BIOTECHNOLOGICALLY RELEVANT ENZYMES AND PROTEINS

Expanding the set of rhodococcal Baeyer–Villiger monooxygenases by high-throughput cloning, expression and substrate screening

A. Riebel · H. M. Dudek · G. de Gonzalo · P. Stepniak · L. Rychlewski · M. W. Fraaije

Abstract To expand the available set of Baeyer–Villiger monooxygenases (BVMOs), we have created expression constructs for producing 22 Type I BVMOs that are present in the genome of *Rhodococcus jostii* RHA1. Each BVMO has been probed with a large panel of potential substrates. Except for testing their substrate acceptance, also the enantioselectivity of some selected BVMOs was studied. The results provide insight into the biocatalytic potential of this collection of BVMOs and expand the biocatalytic repertoire known for BVMOs. This study also sheds light on the catalytic capacity of this large set of BVMOs that is present in this specific actinomycete. Furthermore, a comparative sequence analysis revealed a new BVMO-typifying sequence motif. This motif represents a useful tool for effective future genome mining efforts.

Keywords Biocatalysis · Baeyer–Villiger monooxygenase · Sulfoxidation · Enantioselectivity

A. Riebel · H. M. Dudek · G. de Gonzalo · M. W. Fraaije (✉)
Laboratory of Biochemistry, Groningen Biomolecular Sciences
and Biotechnology Institute, University of Groningen,
Nijenborgh 4,
9747 AG, Groningen, The Netherlands
e-mail: m.w.fraaije@rug.nl
URL: http://www.rug.nl/staff/m.w.fraaije/index

P. Stepniak · L. Rychlewski
Bioinfobank Institute,
Limanowskiego 24A,
60-744, Poznan, Poland

Introduction

In the last few years, interest in using Baeyer–Villiger monooxygenases (BVMOs) as biocatalysts has increased. These flavin-dependent enzymes are able to perform efficient chemo-, regio-, and/or enantioselective oxygenation reactions using stoichiometric amounts of O_2 as oxidant and NAD(P)H as electron donor. It has been shown that BVMOs can be used as biocatalysts for highly enantioselective Baeyer–Villiger oxidations and sulfoxidations (de Gonzalo et al. 2010). In addition, these oxidative biocatalysts have also been shown to catalyze oxygenation of other heteroatoms like nitrogen, phosphor, selenium and boron, as well as epoxidation reactions (Renz and Meunier 1999; Mihovilovic et al. 2004). Such selective oxygenation reactions are difficult to perform selectively with chemical catalysts rendering BVMOs a powerful tool in organic chemistry.

Taking into account the urgent need for environmental friendly processes, the use of biocatalysts is highly favourable due to the mild reaction conditions possible by employing enzymes such as BVMOs. However, there are still some drawbacks concerning the applicability of these enzymes. One of them is the fact that BVMOs are coenzyme dependent and therefore need an NAD(P)H regeneration system. Several different strategies have been applied to solve this problem, like the use of whole cells or of a regeneration enzyme partner (Hollmann et al. 2007; Torres Pazmiño et al. 2009). Another complicating factor is the limited availability of suitable BVMOs. Of the several sequence-related BVMO subclasses, only BVMOs belonging to the Type I class have been produced in a recombinant manner (Kamerbeek et al. 2003; Torres Pazmiño et al. 2010). With the known Type I BVMOs, only a limited set of reactions can be catalyzed. Furthermore, some of these BVMOs that are of interest are

rather unstable. This is already evident from the most studied BVMO: cyclohexanone monooxygenase (CHMO). This BVMO from *Acinetobacter* has been shown to act on a variety of compounds (Stewart 1998; Mihovilovic et al. 2002), but it is very unstable (CHMO: $t_{1/2}=24$ h at 25 °C) (Zambianchi et al. 2002). In contrast, phenylacetone monooxygenase (PAMO) from *Thermobifida fusca* is a very robust BVMO ($t_{1/2}=24$ h at 52 °C) (Fraaije et al. 2005; Secundo et al. 2011). Nevertheless, its substrate scope is rather limited to small aromatic substrates.

The above-mentioned current limitations in BVMO-based biocatalysis has triggered enzyme discovery and engineering studies to generate a larger collection of BVMOs that show a good operational stability while covering a wide variety of compounds. One approach to achieve this is to engineer known BVMOs in order to create supplementing activities and selectivities. For this purpose, several enzyme engineering studies have been reported in recent years (Torres Pazmiño et al. 2010). For example, PAMO has been subjected to structure-based mutagenesis and this has resulted in several mutants that display a somewhat altered substrate range and/or enantioselectivity when compared with the wild-type enzyme (Bocola et al. 2005; Torres Pazmiño et al. 2007; Wu et al. 2010; Dudek et al. 2011). However, these enzyme engineering studies have also shown that it is difficult to introduce substrate acceptance profiles that are significantly different from the parent enzyme.

To increase the number of available and distinct BVMOs, also the natural diversity can be tapped. With the BVMO-specific sequence motif it has become easy to survey sequenced genomes for BVMO-encoding genes (Fraaije 2002). Such identified genes can subsequently be cloned and the corresponding proteins overexpressed for biocatalytic evaluation studies. Genome mining has already delivered a number of novel BVMOs (Torres Pazmiño et al. 2010). Even though BVMOs are widespread in bacteria and fungi, the average distribution is only around one BVMO-encoding gene per microbial genome (de Gonzalo et al. 2010). In this respect, the bacterium *Rhodococcus jostii* RHA1 is a peculiar and promising exception: its predicted proteome contains an unusually high variety in oxidative enzymes, including >20 putative BVMOs (McLeod et al. 2006). Szolkowy et al. (2009) already performed an initial study on the BVMO repertoire of this organism and identified several interesting activities but failed in expressing all BVMOs.

In this paper, we report on the exploration of 22 BVMOs derived from *R. jostii* RHA1. By optimizing expression vectors and expression conditions, all BVMOs were obtained in a soluble form. Using a newly developed method to determine the amount of expressed and active BVMO in a cell extract (CE) in combination with a novel generic activity screening assay, the substrate profiles for a large set of substrates for all 22 BVMOs have been determined.

