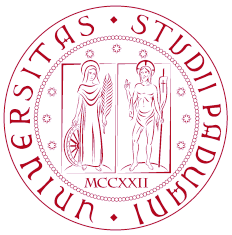


UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede amministrativa: Università degli Studi di Padova
Dipartimento di Biologia

SCUOLA DI DOTTORATO DI RICERCA IN: Bioscienze e Biotecnologie
INDIRIZZO: Biotecnologie
CICLO XXV

Production and Characterization of New Baeyer-Villiger Monooxygenases

Direttore della Scuola: Ch.mo Prof. Giuseppe Zanotti

Coordinatore d'indirizzo: Ch.mo Prof. Giorgio Valle

Supervisore: Dtt.ssa Elisabetta Bergantino

Dottorando: Elisa Beneventi

31 gennaio 2013

The reserach presented in this thesis was carried out in collaboration with F.I.S. - Fabbrica Italiana Sintetici S.p.A.

Abstract

The Baeyer-Villiger reaction is a very important oxidative reaction in organic synthesis that leads to the production of optically pure esters and lactones from ketones. It is typically performed by peroxyacids, oxydative reagents that are toxic, hazardous and lacking selectivity. By using "green chemistry", it is possible to perform the same reaction using "bio-catalyzers", enzymes named Baeyer-Villiger monoxygenases (BVMOs), a group of flavin dependent monooxygenases, that use NAD(P)H and molecular oxygen in order to catalyze chemo-, regio-, and enantio-selective oxidative reactions. BVMOs have been identified in a large number of bacteria and fungi but only a restricted number of them is available in recombinant form. In order to discover novel BVMOs, new sources of enzymes will become important.

Using bioinformatics tools, we have identified and selected five new putative type I BVMOs from different organisms: *Oryza sativa* (Os; plant), *Physcomitrella patens* (Pp; moss), *Cyanidioschyzon merolae* (Cm; red alga), *Trichodesmium erythraeum* (Te; cyanobacterium), *Haloterrigena turkmenica* (Ht; archeabacterium). In particular, basing on most updated literature, the photosynthetic eukaryotes *Oryza, Physcomitrella* and *Cyanidioschyzon* would result very uncommon sources for BVMOs.

The cloning and expression strategy used was the same for all the five putative sequences. The very low solubility of the expressed proteins was the real bottleneck of the work. To overcome the solubility problem many strategies have been performed lowering the inducer concentration, the growth temperature and using chaperones. Adding on the culture an excess of riboflavin, the precursor of the FAD, results successful to obtain soluble flavoproteins from *Physcomitrella patens* and *Cyanidioschyzon merolae*, showing the importance of the flavin cofactor in the folding process.

A biocatalytic characterisation of both expressed proteins was made at the Molecular Enzymology group of the Groningen Biomolecular Science and Biotechnology Institute (GBB), University of Groningen (The Netheralnds) headed by Prof. Dr. M.W. Fraaije, a main expert in flavoenzymes and particularly in Baeyer-Villiger monooxygenses. Once verified their Baeyer-Villiger monooxyge-

nase activity a condition optimization (pH, temperature and stability) has been performed; then the two proteins were tested to investigate their substrate specificity. A steady-state study was carried out in order to obtain kinetics parameters and many conversions were made to determine their selectivity profile.

The work presented led to discover new type I BVMOs enlarging the possibility to find out novel and promising biocatalysts for oxidative reactions.

Riassunto

La reazione di Baeyer-Villiger è una reazione di ossidazione di grande interesse sintetico in chimica organica e consiste nella conversione di composti carbonilici nei corrispondenti esteri o lattoni. L'approccio classico per eseguire la reazione di Baeyer-Villiger prevede l'utilizzo di agenti ossidanti quali perossidi o perossiacidi; questi catalizzatori sono tuttavia il più delle volte intrinsecamente insatabili e/o tossici mancando anche di enantioselettività. Ciò ha portato allo sviluppo di differenti sistemi catalitici che hanno implementano il concetto di "green chemistry" quali gli enzimi, utilizzati come biocatalizzatori. Gli enzimi che in natura catalizzano la reazione di Baeyer-Villiger sono flavoenzimi chiamati "Bayer-Villiger monoossigenasi" (BVMOs). Ad oggi solo pochi geni codificanti per le BVMOs sono stati clonati ed espressi e ciò ha rappresentato sicuramente un grosso ostacolo verso l'applicazione industriale di questa classe di enzimi.

