tr>
<tr>
<td>Bicyclic ketones</td>
<td>9: R¹ = Phe, R²,₃,₄,₅ = H</td>
</tr>
<tr>
<td></td>
<td>10: R² = Phe, R¹,₂,₄,₅ = H</td>
</tr>
<tr>
<td></td>
<td>11: R²,₄,₅ = CH₃, R¹,₂ = H</td>
</tr>
<tr>
<td>Steroids</td>
<td>12: (CH₃)₃CH₃</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>18</td>
</tr>
</tbody>
</table>

To our knowledge, this is the first report of an isolated BVMO that is able to oxidize the keto functionalities at these different positions. Steroid monooxygenase of *Rhodococcus rhodochrous* (STMO) can only convert steroidal carbonyls positioned on the C17 sidechain and not in the rings.23 A homolog from *Cylindrocarpon radicicola* in addition is able to lactonize the D ring.24 CPDMO accepts some steroids with the keto group on the 3 and 17 position, but is inactive on open chain ketones.22 The *T. thermophila* BVMO represents a biocatalyst combining these abilities (Table 1, red arrows) and we therefore named it *polycyclic ketone monooxygenase* (PockeMO).
We further explored the enzyme’s substrate profile using several substrates frequently tested with other BVMOs (Tables 1 and S4.1). For this purpose, we used a highly efficient substrate screening method, based on the conversion of a mixture of compounds and subsequent separation and analysis via GC-MS. We tested a mixture of 16 linear, aromatic, cyclic, and polycyclic ketones (Figure S4.5) and could immediately identify 13 conversion products. Of the unsubstituted cyclic ketones, PockeMO converted cyclohexanone to 100%, but the conversion of cyclo octanone was poor in comparison. The monosubstituted cyclic ketones 9, 10 and 12 were converted with yields close to 100%. Approximately half of the disubstituted cyclic ketone 11 was converted, whereas the unsaturated variant of this molecule (isophorone) was not accepted. The aromatic ketone 3 was fully converted and also the bridged, bicyclic 14 was converted with moderate yield (Table S4.1, Figure 2). Notably, although being only a poorly accepted substrate, the linear 4-octanone (1) was exclusively converted to the abnormal product propyl pentanoate (Figure S4.9). These experiments elegantly demonstrated the activity on an array of substrates of varying degree of bulkiness. At the same time, none of the compounds seemed to exhibit inhibitory effects on the enzyme. To put this further into perspective, we performed the conversion of a mix of 15 ketones with purified CHMO, CPDMO and PockeMO. Figure 2 shows the semi-quantitative comparison of the conversion efficiency of the three enzymes and clearly shows PockeMO’s excellent all-round performance as well as its particular superiority with larger compounds.

For BVMOs, enantio- and regioselectivity has frequently been studied by analyzing the conversion of racemic 13 into the respective lactones. A recent extensive study on BVMO activity in various fungi revealed diverse selectivities
and only a moderate phylogenetic correlation. We found that PockeMO (4 µM enzyme, 2 hours) completely converted 10 mM racemic 13 enantiodivergently, yielding two regioisomeric lactones, with ee values of 100% and 97%, respectively (Figure S4.10). Thus, the enzyme exhibited both high enantio- and regioselectivity as well as high activity towards this substrate. Kinetic analysis confirmed this notion as a $k_{cat}$ of 3.3 s$^{-1}$ was measured ($K_m = 0.4$ mM).

Given these valuable stability and catalytic properties as well as the fact that PockeMO belongs to the structurally uncharacterized CPDMO/CDMO subfamily of BVMOs (Figure S4.1), it was of interest to determine the enzyme's three-dimensional structure. After removal of the SUMO tag, the protein could be crystallized as a complex with FAD and NADP$^+$ and its structure solved at a resolution of 2.0 Å (Table S4.2). Most of the residues could be modelled, apart from loop 607-614.

**Figure 3.** Overall structure of PockeMO (PDB entry 5MQ6). The N-terminal extension (residues 1-73; dark blue) is specific to this BVMO enzyme subclass (Figure S4.1). FAD is yellow and NADP$^+$ green (nicotinamide ring is disordered).