Materials and methods

Reagents and enzymes

Oligonucleotide primers were purchased from Sigma, dNTPs and In-Fusion™ 2.0 CF Dry-Down PCR Cloning Kit from Clontech, Phusion polymerase from Finnzymes. All other chemicals were obtained from Acros Organics, ABCR, Sigma-Aldrich, TCI Europe, and Roche Diagnostics GmbH.

Bacterial strains and plasmids

Escherichia coli TOP10 from Invitrogen was used as a host for DNA manipulations and protein expression. Two expression vectors have been used: (1) a modified pBAD vector (pBAD*N*) in which the *Nde*I site was replaced by the original *Nco*I site (Kamerbeek et al. 2004), and (2) the recently engineered pCRE2 vector which harbors a codon-optimized gene encoding a 18x mutant of phosphite dehydrogenase (PTDH) with an N-terminal His-tag (Torres Pazmiño et al. 2009).

Sequence analysis, cloning and gene expression

The NCBI server was used for DNA sequence retrieval and BLAST searches (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The preliminary alignment of BVMOs was prepared using ClustalW multiple alignment software (Thompson et al. 1994). Next, secondary structure information was extracted from the recently elucidated PAMO structure containing NADP$^+$ and a bound substrate analogue (PDB:2YLT, chain A) (Orru et al. 2011) and used for adjustments of the alignment. The alignments were also validated using GRDB (Pas et al. 2011). For calculation of the Neighbor-Joining trees and bootstrapping (1,500 iterations), we applied the MEGA5 software package (Tamura et al. 2011). The sequence-homology derived branching diagrams were visualized by using the iToL online tool (Letunic and Bork 2011).

Target genes were amplified using the genomic DNA of *R. jostii* RHA1 (originally isolated by Seto et al. in 1995) as template. The genome of *R. jostii* RHA1 has been fully sequenced (McLeod et al. 2006), and information can be found at www.rhodococcus.ca. The genes were cloned into pBAD or pCRE2 expression vectors using the In-Fusion PCR Cloning kit from Clontech, following the recommendations of the manufacturer. Gene #8 was cured by inserting a nucleotide (C) at position 621. In this way, the stop codon of the ORF RHA1_ro08999, encoding a putative N-terminal

domain of a Type I BVMO, was changed into a codon for leucine. This amino acid was found most frequently in BVMOs at the respective position. By the subsequent frame shift ORF RHA1_ro08999 and RHA1_ro08998 are fused at DNA level encoding a full length Type I BVMO. The insertion was achieved by QuikChange® site-directed mutagenesis using the following primers: forward, 5′-caccggtttcgggttcct gaactcccccaa-3′, and reverse, 5′-ttgggggagttcaggaaccc gaaaccggtg-3′.

Expression was tested using 24 deep square well microtiter plates in the sandwich cover system from Enzyscreen®, shaking at 200 rpm. Cell cultures of 2.5 ml in Luria–Bertani medium supplemented with 50 μg ml^{-1} ampicillin were grown at four different temperatures (17 °C for 48 h, 24 °C for 32 h, 30 °C and 37 °C for 16 h) with three different inducer (arabinose) concentrations (0.002%, 0.02%, 0.2%, 0% as control) each. CEs were obtained using DNase/lysozyme in combination with freezing in liquid nitrogen and thawing at 30 °C. CEs as well as cleared cell extracts (CCEs) were analyzed via SDS-PAGE for (soluble) expression of the respective BVMOs.

BVMO concentration determination in cleared cell extract

For the substrate screening assays, cell cultures of 500–800 ml Luria–Bertani medium supplemented with 50 μg ml^{-1} ampicillin and the corresponding amount of arabinose were grown. CE was prepared using sonication with subsequent centrifugation to remove the cell debris and obtain a CCE. An absorption spectrum of the CCE was measured from 700 to 300 nm (oxidized FAD), after which 500 μM NADPH was added. The sample was incubated at 25 °C for 10 min for full flavin reduction, after which again a spectrum was measured. Subtracting the two collected spectra revealed the amount of NADPH-reactive flavin in the CCE. To minimize interfering absorbance due to protein aggregates of *E. coli* proteins, the absorption at 440 nm was corrected by subtracting the absorption value at 700 nm. The amount of flavin-containing BVMO in the CCE was calculated using the difference extinction coefficient derived from PAMO: $\varepsilon_{\text{difference}}= 8.8\ \text{mM}^{-1}\ \text{cm}^{-1}$.

Substrate screening

For activity screening, an indirect activity assay was used. The substrates were dissolved in DMSO or dioxane. The solvent (DMSO or dioxane) concentration was kept at 5% in all tested reactions. For the activity assay, 100 μl 50 mM Tris–HCl (pH 7.5), containing 1 or 5 mM substrate, 100 μM NADPH, 10 μM PTDH, 10 mM sodium phosphite and 1 μM BVMO as CCE was incubated at 25 °C for 2 h. Twenty microliters of the reaction mixture was subsequently mixed with 200 μl molybdate reagent (10 mM $(NH_4)_6Mo_7O_{24}$ $4H_2O$, 100 mM

Zn acetate, pH 5, adjusted with HCl) and 50 μl 10% ascorbic acid (pH 5, adjusted with 40% NaOH). During incubation of 30 min at 30 °C a blue colored complex formed due to the phosphate generated during the oxidation reaction (Saheki et al. 1985). By measuring the absorption at 600 nm, the concentration of produced phosphate could be quantitatively determined using a calibration curve. By being able to quantify the amount of applied BVMO, each conversion could be performed with a standard amount of BVMO which allows a direct comparison.

GC analysis

For GC analysis, 500 μl incubations of 50 mM Tris–HCl (pH 7.5), containing 5 mM substrate, 5% cosolvent (DMSO or dioxane), 100 μM NADPH, 3.1 μM PTDH, 10 mM phosphite and 5 μM BVMO (CCE) was incubated shaking at 24 °C for 2 h. The reactions were then stopped by extracting with ethyl acetate (2×0.5 ml, including 0.1% mesitylene as an internal standard), dried with magnesium sulfate and analyzed directly by GC to determine the degree of conversion. For every tested reaction, control experiments in the absence of enzyme resulted in no conversion.