Attraverso un'analisi bioinformatica è stato possibile identificare cinque sequenze codificanti proteine ritenute BVMOs putative, le quali presentano al loro interno il motivo caratteristico delle BVMOs di tipo I (FxGxxxHxxxW). Gli organismi individuati sono stati i seguenti: *Oryza sativa* (pianta), *Physcomitrella patens* (muschio), *Cyanidioschyzon merolae* (alga rossa), *Trichodesmium erythraeum* (cianobattere), *Haloterrigena turkmenica* (archeobattere). In particolare gli eucarioti fotosintetici *Oryza, Physcomitrella* e *Cyanidioschyzon* risultano fonti di BVMOs alquanto inusuali, considerando che fino ad oggi questi enzimi sono stati identificati solamente in batteri e funghi.

La strategia di clonaggio e di espressione utilizzate è stata la stessa per tutte e cinque le sequenze identificate. La bassa solubilità delle proteine espresse si è rivelata il fattore limitante di tutto il lavoro. Le strategie utilizzate per migliorare la solubilità proteica sono state quelle che hanno portano a una diminuzione della velocità della sintesi proteica come ad esempio la diminuzione della temperatura di crescita di coltura e la diminuzione della concentrazione dell'induttore. Un fattore risolutivo è stato l'aggiunta al terreno di coltura di riboflavina, il precursore del FAD, il quale ha permesso di ottenere in forma solubile le proteine putative da *Physcomitrella patens* e *Cyanidioschyzon merolae*, dimostrando inoltre l'importanza del cofattore flavinico nel processo di ripiegamento proteico.

La caratterizzazione biocatalitica delle proteine espresse è stata eseguita presso il gruppo di Biocatalisi del Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen (The Netheralnds), guidato dal Prof. Dr. M.W. Fraaije, esperto in flavoproteine. Confermata la loro attività come Baeyer-Villiger monoossigenasi è stata eseguita un'ottimizzazione delle condizioni di reazione (in termini di pH, temperatura e stabilità) ed è stata investigata la loro specificità di substrato. In seguito sono stati determinati i parametri cinetici per ciascun enzima e infine sono state eseguite delle bioconversioni con diversi substrati per capire il loro profilo di selettività.

Il lavoro riportato ha portato alla scoperta di nuove Baeyer-Villiger monoossigenasi di tipo I, ampliando il panorama di nuovi promettenti biocatalizzatori ossidoriduttivi.

Contents

Chapter 1

General introduction

1.1 Biocatalysis as a green synthetic tool

1.1.1 Green Chemistry

Traditionally Chemistry is defined as the study of the nature, properties, and composition of matter, and how these undergo changes. In organic chemistry carbon-based compounds are the matter of this study and, chemical reactions between organic compounds are used in nature and in organic synthesis to build-up new organic molecules. In general, these structurally different organic compounds represent not only the base of living organisms but they are important components of many products used in human life, such as drugs, plastics, petrochemicals and their derivatives. Development and optimization of chemical processes, employed for transforming basic organic compounds into useful products for human society, are what industrial chemistry is based on.

Approximately during the two last centuries, chemical industry dramatically transformed the world's economy; in fact chemical processes have created agrochemicals such as pesticides and fertilizers, drugs and pharmaceuticals products, synthetic dyes and chemical fibers, flavors for the food industry, plastics, synthetic resins, photo chemicals and many other chemical products which have played an important role in transforming modern society from a economical, environmental, even social and political point of view. Such a variety of goods has been extremely responsible for the high standard of life in modern industrialized society. On the other hand, all these benefits for human living led to a big exploitation of non-renewable natural sources and a massive use of toxic and hazardous feedstock and catalysts, generating dangerous by-products for the environment.