PockeMO exhibits the typical BVMO domain organization with an FAD-domain (residues 1-229 and 479-655), an NADP-domain (residues 230-314 and 417-478), and a helical domain (residues 315-416) (Figure 3). The closest structural homologs are STMO$^{26}$ and PAMO (RMSD of 1.91 and 1.75 Å for 523 and 533 equivalent Ca atoms, respectively, with 30% sequence identity to both). The location of the FAD and NADP cofactors is consistent with that of other BVMOs and the flavin shows no significant deviation from planarity (Figure S4.11).
Polycyclic Ketone Monooxygenase

Figure 4. Characteristic structural features of PockeMO. (A) Residues 316-388 (bright blue) cover the active site and adopt a conformation not observed in other BVMO structures as exemplified by the superimposed equivalent residues of PAMO (PDB 2YLT, 234-298; orange; Figure S4.11). (B) PockeMO structure cut along the substrate entrance to outline the wide tunnel forming the active site. (C) PAMO active site features a narrower channel, ranging from the solvent exposed surface to the active site with MES bound.

Despite this clear sequence and structural homology with other BVMOs and the conservation of the overall fold, PockeMO features some striking structural peculiarities. Firstly, a 70-residues N-terminal extension forms a long α-helix followed by an elongated stretch that wraps around the enzyme in the vicinity of the FAD-domain (Figure 3). The interactions between the 44 N-terminal residues and the rest of the protein are extensive and establish 24 H-bonds, 5 salt bridges, and many hydrophobic contacts. Notably, this elongated N-terminus is a feature of all BVMOs in the phylogenetic group of enzymes acting on complex molecules (Figures S4.1 and S4.12A). Secondly, there is a structural element ranging from residue 316 to 388 whose characteristic conformation reshapes the active site. This region has low sequence similarity to the canonical BVMOs (Figure S4.12B), in which this portion is structurally conserved (as exemplified by PAMO, Figure 4A). In PockeMO, this stretch forms a characteristic conformation of alternating loops and α-helices. As a result, the substrate-binding site of PockeMO is a long and wide tunnel-like cavity, which starts from the loop 587-624 and ends at the re-face of FAD (Figures 4B-C). The flexible loop delimited by residues 606 and 615 serves as a gate-like mechanism, constituting a longer element when compared to the same topological loop in PAMO (residues 495-515). Modeling shows that bulky
molecules can readily be accommodated in the PockeMO’s substrate tunnel (Figure S4.13). Thus, the enzyme exhibits characteristic structural features, which are fully consistent with the notion that it belongs to a distinct subgroup of BVMO enzymes (Figure S4.1).

Conclusions

In summary, we identified a novel, robust and versatile enzyme performing Baeyer-Villiger oxidations. PockeMO combines thermostability and organic solvent tolerance with a broad substrate profile. Specifically, the enzyme accepts bulky and complex substrates and converts them with high efficiency, making it a promising candidate for application as an industrial biocatalyst. In addition, the elucidated crystal structure not only lays the basis for enzyme engineering, but can also be regarded as a prototype of an evolutionary and structurally distinct group of BVMOs. Careful further analysis should contribute to a better understanding of the still largely unknown mechanism by which substrate selectivity in BVMO is tuned. The structural and sequence features of PockeMO can also be exploited to identify new and diverse BVMOs evolved to process relatively large substrates.

Materials and methods

General methods

All chemical reagents were purchased from Sigma-Aldrich or TCI Europe, unless otherwise stated. Oligonucleotide primers were synthesized by Sigma-Aldrich or Eurofins. DNA sequencing was performed by GATC (Konstanz, Germany).

Phylogenetic analysis

Phylogenetic analysis with 79 BVMO sequences was performed using MEGA 7.0 to generate the alignment using the ClustalW algorithm with the BLOSUM protein weight matrix. Based on this alignment, a phylogenetic tree was generated using the Maximum-Likelihood (ML) method using these parameters: model: Jones-Taylor-Thornton; uniform rates, heuristic model: nearest-neighbor-interchange. The robustness was then tested with 500 bootstraps and the cut-off for displaying a common node in the final tree was 30%.