Results

Identification of putative BVMO-encoding genes

By using protein sequences of PAMO and CHMO as baits for BLASTP searches in the predicted proteome of *R. jostii* RHA1 (McLeod et al. 2006), and inspecting individual putative BVMO sequences for known sequence motifs, 23 putative Type I BVMO genes could be identified (Table 1). For consistency, we numbered the BVMOs according to the numbering employed in a previous study by the Grogan group (Szolkowy et al. 2009). Compared with this study, our set of BVMOs contains one more putative BVMO-encoding gene that was not identified before (BVMO24). Furthermore, it was previously suggested that BVMO8 represented a BVMO lacking a large part of the C-terminal domain which binds the NADPH coenzyme (RHA1_ro08998). Careful inspection of the sequenced genome revealed that a large part of a C-terminal BVMO sequence (ORF RHA1_ro08999) resides upstream of RHA1_ro08998 in the genome. Sequence analysis revealed that only one nucleotide was missing to create a gene encoding a full length BVMO, fusing both these open reading frames. The corrected BVMO8 gene was generated and included in this study (see Materials and methods). Except for these two newly reported BVMOs, BVMO8 and BVMO24, we also discarded one gene from our study that was previously assigned as a Type I BVMO (BVMO22 in Szolkowy et al. 2009). Analysis of the respective protein

Table 1 All identified BVMOs and their sequence characteristics from the predicted *R. jostii* RHA1 proteome

BVMO	Accession number	Rossman motif	BVMO motif	Rossman Motif	Length (aa)	Soluble expression[b]	Medium color
1	ro06679	GxGxxG	FxGxxxHxxxWP	GxGxxG	538	+ B	–
2	ro04304	GxGxxG	FxGxxxHxxxWN	GxGxxA	496	++ C	–
3	ro03247	GxGxxG	FxGxxxHxxxWP	GxGxxG	543	++ C	Blue
4	ro03063	GxGxxG	YxGxxxHxxxWR	GxGxxA	549	++ B	–
5	ro02109	GxGxxG	FxGxxxHxxxWN	GxGxxA	491	+ C	Blue
6	ro01874	GxGxxG	FxGxxxHxxxWN	GxGxxA	493	++ B	–
7	ro06008	GxGxxG	FxGxxxHxxxWN	GxGxxA	537	++ C	Blue
8[a]	ro08998/9	GxGxxG	FxGxxxHxxxWD	GxGxxA	545	++ C	–
9	ro09035	GxGxxG	FxGxxxHxxxWP	GxGxxG	541	++ B	–
10	ro09039	GxGxxG	FxGxxxHxxxWD	GxGxxS	663	++ B	–
11	ro06698	GxGxxG	FxGxxxHxxxWP	GxGxxG	545	++ B	–
12	ro07112	GxGxxG	FxGxxxHxxxWD	GxGxxA	514	++ C	Blue
13	ro03773	GxGxxG	FxGxxxHxxxWP	GxGxxA	515	+ C	–
14	ro03437	GxGxxG	FxGxxxLxxxWP	GxGxxG	547	++ B	–
15	ro02492	GxGxxG	FxGxxxHxxxWP	GxGxxG	553	++ B	–
16	ro02919	GxGxxG	FxGxxxHxxxWP	GxGxxG	498	++ C	–
17	ro05228	GxGxxG	FxGxxxHxxxWD	GxGxxA	603	+ C	Blue
18	ro05396	GxGxxG	FxGxxxHxxxWP	GxGxxS	652	++ C	–
19	ro05522	GxGxxG	FxGxxxHxxxWN	GxGxxA	515	+ C	Brown/blue
20	ro08137	GxGxxG	FxGxxxHxxxWP	GxGxxG	548	++ B	Brown/blue
21	ro10187	GxGxxG	FxGxxxHxxxWP	GxGxxG	546	++ B	Brown/blue
23	ro08185	GxGxxG	FxGxxxHxxxWP	GxGxxG	546	n.d	n.d.
24	ro05323	GxGxxG	FxGxxxSxxxWN	GxGxxG	564	++ B	Yellow

n.d. not determined

[a] This gene has been cured, see text

[b] ++ indicates a clear visible protein band using SDS-PAGE; + indicates that no clear band was visible but flavin reduction was observed; B indicates that the respective BVMO has been expressed using the pBADN vector; C indicates the use of the pCRE2 expression vector. The *underlined* BVMOs were previously expressed in soluble form in a pET expression vector (Szolkowy et al. 2009)

sequence revealed that it represents a typical flavin-containing monooxygenase (FMO), which is member of a flavoprotein family that is distinct from BVMOs (Fraaije et al. 2002). Furthermore, two BVMOs (BVMO21 and BVMO23) differed in only one amino acid out of 546 amino acids. Based on sequence alignment with PAMOs, this residue (Leu or Arg at position 480) is predicted to be on the surface of the protein and therefore will have little effect on the catalytic properties of these BVMOs. This has, in fact, been confirmed by the previous study of the Grogan group (Szolkowy et al. 2009). For this reason, BVMO23 was excluded from this study.

All identified BVMO sequences contained two conserved Rossmann fold motifs: one near the N terminus and one in the middle of the sequence (Table 1). In BVMO10 and BVMO18, the N-terminal GXGXXG motif is preceded by ~120 amino acids instead of only ~20 amino acids. This indicates that these enzymes have acquired an additional domain that is also found in a known BVMO, 4-hydroxyacetophenone monooxygenase (HAPMO), and which has been suggested to play a role in dimerization

(Kamerbeek et al. 2005). As described before, Type I BVMOs contain a typifying short sequence motif: FXGXXXHXXXW[P/D] (Fraaije et al. 2002). By allowing one or two conservative mutations in this motif, all 22 sequences were found to contain such sequence motif. It shows that not every residue of the previously described motif is strictly conserved. While the glycine and tryptophan residues are strictly conserved, some mutations (e.g., F>Y in BVMO4), also occur which reflect the origin of the sequence conservation within this motif. The BVMO motif is located in a loop region of BVMOs that allows spatial rearrangements during the catalytic cycle while it is not directly involved in catalysis (Orru et al. 2011).