In the 1960s it had become evident that air, water and land pollution was overcoming the limits, so many countries started a notable legislative action to reduce pollutants from wasted sites. Many advanced technologies to control pollution were implemented in the so-called "command and control" approach, improving successfully the environmental quality. However, a better system would consist in preventing pollutions at its source rather than acting afterwards to minimize it.

In 1990 the Pollution Prevention Act was passed in the United States putting

the bases for the Green Chemistry approach, an innovative and philosophical way toward sustainability. In the same years, Anastas of the US Environmental Protection Agency (EPA) coined the term "Green Chemistry" [4] and in 1993 the EPA officially adopted the name "US Green Chemistry Program". Green chemistry can be applied to any chemical sector, redesigning products and processes that minimize the use of hazardous substances and maximizing the efficiency of any chemical choice [3, 6, 5, 4]. The essence of this new approach has been defined as follows [98]:

Green chemistry efficiently utilizes (preferably renewable) raw materials, eliminates waste and avoids the use of toxic and/or hazardous reagents and solvents in the manufacture and application of chemical products.

This concept is embodied into 12 basic Principles:

1. Waste prevention instead of remediation
2. Atom efficiency
3. Less hazardous/toxic chemicals
4. Safer products by design
5. Innocuous solvents and auxiliaries
6. Energy efficient by design
7. Preferably renewable raw materials
8. Shorter syntheses (avoid derivatization)
9. Catalytic rather than stoichiometric reagents
10. Design products for degradation
11. Analytical methodologies for pollution prevention
12. Inherently safer processes

Maximizing the percent yield was the traditional goal to evaluate a chemical process, while the new developed concept is focused in minimizing waste to the molecular level; this means that, in a designed process, the amount of starting material should end up in the generated product without atom waste. Atom economy can be improved carefully choosing suitable starting materials and cat-

alyst systems. In fact in many cases waste, generated in chemical processes, is a consequence of required stoichiometric amounts of reagents and catalysts. The use of catalytic quantities represents a solution to "clean" classical stoichiometric methodologies and if the catalyst is really "green" it can be recovered and reused many times in the same process [97, 27, 28].

1.1.2 Biocatalysis

Biocatalysis, that is the use of natural catalysts such as enzymes to perform chemical transformations on organic compounds, represents a good opportunity from the green chemistry point of view, because enzymes usually operate in mild reaction conditions (room temperature and physiological pH), are biodegradable catalysts derived from a renewable sources and use water as solvent [116]. Moreover, their chiral nature results in high chemo-, regio- and stereoselectivity with excellent rate acceleration, a very important feature for chemical organic synthesis. However, the fulfillment of an enzymatic reaction is not so obvious and can be affected by several parameters (Fig. 1.1), such as specific activity (quantified by k_{cat}), specificity (determined by k_{cat}/K_M), stability and the degree of inhibition by substrate or product. The ideal biocatalyst would have high specific activity, high selectivity and stability with a minimal substrate and product inhibition (Fig. 1.1) [57].

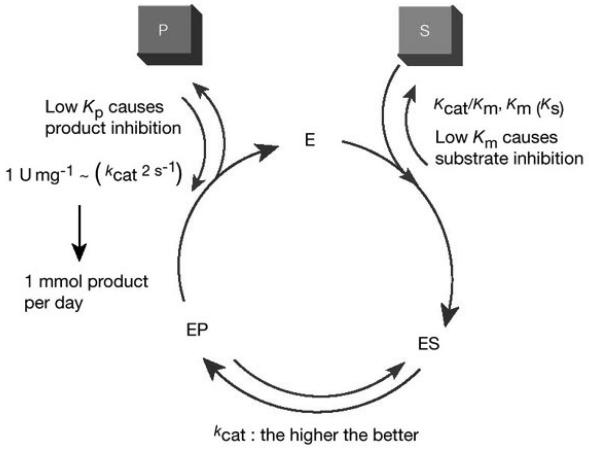


Figure 1.1: Parameters to be considered for enzymatic synthesis. (Image from Nature 409, 232-240 (11 January 2001), doi:10.1038/35051706).