Cloning

A synthetic gene (GenScript - New Jersey, USA) was cloned in a pET-SUMO vector. For cloning into the pCRE vector, the gene was amplified via PCR with 0.2 µM forward primer GACTCGAGATCTGCTGCTGGTATGGCACCGTCTG and reverse primer GCGTTTCACTTCTGCTGTTGATGGCATGGCACCACGTCTG using the PfuUltra II Hotstart PCR Master Mix (Agilent Technologies) with 50 ng pET-SUMO-pockeMO vector DNA. The thermocycler conditions were 92 °C - 2' / [92 °C - 30'' / 55 °C - 30 '' / 72 °C - 4'] x 30 / 72 °C - 10'. The
PCR product and the vector were digested with XhoI and HindIII (New England Biolabs) overnight. The vector was then digested with alkaline phosphatase (Roche) overnight. The PCR product was purified and the backbone of the digested vector was cut from an agarose gel and also purified. The two fragments were ligated with T4 ligase (Promega) for 6 hours and 5 µl of the ligation were used to transform 100 µl of competent E. coli NEB 10β cells. Plasmid was prepared using a QIAprep Spin Miniprep Kit. In-frame insertion and sequence integrity was confirmed by sequencing (GATC-Biotech, Germany). Unless otherwise indicated, all enzymes were used according to manufacturer’s instructions. In-frame insertion and sequence integrity was confirmed by sequencing (GATC-Biotech, Germany).

Protein expression and purification

**PockeMO**

A baffled Erlenmeyer flask (200 mL – 5 L) was filled up to 10% of the volume with TB medium and inoculated with 5 mL of an overnight culture of E. coli. For expression of the SUMO fusion, BL21 (DE3) transformed with pET-His6-SUMO-pockeMO was used and the medium was supplemented with 50 µg/mL kanamycin. Cells were grown at 37 °C until an OD_{600} of ~1, and then expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Cells were shifted to 30 °C and allowed to grow 6-12 hours before harvest. For expression of the phosphite dehydrogenase fusion, chemocompetent E. coli NEB 10β cells transformed with pBAD-His6-PTDH-Linker-pockeMO (pCRE-pockeMO) were used to inoculate the main culture. 0.02% L-arabinose and 50 µg/mL ampicillin were added and the flask was incubated at 24 °C for 36 h with shaking.

Cells were harvested by centrifugation (6,000×g for 15 min at 4 °C, JA-10.5 rotor, Beckman Coulter) and frozen at −20 °C for later use. Cells were then disrupted by sonication and subsequently subject to centrifugation (15,000×g for 45 min at 4 °C, JA-17 rotor, Beckman Coulter). The cell free extract was filtered and loaded on 2 mL of Ni2+ Sepharose HP (GE Healthcare) pre-equilibrated with buffer and incubated for 1 h at 4 °C with rotating. Then, the column was washed with three column volumes of 50 mM Tris/HCl pH 7.5, followed by three column volumes of 50 mM Tris-HCl pH 7.5 containing 5 mM imidazole. The protein was eluted using 50 mM Tris-HCl pH 7.5 with 500 mM imidazole. Fractions containing yellow protein were applied on a pre-equilibrated Econo-Pac 10DG desalting column (Bio-Rad). The desalted sample was used directly or flash frozen in liquid nitrogen after optional addition of 10% glycerol (final concentration) and stored at −80 °C. Purity and integrity of the purified enzyme batch was confirmed by SDS-Gel analysis and recording of a UV-visible spectrum.