For illustrating the sequence relationships between the identified *R. jostii* sp. RHA1 BVMOs and some of the best characterized BVMOs, all these sequences were aligned (see Fig. S1). This revealed that the rhodococcal BVMOs are representatives of many different clades of the Type I BVMO family. Careful inspection of the structure-based multiple sequence alignment also revealed another interesting conserved sequence motif.
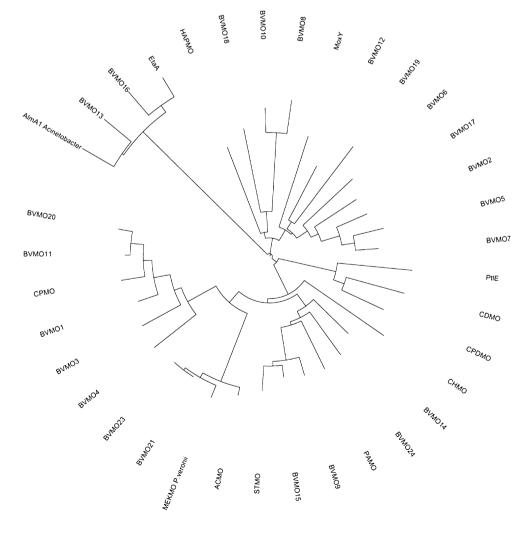
Except for the two Rossmann fold motifs and the BVMO motif mentioned above, there is one particular region of the sequences that shows remarkable conservation, located between the N-terminal GxGxxG motif and the known BVMO motif. This region can be defined by [A/G]GxWxxxx[F/Y]P[G/M]xxxD. The full conservation of the aspartate is in full agreement with the recent finding that Asp66 in PAMO plays a crucial role in catalysis by direct interaction with the NADP$^+$ coenzyme. Mutating this residue into an alanine resulted in inactivation of the enzyme by decreasing the rate of NADPH-mediated flavin reduction (Orru et al. 2011). Combined with Arg337 (in PAMO), these two residues represent the only fully conserved active site residues, both crucial for catalysis. The conserved doublet of glycine residues in this newly recognized motif is a known motif for a Rossmann fold domain and in the BVMOs they support effective binding of the FAD cofactor. The other conserved residues in the motif center around the flavin cofactor and appear to be essential for creating a productive active site architecture (see Fig. S2). This newly recognized motif appears to be more discriminative for identifying BVMO sequences as it entails more and more conserved residues. It also clearly differentiates between BVMO and FMO sequences as the structure and sequence region around Asp66 (in PAMO) in FMOs is totally different (Fraaije et al. 2002; Alfieri et al. 2008). Therefore, this motif represents a new and more reliable tool to annotate putative BVMO-encoding genes.

In order to obtain a better view on the sequence–function relationship among the tested set of BVMOs, another sequence alignment was also performed. Using the recently elucidated crystal structure of PAMO, which contains NADP$^+$ and a substrate analogue (MES), we identified all residues that are within a 12-Å radius from the sulfur atom of the substrate analogue (Orru et al. 2011). This yielded a set of 59 residues that form the core of the substrate binding pocket and that are spread over the PAMO sequence. Upon performing a structure-based multiple sequence alignment, we filtered for all respective 59 residues for generating a separate multiple sequence alignment. We anticipated that such an approach of comparing partial sequences would have a higher probability to disclose similarities in substrate specificities. In Fig. 1, the branching diagram representation of this focussed sequence alignment of all tested and known

Fig. 1 Radial branching diagram of the comparison of active site residues of the selected BVMO protein sequences. The included sequences are: *AlmA1, Acinetobacter* sp. DSM 17874 (ABQ18224.1); *EtaA, Mycobacterium tuberculosis* H37Rv (NP_218371.1); *HAPMO, Pseudomonas fluorescens* ACB (Q93TJ5.1); *MoxY, Aspergillus flavus* (AAS90037.1); *PtlE, Streptomyces avermitilis* MA-4680 (NP_824170.1); *CDMO, Rhodococcus ruber* SCI (AAL14233.1); *CPDMO, Pseudomonas* sp. strain HI-70 (BAE93346.1); *CHMO, Acinetobacter calcoaceticus* NCIMB 9871 (BAA86293.1); *PAMO, T. fusca* (YP_289549.1); *STMO, Rhodococcus rhodochrous* (BAA24454.1); *ACMO, Gordania* sp. strain TY-5 (BAF43791.1); *MEKMO, Pseudomonas veronii* MEK700 (ABI15711.1); *CPMO, Comamonas* sp. strain NCIMB 9872 (Q8GAW0)

BVMOs is shown. The grouping differed slightly from the global multiple sequence alignment and was used for the discussion below.

Expression of the putative BVMOs

All 22 putative BVMO-encoding genes were cloned into a slightly modified pBAD expression vector (Kamerbeek et al. 2004) using a ligation free cloning method. Subsequently, expression tests were performed in which the temperature and arabinose concentration was varied. With these constructs, 11 BVMOs could be obtained in soluble form with good expression levels, using *E. coli* TOP10 as host (Table 1). A similar success rate was previously obtained for roughly the same set of genes using a pET-based *E. coli* expression system (Szolkowy et al. 2009) (see Table 1). In the other cases, the expressed protein was either insoluble or the expression levels of soluble protein were very low. To increase the expression level and/or solubility, the 11 BVMO genes that could not be expressed using pBAD were cloned into the recently developed pCRE2 expression vector and tested using various expression conditions (Torres Pazmiño et al. 2009). This resulted in good expression levels of soluble protein of most of these genes. Only in five cases was poor soluble protein expression observed. Nonetheless, by using the NADPH-dependent flavin reduction method (vide infra) it was found that also these constructs led to expression of detectable levels of soluble NADPH-reactive proteins. Taken together, by using two expression vectors and careful tuning of expression conditions, we succeeded in producing all 22 targeted BVMOs in a soluble form.

Interestingly, for a number of constructs it was found that, using the optimal expression conditions, the growth medium changed significantly in color (Table 1). Such a phenomenon had been observed before in a BVMO mutagenesis study where the M446G PAMO mutant was created (Torres Pazmiño et al. 2007). This mutant BVMO was found to form indigo blue by oxidizing indole which is present in the cytosol of *E. coli*. Also for some of the expressed rhodococcal BVMOs a blue color was observed while some other constructs led to a more brownish or intense yellow color. Apparently, some of the expressed BVMOs act on endogenous compounds present in the culture medium or in the cells which confirms functional expression of these monooxygenases.