Over the last two decades, advances in DNA recombinant technologies have enabled manipulation of enzymes and their consequent application in industrial organic synthesis. In fact, protein-engineering techniques have permitted alteration of their substrate specificity and stability ,and increase of their specific activity by developing enzyme variants with modified properties, adapted for new process conditions [13]. However the success of biocatalysis depends on the feasibility of the specific process from an economical point of view. Comparing conventional chemical methods with biocatalytic processes, the development of the latter ones is usually much longer to be established and requires inputs from different science fields such as chemical engineering, chemistry, biochemistry and molecular biology.

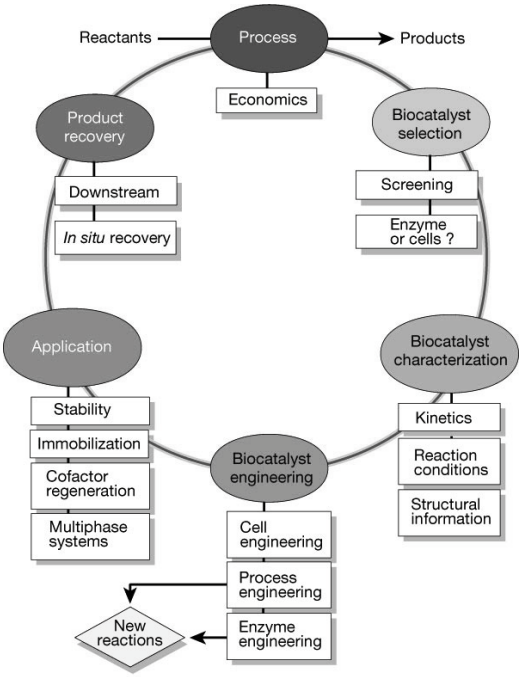


Figure 1.2: The biocatalysis cycle. (Image from Nature 409, 258-268(11 January 2001), doi:10.1038/350517369).

To develop a biocatalytic process (Fig 1.2) the starting point should usually be the product, which can be produced by one or more biocatalytic reactions. One or more biocatalysts can be identified or developed, but the key point is that resulting bioconversions need to be economically feasible. Thus, an economic analysis-based approach is necessary to identify process bottlenecks and try to

overcome these limitations in an iterative way, gradually leading to an efficient industrial process [96]. On the basis of the impact of biocatalysis on the synthesis of chiral drug candidates, several pharmaceutical companies have dedicated great emphasis on biocatalysis for organic synthesis. Some examples are BASF, DSM, Lonza, Merck, Pfizer, Bristol Myers Squibb and Codexis.

1.2 Monooxygenases

Monooxygenases (EC 1.13.x.x and EC 1.14.x.x) are oxygenase enzymes able to catalyze the transfer of one atom of oxygen to an organic substrate, and reduce the other atom of oxygen to water. Monooxygenases represent very interesting enzymes for synthetic purpose because the insertion of one atom of oxygen into an organic molecule is difficult to obtain in a chemical way. In fact molecular oxygen by itself cannot perform a direct oxidation reaction because of its spin-state, but it needs activation. Oxidizing enzymes such as monooxygeanases (but also dioxygenases) are able to carry out activation providing electrons to molecular oxygen, after which oxygenation of the organic substrate can occur [17]. There are many reactions catalyzed by these enzymes interesting from a synthetic point of view such as heteroatom oxidation, aromatic and aliphatic hydroxylation, epoxidation and Baeyer-Villiger oxidation (Fig 1.3).

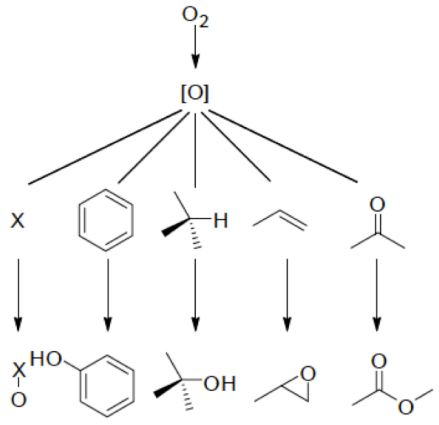


Figure 1.3: Examples of reactions catalysed by monnoxygenase. (Image from Enzyme Catalysis in Organic Synthesis. A Comprehensive Handbook, Vol. III).