**CPDMO and CHMO**

Cyclopentadecanone monooxygenase from Pseudomonas sp. HI 70 (CPDMO) and cyclohexanone monooxygenase from Acinetobacter calcoaceticus NCIMB 9871 (CHMO) were expressed from an identical to above pBAD-His6-PTDH-Linker-BVMO plasmid, using the same conditions with 0.02% L-arabinose, 50 µg/mL ampicillin and incubation at 24 °C for 36 h with shaking. Purification was also performed exactly as described above for PockeMO.
**Determination of extinction coefficient**

Purified enzyme was diluted to a concentration of 5 - 10 µM and distributed to two reaction tubes. To one of the samples, sodium dodecyl sulfate was added to a final concentration of 0.1% (w/v). After 10 minutes the UV-visible spectrum of both samples was recorded. The concentration of holoenzyme originally present was then defined as equal to the concentration of free FAD determined in the SDS sample, by using an extinction coefficient of 11,300 M$^{-1}$ cm$^{-1}$. The extinction coefficient of the enzyme was then calculated for the wavelength of highest absorption using the Beer-Lambert law.

**NADPH consumption assay and kinetic parameters**

To determine enzymatic activity and kinetic parameters, reactions were followed spectrophotometrically by measuring the decrease in absorption at 340 nm. The reaction mixture (100 - 200 µl) typically contained 50 mM Tris-HCl pH 7.5, 2 µM isolated enzyme, 100 µM NADPH, and varying amounts of substrate, solubilized in 2.5 or 10% cosolvent, if required. All components except for the substrate were kept on ice until immediately before use and the reaction was started by adding NADPH. The mix was then quickly transferred to a 0.1 mL quartz cuvette and the absorbance at 340 nm was measured for at least 40 seconds. Kinetic parameters were obtained by fitting the obtained data to the Michaelis-Menten equation using the software GraphPad Prism 6. In every experiment, control reactions to monitor uncoupling rates were carried out by measuring NADPH consumption in the absence of substrate in buffer (without or with the co-sovlent, depending on the experiment).

**pH profile**

Activities were measured by first preparing samples of 1 mM bicyclo[3.2.0]hept-2-en-6-one in 100 mM Tris-HCl buffer of the desired pH. Then, 1 µM enzyme (final) and 100 µM NADPH (final) buffered in 10 mM Tris-HCl pH 7.5 were added from concentrated stock solutions in a volume ratio substrate:enzyme:NADPH 8:1:1. Following mixing, NADPH consumption rate was followed spectrophotometrically at 340 nm. The reactions were carried out at 25 °C.

**Temperature profile**

Activity at different temperatures was measured with samples of 1 mM bicyclo[3.2.0]hept-2-en-6-one in 50 mM Tris-HCl pH 7.5, mixed with 1 µM of enzyme (final) and 100 µM NADPH (final). Except for the substrate, before mixing, solutions were kept on ice and then the mix was transferred into a spectrophotometer cuvette. The spectrophotometer cell temperature was beforehand set to the desired value and the cuvette was allowed to heat up for 10 seconds, before starting the measurement at 340 nM to follow NADPH consumption.

**Tm determination via ThermoFAD**

To determine the apparent melting temperature ($T_m$) of PockeMO in different conditions, duplicate samples of 25 µl were prepared in a 96-well thin wall PCR plate.
A 50 mM Tris buffer with HCl-adjusted pH containing the desired additives was used. The samples contained 1 mg/mL purified enzyme. The plate was heated from 20 °C to 90 °C, increasing temperature by 0.5 °C every 10 seconds, using an RT-PCR machine (CFX96-Touch, Bio-Rad Laboratories) that measured fluorescence using a 450–490 excitation filter and a 515–530 nm emission filter. The melting point was defined as the temperature when the first derivative of the observed fluorescent signal showed a maximum.

**Temperature and solvent stability**

Activity measurements were performed in 50 mM Tris/HCl using 1 mM bicyclo[3.2.0]hept-2-en-6-one as the substrate. NADPH consumption was followed spectrophotometrically at 340 nm for 1 min. Purified enzyme was kept on ice until the first activity measurement and subsequently incubated without shaking at 25 or 35 °C in a theromixer (Eppendorf). Several measurements were performed for up to 24 h after incubation start.