Substrate profiling

To establish the biocatalytic potential of this large set of BVMOs, we aimed at testing each of these BVMOs for activity on a large panel of potential substrates. For such substrate profiling study, we first set out to quantify the amount of active enzyme in CCEs. Based on the fact that all Type I BVMOs contain a FAD cofactor that is specifically reduced by NADPH, we developed a simple method for accurate BVMO concentration determination. This approach is somewhat analogous to the well-known method to quantify the amount of P450 monooxygenases by using the specific heme absorption features of reduced P450 monooxygenases with carbon monooxide. In *E. coli* only a minority of the native proteins contains a flavin cofactor and most of them will not be reactive with NADPH. In contrast, the overexpressed BVMOs will be dominant in the respective extracts, outnumbering the endogenous flavoproteins of *E. coli*. By measuring the absorbance spectrum of the CCE before and after the addition of NADPH the amount of NADPH-reactive flavin cofactor in the expressed BVMO can be determined (Fig. 2). Therefore an extinction coefficient is needed which reflects the absorbance difference between an oxidized and reduced BVMO. For this, we used the extinction coefficients for oxidized PAMO ($\varepsilon_{441\ nm\ ox}$= 12.3 mM^{-1} cm^{-1}) and reduced PAMO ($\varepsilon_{441\ nm\ red}$= 3.6 mM^{-1} cm^{-1}) to calculate an estimate of the difference extinction coefficient for FAD containing BVMOs at 440 nm ($\varepsilon_{440\ nm,\ BVMO}$=8.8 mM^{-1} cm^{-1}). We decided to use the absorption at 440 nm and to neglect the small differences caused by different flavin microenvironments in individual BVMOs. This corroborates with flavin absorbance spectral data of known BVMOs from literature which display highly similar absorbance maxima and extinction coefficients (Table 2). Validation of the method by using purified enzyme with a known concentration showed that the method is accurate. Furthermore, using extracts of *E. coli* devoid of any expressed BVMO did not show any significant A_{440} change after incubation with NADPH, indicating that there is no interfering absorbance. Using this method, it was possible to determine the concentration of soluble BVMO in all 22 samples, independent of the level of BVMO overexpression. In this way, we could normalize the amount of
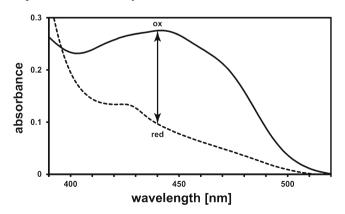

Fig. 2 Absorbance spectra of a cleared cell extract containing overexpressed BVMO. The spectra of the oxidised flavin cofactor in BVMO14 (*continuous line*) and after reduction by NADPH (*broken line*) are shown. The difference in absorbance at 440 nm that can be used to determine the BVMO concentration is indicated

Table 2 Collection of literature and measured values for λ_{max} and ε_{ox} of known BVMOs

	PAMO (Fraaije et al. 2005)	HAPMO (Kamerbeek et al. 2001)	CHMO (Sheng et al. 2001)	STMO (Miyamoto et al. 1995)	CPMO[a]	CPDMO[a]
Buffer	50 mM Tris/Cl, pH 7.5	50 mM phosphate, pH 7.0	50 mM phosphate, pH 7.2	100 mM phosphate pH 7.4	50 mM Tris/Cl pH 7.5	50 mM Tris/Cl pH 7.5
λ_{max} (nm)	441	439	440	439	444	442
ε_{ox} (mM^{-1} cm^{-1})	12.3	12.4	13.8	n.r.	14.5	12.8

n.r. no extinction coefficient has been reported

[a] These BVMOs were purified as PTDH fusion enzymes and analysed spectrophotometrically (this study)

recombinant enzyme in the substrate profiling experiments while using CCEs.

For substrate profiling we used 39 substrates, representing different compound classes: linear aliphatic ketones, cyclic aliphatic ketones, aromatic ketones, aromatic amines, and aromatic sulfides (Table 3). All substrates were used at two different concentrations (5 and 1 mM) and a fixed amount of cosolvent was used. For determining activity for each tested compound, we initially used the NADPH depletion assay. Consumption of NADPH by BVMO can be monitored by measuring the absorbance of NADPH at 340 nm. However, this assay could not be used for compounds that absorbed in the same wavelength regime of NADPH (300–390 nm) and only allowed to measure the conversion of relatively small concentrations of NADPH and substrate. Therefore, we decided to use an assay that is coupled to the activity of the employed regeneration enzyme which allows us to monitor conversion over a longer time period. For this, PTDH was used which generates one molecule of phosphate for the regeneration of each NADPH molecule. The formed phosphate could be quantified by using a chromogenic reaction allowing spectrophotometric detection of the degree of conversion. The conversions were incubated for 2 h before measuring the phosphate concentrations. As BVMOs can also act as NADPH oxidases in the absence of a suitable substrate, formation of phosphate was only regarded as significant when it was significantly more when compared with the incubation without a test compound. By this, poor substrates will be overlooked while establishing profiles of efficient substrates for all tested BVMOs.

The observed substrate profiles for all tested BVMOs are summarized in Table 3. The results disclose a broad substrate scope for the produced set of 22 BVMOs. To verify the results, we checked a selection of substrates for conversion with several BVMOs by GC analysis which confirmed the majority of the identified substrates. For the majority of the tested compounds, we could find at least one active enzyme, even though the intensity of the activity varies a lot between the tested BVMOs. Nonetheless, for eight BVMO (BVMOs 1, 3, 6, 7, 8, 13, 16, and 17) no significant activity was observed for any of the tested compounds. However, it is worth noting that of these BVMOs, BVMO3, BVMO7, and BVMO17 appear to produce some indigo blue when expressed in *E. coli* (Table 1). This suggests that these BVMOs are indeed active while not very efficient on any of the compounds in our test panel. Another relevant observation is the fact that conversions using BVMO3, BVMO8, BVMO13 and BVMO19 in the absence of substrate results in a 4- to 5-fold higher background reaction when compared with all other tested BVMOs (~9 mM phosphate formed instead of ~2 mM). For BVMO13 and BVMO19 this can be explained by the fact that their expression was poor which dictated the use of relatively high amounts of extract which leads to a relatively high background NADPH consumption. Contrarily, BVMO3 and BVMO8 are well expressed. This suggests that these BVMOs display an exceptionally high uncoupling rate and therefore substrates may be unnoticed by monitoring NADPH consumption or phosphate formation. It also agrees with the fact that Szolkowy et al. (2009) have found that BVMO3 is indeed active on a number of ketones. For the BVMOs that did not display significant background NADPH consumption or uncoupling rates, the substrate acceptance profiles varied considerably. Among the tested BVMOs, only a small set of five BVMOs were shown to be active on ≥10 tested compounds, leaving only five potent BVMOs. This contrasts the catalytic potential of well characterized BVMOs that were also tested with the panel of test compounds (Table 3). As known from the literature, CHMO, CPMO, PAMO and HAPMO were indeed found to accept a broad range of compounds. Only STMO, for which a substrate profile had not yet been established, appeared rather selective by converting only nine of the tested compounds. Similar to the known prototype BVMOs, two rhodococcal BVMOs, BVMO4 and BVMO24, appeared to be very substrate tolerant as they both accepted the majority of the tested compounds. Their substrate profiles resemble, to a large extent, that of CPMO, with BVMO24 exhibiting greater efficiency when compared with BVMO4. It is worth noting that BVMO24 had escaped the previous rhodococcal BVMO exploration study by the