Generally they utilize NADH or NADPH to provide reducing potential (external monooxygenases) together with a wide range of prosthetic groups/cofactors, including iron (e.g., cytochrome P450), copper (e.g., dopamine β monooxygenase), pterins, and flavins. In some cases no cofactor is even present; such internal monooxygenases provide electrons from the substrate itself.

Internal monooxygenase reaction: $XH_2 + O_2 \rightarrow XO + H_2O$
External monooxygenase reaction: $X + O_2 + DH + H^+ \rightarrow XO + H_2O + D^+$

1.3 Baeyer-Villiger Monooxygenases

1.3.1 Baeyer-Villiger reaction

The Baeyer-Villiger reaction was first reported by Adolf von Baeyer and Victor Villiger in 1899 [8]. In this organic reaction ketones are oxidized into corresponding esters or lactones by peroxyacids resulting in oxygen insertion next to the carbonyl group. The reaction mechanism consists in the nucleophilic attack of peroxy acid to the carbonyl function forming a tetrahedral intermediate called the "Criegee intermediate" which undergoes rearrangement to the corresponding ester or lactone; one of the group attached to carbonyl carbon migrates to the electron deficient oxygen atom with the simultaneous dissociation of the O-O bond; proper alignment of orbitals is required for the rearrangement step, which is the rate-limiting step of the reaction [32].

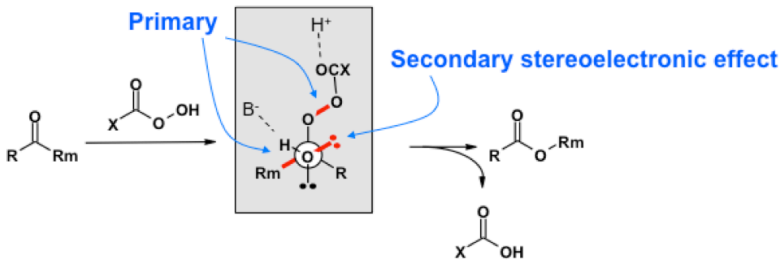


Figure 1.4: Mechanism of the Baeyer-Villiger reaction.

The regiochemistry of the reaction depends on the relative migratory ability

of the substituents attached to the carbonyl group (Fig. 1.4). Substituents which are able to stabilize a positive charge migrate more readily, so that the migratory aptitude is: tertiary alkyl > cyclohexyl > secondary alkyl > phenyl > primary alkyl > CH3 [90]. Stereoeletronic factors can also affect the regiochemical result; indeed, in the transition state the antiperiplanar alignment of the migrating substituent (Rm) and the O-O bond of the leaving group (primary stereoeletronic effect) and another antiperiplanar alignment of alone pair of oxygen with the migrating group (secondary stereoeletronic effect), are required. Moreover, if the migrating carbon is chiral the stereochemistry is retained [30]. As above mentioned, reagents able to carry out the Baeyer-Villiger reaction are peroxy acids such as meta-chloroperoxybenzoic acid (m-CPBA), trifluoroperoxyacetic acid (TFPPA), peroxyacetic acid, monopermalic acid, p-nitrobenzoic acid and hydrogen peroxide. Organic peroxy acids are expensive, toxic and hazardous; they are strong oxidants and react with other functional groups. Moreover being achiral they lack selectivity, a very important feature in organic synthesis, because of the need to produce enantiopure compounds. Many therapeutic drugs, but also fragrances and agrochemicals can achieve their biological effect only in a specific enantiomeric form and enzymatic reactions are very good tools to perform asymmetric catalysis.

1.3.2 Enzyme-catalysed Baeyer-Villiger reaction

Baeyer-Villiger monooxygenases are flavin-dependent monooxygenases able to catalyse nucleophilic oxidation of carbonyl groups but also the electrophilic oxidation of heteroatoms such as sulphur, nitrogen and boron, with good enantioselctivity. In fact, from an industrial point of view, the great potential of these enzymes is the possibility to produce chiral lactones and sulfoxides by asymmetric synthesis, attractive for organic synthesis. BVMOs contain flavin cofactor, FAD or FMN, not covalently but tightly bound to the enzyme. This flavin-cofactor has to be activated by electron donors, NADH or NADPH, which give reduction equivalents allowing molecular oxygen binding. Depending on cofactors usage, BVMOs can be classified into two types [114].