**Bioconversions**

A typical conversion mix contained 50 mM Tris/HCl pH 9.0 buffer, 10 µM purified phosphite dehydrogenase-fused enzyme, 10 mM sodium phosphite and 100 µM NADPH. In a closed 20 mL glass vial, 1 mL of that mix was shaken at 30 °C for 24 h before analysis. For the large cyclic ketones and steroids the substrate stock solution concentration was 8 mM solubilized in 100% dioxane. Conversion of ketone mixes were prepared with a substrate mix stock solution in 100% dioxane that contained each of the ketones in a concentration of 4 mM (mix of 4-7 substrates) or 2 mM (mix of 15 substrates). The stock was then diluted 1:10 in the final reaction mix.

Analysis of isomeric bicyclo[3.2.0]hept-2-en-6-one lactones by chiral GC

Bioconversion was performed as described above, but for only two hours, with 4 µM enzyme and 10 mM racemic bicyclo[3.2.0]hept-2-en-6-one. 500 µl of the conversion mixes were extracted three times with 500 µl of tert-butyl methyl ether supplemented with 0.1% (v/v) mesitylene as an internal standard, dried over anhydrous sodium sulfate, and analyzed by chiral gas chromatography (GC) using a 7890A GC System (Agilent Technologies) and a CP Chiralsil Dex CB column (25 m x 0.25 mm x 0.25 µm, Agilent) chiral column. The temperature program of the column oven was from 40 °C to 130 °C at 10 °C/min, then 130 °C for 15 min; and finally from 130 °C to 40 °C at 10 °C/min. Configurations of the products were assigned by comparison of the GC retention times order with published data.

**Analysis of products by GC-MS**

600 µl of the conversion mixes were extracted three times with equal amounts of ethyl acetate, dried over anhydrous sodium sulfate and the solvent evaporated using a speed vac concentrator (Savant). Compounds were then resuspended in 300 µl ethyl acetate and analyzed via a GC-MS QP2010 ultra (Shimadzu) with electron ionization and quadrupole separation. The column used was a HP-1 (Agilent, 30 m x 0.32 mm x 0.25 µm). The programs for the GC can be inferred from the following table:
### Polycyclic Ketone Monooxygenase

<table>
<thead>
<tr>
<th>Compound(s)</th>
<th>Program</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyclododecanone, Cycloundecanone, Cyclopentadecanone</td>
<td>Injection temp.: 300 °C</td>
</tr>
<tr>
<td></td>
<td>Oven program: 100 °C; 5 °C/min until 200 °C; 200 °C, 2 mins.</td>
</tr>
<tr>
<td>Steroids</td>
<td>Injection temp.: 250 °C</td>
</tr>
<tr>
<td></td>
<td>Oven program: 190 °C, 2 mins; 12 °C/min until 325 °C; 325 °C, 3 mins.</td>
</tr>
<tr>
<td>Mix of ketones without steroids</td>
<td>Injection temp.: 250 °C</td>
</tr>
<tr>
<td></td>
<td>Oven program: 30 °C, 3 mins; 7.5 °C/min until 200 °C; 200 °C, 3 mins; 15 °C/min until 325 °C; 325 °C, 3 mins.</td>
</tr>
<tr>
<td>Mix of ketones with steroids</td>
<td>Injection temp.: 260 °C</td>
</tr>
<tr>
<td></td>
<td>Oven program: 30 °C, 5 mins; 5 °C/min until 70 °C; 70 °C, 5 mins; 5 °C/min until 130 °C; 130 °C, 5 mins; 15 °C/min until 325 °C; 325 °C, 1 min.</td>
</tr>
</tbody>
</table>

In all cases, 1 µl was injected into the GC and the split ratio was 5.0. The software to analyze chromatograms, MS spectra and to generate the figures was GCMSsolution Postrun Analysis 4.11 (Shimadzu). The library for the MS spectra was NIST11.