Table 3 Substrate profiling

		2	4	5	9	10	11	12	14	15	18	19	20	21	24	PA	CP	CH	ST	HAP
1	Acetone[a]		+								+				++		+			+
2	2-Butanone		+												++		+			+
3	2-Octanone		+	+		+							+	+	++		++	+		+++
4	2-Dodecanone		++	+									+				+++	+++		+++
5	3-Octanone		+												++		++			+++
6	4-Octanone	+	+									+			++		++	+		+
7	Methylvinylketone		+		+					+			+	+	++	+	++		+	+++
8	3-Methyl-2,4-pentanedione		+							+	+		+		++		++			+++
9	Cyclobutanone		+							+			+		++		++			+++
10	Cyclopentanone	+	+										+	+	++		++	++		+
11	Cyclohexanone	+	+										+		++		+++	++		+
12	Cyclopentadecanone		+														++	+		+
13	Oxocyclohexanecarbonitrile	+	+												++		++			+++
14	4-Methylcyclohexanone		+														+++	++		+
15	2-Propylcyclohexanone		+	+											+		+++	++	+	+
16	Dihydrocarvone		+			+									+		++	++		+++
17	Cyclopropylmethylketone	+	+										+	+			++			+++
18	Norcamphor		+										+	+	++		+++	++		+
19	Bicycloheptenone					+++		+					+	+	++		+++	++		+++
20	Thioanisole		+							+							++	++		+++
21	Benzylethyl sulfide		+					+	+						++		++	++		+++
22	Benzylphenyl sulfide		+						+						++		++	+		+++
23	Ethionamide		+		++					+					++	+++	+		+++	+++
24	Modafinil precursor		+											+	+++	+	+			
25	Thiacetazone												+	+	++	+	+			+
26	Nicotin		+				+										++	+		
27	Indole															+	+			
28	3-Acetylindole				++					+							+			+++
29	Acetophenone		+										+		++		++		+++	+++
30	4-Hydroxyacetophenone		+		+	+				+++					+		+	++		+++
31	2,6-Dihydroacetophenone		+														+	++	+++	
32	3-Phenylpentane-2,4-dione									+					+		+++	++	+	+
33	Phenylacetone		+		+					+++			+	+	+++		++	++	+	+++
34	Raspberry ketone		+		++					+				+	+++	+	+		+++	
35	2-Phenylcyclohexanone																+	++	++	+
36	Benzoin																+			+++
37	Phenendione					+++										+	+			
38	2-Indanone	+													+		++	+++		+
39	1-Indanone												+	+			++			+
	Total substrates	6	29	3	6	5	1	1	2	10	3	1	14	11	25	12	38	19	9	33

Activities as measured by phosphate formation are indicated for the rhodococcal BVMOs and prototype BVMOs. BVMOs that did not show any activity on the tested compounds were excluded. The rhodococcal BVMOs are indicated by their respective numbers. The observed activities are indicated as +, ++ or +++ and reflect 1.2-, 2- or 5-fold the phosphate formation (cq. conversion), respectively, when compared with incubations lacking the tested compound

PA PAMO, *CP* CPMO, *CH* CHMO, *ST* STMO, *HAP* HAPMO

[a] For the structural formulas and other details, see Supplementary information

Grogan group. Of the other BVMOs, BVMO2 and BVMO20 also were found to be active on a number of cycloalkanones. BVMO9 and BVMO15 showed high activity with a number of aromatic ketones, resembling to some extend the substrate

scope of HAPMO. Other BVMOs appeared to be rather restricted in substrate scope. For example, BVMO10 appears to be very selective in acting on bulky ketones converting bicycloheptenone and phenendione. Still, these BVMOs may be valuable biocatalysts exhibiting new substrate specificities and chemo-, regio-, or enantioselectivities.

To have a better insight into the biocatalytic potential of some of the BVMOs, a selected set was studied for the conversion of some typical BVMO substrates, including prochiral sulfides. Table 4 shows that all these BVMOs convert the typical BVMO substrates phenylacetone and bicycloheptenone. BVMO8 gives poor conversion of phenylacetone while none of the other compounds are converted. This confirms the outcome of the activity screening which did not reveal any substrate for this enzyme. The fact that for phenylacetone conversion is observed is due to the fact that the conversions have been performed over a relatively long time period. For 2-indanone, only BVMO24 was found to be reactive which is also in line with the substrate profiling study (Table 3). Except for BVMO15, BVMO20 and BVMO8, the two aromatic sulfides were efficiently converted with good to excellent enantioselectivities. Strikingly, most of the BVMOs show the same enantioselectivity forming predominantly one specific sulfoxide enantiomer of each of the two aromatic sulfides. Only BVMO24 is again an exception and shows opposite enantioselectivity.

Discussion

By employing two expression vectors and screening for proper expression conditions, we have been able to express 22 Type I BVMOs from *R. jostii* RHA1. This is the largest set of recombinant BVMOs produced from one microbial origin reported so far. The collection supplements the previous set of 13 BVMOs produced by the Grogan group. Except for establishing the expression conditions for BVMOs that could not be expressed before, we also identified two new BVMO genes (BVMO8 and BVMO24). The biocatalytic exploration of this new collection of BVMOs

was aided by the development of two new methods: (1) by exploiting the distinct absorbance features of oxidized and reduced protein-bound FAD, the BVMO content in the extract could be determined; and (2) by performing the conversions coupled to NADPH regeneration by PTDH, conversion could be quantified by colorimetric phosphate detection. Sequence analysis also revealed a novel BVMO-typifying motif: [A/G]GxWxxxx[F/Y]P[G/M]xxxD. This motif includes residues that are part of the active site and therefore represents a useful tool for annotating BVMO-encoding genes. The motif is absent in FMO sequences and, therefore, allows discrimination between members of the two sequence-related flavoprotein subfamilies, BVMOs and FMOs.