Type I BVMOs are FAD and NADPH dependent and contain two-dinucleotide

binding-domain ($\beta\alpha\beta$-folds) known as Rossman motifs, one for the FAD binding and the other one for the NADPH binding. They are composed of only one polypeptide. Type II BVMOs are FMN and NADH dependent and consist into two distinct subunits, a dehydrogenase subunit using NADH to reduce FMN and a second subunit able to perform the Baeyer-Villiger reaction using reduced flavin. There is another monooxygenase, MtmOIV, involved in the biosynthesis of mithramycin, able to perform the Baeyer-Villiger reaction. Mithramycin is a polyketide anticancer antibiotic produced by the soil bacterium *Streptomyces argillaceus* (ATCC 12956) and other bacteria of the genus *Streptomyces* [9]. Sequence analysis and crystal structure revealed that MtmVIO cannot be classified as type I nor type II BVMO but it appears to be an atypical BVMO belonging to another flavoprotein family (for flavoprotein monooxygenases classification, see [12]. The main described and characterized enzymes are type I BVMOs, which display good substrate promiscuity and are thus attractive from a synthetic point of view [74].

1.3.3 Mechanism of action of type I BVMOs

The first general accepted mechanism, proposed by Walsh et al. [92], was designed for cyclohexanone monooxygenases (CHMO) from A. *calcoaceticus* NCIMB 9871. Mechanistic studies were then performed by Massey et al. on the same enzyme [99], elucidating the presence of intermediates involved in catalysis. Further investigation by Torres Pazmiño et al. on the kinetic mechanism of PAMO has elucidated the ability of these enzymes to catalyze not only the nucleophilic attack on the carbonyl group but also the eletrophilic reaction according to the peroxy-intermediate formed [79].

The biocatalytic cycle (Fig. 1.5) starts with FAD reduction mediated by NADPH to form an enamine intermediate, which readily reacts with molecular oxygen leading to the nucleophilic 4a-peroxyanion in equilibrium with the corresponding electrophilic 4a-hdroperoxyflavin [99]. Thus, depending on the protonation state of the formed oxygenated flavin-intermediate, flavoprotein monooxygenases are able to perform Baeyer-Villiger oxidation or heteroatom oxidation, showing the double oxidative functionality (nucleophilic and electrophilic) of the

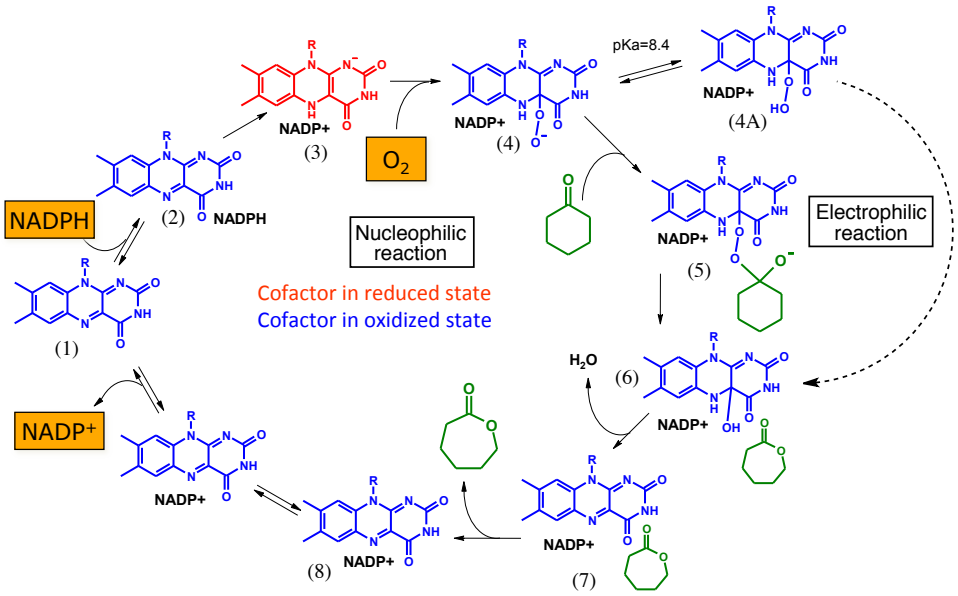


Figure 1.5: Mechanism of action of flavin-dependent type I BVMOs.