### NMR analysis

To produce sufficient amount for ¹H NMR, bioconversion was performed on a multi milliliter scale, with reaction mixes of 15 mL and 40 mL for stanolone and pregnenolone respectively. For stanolone, the standard concentration of 800 µM, and for the highly insoluble pregnenolone, a concentration of 500 µM was used in the reaction. Extraction was performed three times with ethyl acetate in a separation funnel, and the solvent was removed in a rotary evaporator. Both steroids afforded a white/light brown pellet which was dried over N2 overnight and then resuspended in 1 mL CDCl₃. NMR analysis was performed using a Varian Unity Plus (¹H: 400 MHz) spectrometer. Chemical shifts are denoted in δ-units (ppm) relative to the residual solvent peak (CDCl₃: 1H δ =7.26). The splitting parameters are named as follows: s = singlet, d = doublet, t = triplet, m = multiplet, dd = doublet of doublets.

### 3.15 Crystallography

Purified PockeMO used for crystallization experiments was concentrated to 12 mg mL⁻¹ of enzyme in 20 mM Tris-HCl buffer at pH 7.5. For initial crystallization screening, we tested the kits Classic 1 and 2 from Jena Bioscience, Midas 1 and 2 from Molecular Dimensions, and the pH clear and AmSO4 suites from Qiagen. An Oryx 8 crystallization robot (Douglas Instruments, UK) was used for robotic screening. Crystallization hits were optimized manually using the sitting-drop setup. After optimization, the best diffracting crystals were found in a condition containing 2.7 M ammonium sulfate, 100 mM MES sodium salt pH 6.5, and 5% v/v PEG400. Crystals grew from drops prepared 1:1 pockeMO (9 mg mL⁻¹) and reservoir at 20 °C in 48 h as bright yellow solids, shaped as cut diamonds with hexagonal base. These crystals were extremely reproducible and co-crystallization with NADP⁺ was not impairing or improving their formation. Crystals were harvested from the mother liquor using nylon cryoloops (Hampton...
research, USA) and flash-cooled in liquid nitrogen after a short soak in a solution containing 20% v/v PEG400, 5 mM NADP+, 3 M ammonium sulfate, 100 mM MES sodium salt pH 6.5. X-ray diffraction data used for structure determination and refinement were collected at PX beamline of the Swiss Light Source in Villigen, Switzerland (SLS) and at the ID23-EH1 beamline of the European Synchrotron Radiation Facility in Grenoble, France (ESRF). Data were processed and scaled using the CCP4 programs iMOSFLM for indexing and integration and AIMLESS for space group assignment and scaling. The space group symmetry together with final data-collection and processing statistics are listed in Table S1. The structures were solved with MOLREP, and the coordinates of PAMO (PDB: 2YLT) as search model. Manual building, addition of water molecules, and crystallographic refinement were performed with COOT, REFMAC5 and other programs of the CCP4 suite. Structure quality and validation were assessed using the wwPDB validation server. The final model coordinates were deposited in the Protein Data Bank with accession code 5MQ6. Figures were created with Chimera.

Acknowledgements

The research for this work has received funding from the European Union (EU) project ROBOX (grant agreement n° 635734) under EU’s Horizon 2020 Programme Research and Innovation actions H2020-LEIT BIO-2014-1.

References

7. Berka, RM; Grigoriev, IV; Otiilar, R; Salamov, A; Grimwood, J; Reid, I; Ishmael, N; John, T; Darmond, C; Moisan, MC; Henrissat, B; Coutinho, PM; Lombard, V; Natvig, DO; Lindquist, E; Schmutz, J; Lucas, S; Harris, P; Powlowski, J; Bellemare, A; Taylor, D; Butler, G; de Vries, RP; Allijn, IE; van den Brink, J; Ushinsky, S; Storms, R; Powell, AJ; Paulsen, IT; Elbourne, LD; Baker, SE; Magnuson, J; Laboissiere, S; Clutterbuck, AJ; Martinez, D; Wogulis, M; de Leon, AL; Rey, MW; Tsang, A. Comparative Genomic Analysis of the Thermophilic Biomass-Degrading Fungi Myceliophthora thermophila and Thielavia Terrestris. Nat. Biotechnol. 2011 (29) 922-7.


Matsuda, Y; Awakawa, T; Wakimoto, T; Abe, I. Spiro-Ring Formation Is Catalyzed by a Multifunctional Dioxygenase in Austinol Biosynthesis. *J. Am. Chem. Soc.* 2013 (135) 10962-5.