The unraveled substrate profiles of the studied BVMOs differed significantly among the produced BVMOs. For some BVMOs no substrate could be identified by using a test set of 39 compounds. This could indicate that the natural substrates for these enzymes differ too much from the tested compounds. However, it may also be that these BVMOs only display very low oxygenation activities that could not be detected with the employed assay. Such low activity has also been observed for EtaA from *Mycobacterium tuberculosis* (Fraaije et al. 2004). This drug-resistance related enzyme was shown to act on a variety of typical BVMO substrates but with very low activity. A small number of the produced BVMOs were active on ≥10 compounds: BVMO4, BVMO15, BVMO20, BVMO21, BVMO24. BVMO24 appeared to be a very potent biocatalyst as it showed very good activities on 26 compounds. BVMO3 may also belong to this group of potent BVMOs but appears to act as NADPH oxidase, which prevented establishing its true substrate profile. To correlate sequence to substrate profiles, we generated a branching diagram that is based on the comparison of residues that make up the core of the active site. This revealed that all BVMOs that displayed a relaxed substrate acceptance profile were closely related to the clades that include the well known prototype BVMOs (CHMO, CPMO, PAMO and STMO). The above mentioned potent rhodococcal BVMOs, including BVMO3, are all closely related to these prototype BVMOs (Fig. 1).The

Table 4 GC analyses of substrate conversions by selected BVMOs

Substrate	BVMO							
	3	8	9	14	15	20	21	24
Phenylacetone	99	27	99	99	92	45	99	38
2-Indanone	<3	<3	<3	<3	<3	<3	<3	62
Bicycloheptenone	97	<3	36	97	90	97	97	93
Thioanisole	87	<3	21	99	76	24	55	80
Enantioselectivity	*90 (S)*	*n.d.*	*61 (S)*	*n.d.*	*45 (S)*	*5 (S)*	*63 (S)*	*82 (R)*
Benzylethyl sulfide	60	<3	<3	72	14	4	92	91
Enantioselectivity	*75 (R)*	*n.d.*	*n.d.*	*90 (R)*	*20 (R)*	*n.d.*	*93 (R)*	*>99 (S)*

The conversions are indicated as percentage (%). Enantioselectivities are indicated as enantiomeric excess (%). For some incubations, the enantioselectivity has not been determined (*n.d.*)

rhodococcal BVMOs that are part of sequence related groups that include other known but less explored BVMOs (HAPMO, MoxY, EtaA, AlmA) appear to be less promising biocatalysts. This observation is valuable for future BVMO discovery efforts: it appears more productive to search for homologs of the typical prototype BVMOs when looking for potent biocatalysts.

Two BVMOs display a peculiar behavior: BVMO3 and BVMO8 act as NADPH oxidases when no potential substrate is present. For the other BVMOs and all previously characterized Type I BVMOs, this futile oxidation of NADPH was typically very slow. The relatively high NADPH oxidase activity may reflect the fact that these respective BVMOs do not function properly when taken out of their natural context. Alternatively, it may indicate that these enzymes have acquired NADPH oxidase activity for a certain purpose, e.g., to regenerate $NADP^+$ in the cell or to produce H_2O_2. Production of H_2O_2 complies with the recent discovery that actinomycetes are relatively rich in peroxidases (van Bloois et al. 2010). In line with this, a DyP-type lignin peroxidase from *R. jostii* sp. RHA1 has recently been identified and characterized (Ahmad et al. 2011). Except for displaying NADPH oxidase activity, BVMO3 also was found to be acting as an oxygenase as the cells expressing this enzyme developed a blue color indicative of indole oxygenation. This would imply that BVMO3 and BVMO8 have a dual function in *R. jostii* RHA1: acting as oxygenases or NADPH oxidases. In this context, it is worth noting that, due to a missing nucleotide, the gene encoding BVMO8 is disrupted. This may reflect that due to the relative high NADPH oxidase activity, generating toxic hydrogen peroxide, it was beneficial to silence the gene. Further work on the enzymes reported in this study will disclose details on their catalytic properties and biocatalytic potential.

References

Ahmad M, Roberts JN, Hardiman EM, Singh R, Eltis LD, Bugg TD (2011) Identification of DypB from *Rhodococcus jostii* RHA1 as a lignin peroxidase. Biochemistry 50:5096–50107

Alfieri A, Malito E, Orru R, Fraaije MW, Mattevi A (2008) Revealing the moonlighting role of NADP in the structure of a flavin-containing monooxygenase. Proc Natl Acad Sci USA 105:6572–6577

Bocola M, Schulz F, Leca F, Vogel A, Fraaije MW, Reetz MT (2005) Converting phenylacetone monooxygenase into phenyl-cyclohexanone monooxygenase by rational design: towards practical Baeyer–Villiger monooxygenases. Adv Synth Catal 347:979–986

de Gonzalo G, Mihovilovic MD, Fraaije MW (2010) Recent developments in the application of Baeyer–Villiger monooxygenases as biocatalysts. ChemBioChem 11:2208–2231

Dudek HM, de Gonzalo G, Torres Pazmiño DE, Stępniak P, Wyrwicz LS, Rychlewski L, Fraaije MW (2011) Mapping the substrate binding site of phenylacetone monooxygenase from *Thermobifida fusca* by mutational analysis. Appl Environ Microbiol 77:5730–5738

Fraaije MW, Kamberbeek NM, van Berkel WJ, Janssen DB (2002) Identification of a Baeyer–Villiger monooxygenase sequence motif. FEBS Lett 518:43–47

Fraaije MW, Kamerbeek NM, Heidekamp AJ, Fortin R, Janssen DB (2004) The prodrug activator EtaA from Mycobacterium tuberculosis is a Baeyer–Villiger monooxygenase. J Biol Chem 279:3354–33560

Fraaije MW, Wu J, Heuts DPHM, van Hellemond EW, Spelberg JHL, Janssen DB (2005) Discovery of a thermostable Baeyer–Villiger monooxygenase by genome mining. Appl Microbiol Biotechnol 66:393–400

Hollmann F, Taglieber A, Schulz F, Reetz MT (2007) A light-driven stereoselective biocatalytic oxidation. Angew Chem 119:2961–2964

Kamerbeek NM, Moonen MJH, van der Ven JGM, van Berkel WJH, Fraaije MW, Janssen DB (2001) 4-Hydroxyacetophenone monooxygenase from *Pseudomonas fluorescens* ACB. A novel flavoprotein catalyzing Baeyer–Villiger oxidation of aromatic compounds. Eur J Biochem 268:2547–2557

Kamerbeek NM, Janssen DB, van Berkel WJH, Fraaije MW (2003) Baeyer–Villiger monooxygenases, an emerging family of flavin-dependent biocatalysts. Adv Synth Catal 345:667–678

Kamerbeek NM, Fraaije MW, Janssen DB (2004) Identifying determinants of NADPH specificity in Baeyer–Villiger monooxygenases. FEBS Lett 271:2107–2116

Kamerbeek NM, van der Ploeg R, Olsthoorn AJJ, Tahallah N, Heck AJR, Malito E, Janssen DB, Fraaije MW (2005) Exploring the role of the N-terminus of 4-hydroxyacetophenone monooxygenase. In: Nishino T, Miura R, Tanokura M, Fukui K (eds) Flavins and flavoproteins 2005, pp. 173–178. Tokyo

Letunic I, Bork P (2011) Interactive tree of life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acids Res 39:W475–W478

McLeod MP, Warren RL, Hsiao WWL, Araki N, Myhre M, Fernandes C, Miyazawa D, Wong W, Lillquist AL, Wang D, Dosanjh M, Hara H, Petrescu A, Morin RD, Yang G, Stott JM, Schein JE, Shin H, Smailus D, Siddiqui AS, Marra MA, Jones SJM, Holt R, Brinkman FSL, Miyauchi K, Fukuda M, Davies JE, Mohn WW, Eltis LD (2006) The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad Sci USA 103:15582–15587

Mihovilovic MD, Mueller B, Stanetty P (2002) Monooxygenase-mediated Baeyer–Villiger oxidations. ChemInform 34:3711–3730

Mihovilovic MD, Rudroff F, Grotzl B (2004) Enantioselective Baeyer–Villiger oxidations. Curr Org Chem 8:1057–1069

Miyamoto M, Matsumoto J, Iwaya T, Itagaki E (1995) Bacterial steroid monooxygenase catalyzing the Baeyer–Villiger oxidation of C21-ketosteroids from *Rhodococcus rhodochrous*: the isolation and characterization. Biochim Biophys Acta 1251:115–124

Orru R, Dudek HM, Martinoli C, Torres Pazmiño DE, Royant A, Weik M, Fraaije MW, Mattevi A (2011) Snapshots of enzymatic Baeyer–Villiger catalysis: oxygen activation and intermediate stabilization. J Biol Chem 286:29284–29291

Pas J, Stepniak P, Wyrwicz L, Ginalski K, Rychlewski L (2011) GRDB - gene relational dataBase. Bioinfobank Library Acta 11:2659. doi:10.3984/rep2659

Renz M, Meunier B (1999) 100 Years of Baeyer–Villiger oxidations. Eur J Org Chem 4:737–750

Saheki S, Takeda A, Shimazu T (1985) Assay of inorganic phosphate in the mild pH range, suitable for measurement of glycogen phosphorylase activity. Anal Biochem 148:277–281

Secundo F, Fiala S, Fraaije MW, de Gonzalo G, Meli M, Zambianchi F, Ottolina G (2011) Effects of water miscible organic solvents on the activity and conformation of the Baeyer–Villiger monooxygenases from *Thermobifida fusca* and *Acinetobacter calcoaceticus*: a comparative study. Biotechnol Bioeng 108:491–499

Seto M, Kimbara K, Shimura M, Hatta T, Fukuda M, Yano K (1995) A novel transformation of polychlorinated biphenyls by *Rhodococcus* sp. strain RHA1. Appl Environ Microbiol 61:3353–3358

Sheng D, Ballou DP, Massey V (2001) Mechanistic studies of cyclohexanone monooxygenase: chemical properties of intermediates involved in catalysis. Biochemistry 40:11156–11167

Stewart JD (1998) Cyclohexanone monooxygenase: a useful reagent for asymmetric Baeyer–Villiger reactions. Curr Org Chem 2:195–216

Szolkowy C, Eltis LD, Bruce NC, Grogan G (2009) Insights into sequence–activity relationships amongst Baeyer–Villiger monooxygenases as revealed by the intragenomic complement of enzymes from *Rhodococcus jostii* RHA1. ChemBioChem 10:1208–1217

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2279

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Torres Pazmiño DE, Snajdrova R, Rial DV, Mihovilovic MD, Fraaije MW (2007) Altering the substrate specificity and enantioselectivity of phenylacetone monooxygenase by structure-inspired enzyme redesign. Adv Synth Catal 349:1361–1368

Torres Pazmiño DE, Riebel A, de Lange J, Rudroff F, Mihovilovic MD, Fraaije MW (2009) Efficient biooxidations catalyzed by a new generation of self-sufficient Baeyer–Villiger monooxygenases. ChemBioChem 10:2595–2598

Torres Pazmiño DE, Dudek HM, Fraaije MW (2010) Baeyer–Villiger monooxygenases: recent advances and future challenges. Curr Opin Chem Biol 14:138–144

van Blois E, Torres Pazmiño DE, Winter RT, Fraaije MW (2010) A robust and extracellular heme-containing peroxidase from *Thermobifida fusca* as prototype of a bacterial peroxidase superfamily. Appl Microbiol Biotechnol 86:1419–1430

Wu S, Acevedo JP, Reetz MT (2010) Induced allostery in the directed evolution of an enantioselective Baeyer–Villiger monooxygenase. Proc Natl Acad Sci USA 107:2775–2780

Zambianchi F, Pasta P, Carrea G, Colonna S, Gaggero N, Woodley JM (2002) Use of isolated cyclohexanone monooxygenase from recombinant *Escherichia coli* as a biocatalyst for Baeyer–Villiger and sulfide oxidations. Biotechnol Bioeng 78:489–496