Seo, MJ; Zhu, D; Endo, S; Ikeda, H; Cane, DE. Genome Mining in *Streptomyces*: Elucidation of the Role of Baeyer–Villiger Monooxygenases and Non-Heme Iron-Dependent Dehydrogenase/Oxygenases in the Final Steps of the Biosynthesis of Pentalenolactone and Neopentalenolactone. *Biochemistry* 2011 (50) 1739-1754.

Jiang, J; Tetzlaff, CN; Takamatsu, S; Iwatsuki, M; Komatsu, M; Ikeda, H; Cane, DE. Genome Mining in *Streptomyces avermitilis*: A Biochemical Baeyer–Villiger Reaction and Discovery of a New Branch of the Pentalenolactone Family Tree. *Biochemistry* 2009 (48) 6431-6440.

Hu, Y; Dietrich, D; Xu, W; Patel, A; Thuss, JA; Wang, J; Yin, WB; Qiao, K; Houk, KN; Vederas, JC; Tang, Y. A Carbonate-Forming Baeyer–Villiger Monooxygenase. *Nat. Chem. Biol.* 2014 (10) 552-554.


Szolkowy, C; Eltis, LD; Bruce, NC; Grogan, G. Insights into Sequence–Activity Relationships Amongst Baeyer–Villiger Monooxygenases as Revealed by the Intragenomic Complement of Enzymes from *Rhodococcus jostii* Rha1. *ChemBioChem* 2009 (10) 1208-1217.


Torres Pazmiño, DE; Riebel, A; de Lange, J; Rudroff, F; Mihovilovic, MD; Fraaije, MW. Efficient Biooxidations Catalyzed by a New Generation of Self-Sufficient Baeyer-Villiger Monooxygenases. *ChemBioChem* 2009 (10) 2595-8.

Forneris, F; Orru, R; Bonivento, D; Chiarelli, LR; Mattevi, A. Thermofad, a Thermofluor-Adapted Flavin *ad hoc* Detection System for Protein Folding and Ligand Binding. *FEBS J.* 2009 (276) 2833-2840.


Leipold, F; Rudroff, F; Mihovilovic, MD; Bornscheuer, UT. The Steroid Monooxygenase from Rhodococcus rhodochrous; a Versatile Biocatalyst. *Tetrahedron: Asymmetry* 2013 (24) 1620-1624.


Butinar, L; Mohorčič, M; Deyris, V; Duquesne, K; Iacazio, G; Claeyx-Bruno, M; Friedrich, J; Alphand, V. Prevalence and Specificity of Baeyer–Villiger Monooxygenases in Fungi. *Phytochemistry* 2015 (117) 144-153.


29 Winn, MD; Ballard, CC; Cowtan, KD; Dodson, EJ; Emsley, P; Evans, PR; Keegan, RM; Krissinel, EB; Leslie, AG; McCoy, A. Overview of the Ccp4 Suite and Current Developments. *Acta Cryst. D* **2011** (67) 235-242.


31 Orru, R; Dudek, HM; Martinoli, C; Torres Pazmiño, DE; Royant, A; Weik, M; Fraaije, MW; Mattevi, A. Snapshots of Enzymatic Baeyer–Villiger Catalysis: Oxygen Activation and Intermediate Stabilization. *J. Biol. Chem.* **2011** (286) 29284-29291.


33 Murshudov, GN; Skubák, P; Lebedev, AA; Pannu, NS; Steiner, RA; Nicholls, RA; Winn, MD; Long, F; Vagin, AA. Refmac5 for the Refinement of Macromolecular Crystal Structures. *Acta Cryst. D* **2011** (67) 355-367.

34 Pettersen, EF; Goddard, TD; Huang, CC; Couch, GS; Greenblatt, DM; Meng, EC; Ferrin, TE. Ucsf Chimera—a Visualization System for Exploratory Research and Analysis. *Journal of computational chemistry* **2004** (25) 1605-1612.