reactive intermediate species. In the nucleophilic oxidation, 4a-peroxyflavin attacks the keto group (C=O) of the substrate to produce a tetrahedral Criegee intermediate, which rearranges forming an ester or lactone and a hydroxyflavin (FAD-OH); it then eliminates water to close the catalytic cycle. It is known that the oxidized $NADP^+$ remains in the active site till the oxygenation occurs, stabilizing the negatively charged peroxyflavin-intermediate. In the protonated form, hydroperoxyflavin is able to oxidize the substrate by electrophilic addition and, as the product is released from the enzyme complex, FAD-OH is formed as for the nucleophilic process [79].

1.3.4 Structural features

The first crystal structure of a BVMO was solved by Malito et al. in 2004 [69], who published the three dimensional structure of PAMO (pdb-code 1W4X). The enzyme has two domains with the active site located in a cleft at the domain interface. The FAD-binding domain is linked to the NADPH-binding domain by a linker region containing the "fingerprint" motif FxGxxHxxxWP/D, a conserved sequence shared by all type I BVMOs. Moreover each domain contains a

dinucleotide-binding motif (part of Rossman fold), one for the FAD and the other for NADPH/NADP$^+$. The latter one interacts with two important residues, R217 and K336, having a role in affinity and recognition of the ribose moiety of NADPH. Located on the *re* side of the flavin ring there is R337, another conserved residue among BVMOs. It adopts two alternative conformations during the biocatalytic cycle: a IN position, which stabilizes the negatively charged flavin-peroxide intermediate, and an OUT position, responsible for the positioning to perform FAD reduction. Finally H173, which is part of the "fingerprint" motif, has a critical role in FAD binding, domain rotation and conformational changes during the biocatalysis [69]. Since 2011 different 3D-structures of PAMO were solved elucidating the structural mechanism of the oxidation process [83].

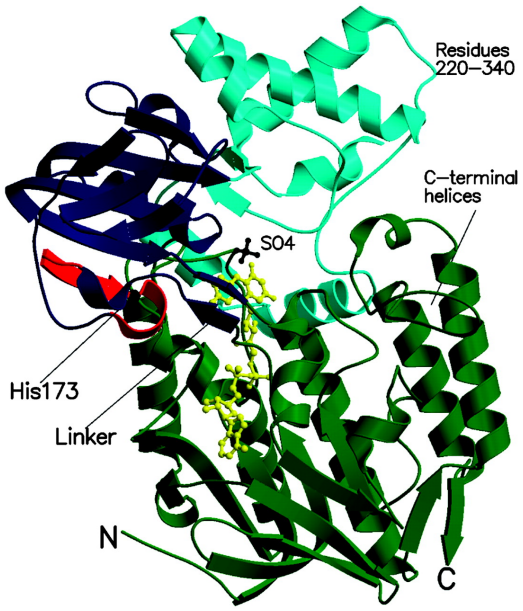


Figure 1.6: Crystal structure of PAMO dimer. (Image from PNAS,101(36),2004) [69].

In addition to PAMO, other BVMO structures were solved in recent years. In 2009 MtmOIV from *Streptomyces argillaceus* (ATCC 12956) was crystallized by Beam et al. (pdb-code 3fmw) [9]. The enzyme, involved in mithramycin biosynthetic pathway, is a dimer lacking the characteristic conserved "fingerprint" motif of type I BVMOs. It is considered an example of atypical class A flavoprotein monooxygenase catalyzing a Baeyer-Villiger oxidation, since it requires perox-

yflavin intermediate for the nucleophilic attack to the carbonyl function instead of the hydroxyl-flavin to perform an electrophilic oxidation. Its active site contains R52, the equivalent of R337 in PAMO, but located at the *si* side of the flavin ring.

Another crystal structure is the one of CHMO from a soil actinomycete *Rhodococcus* sp. HI-31 solved in two different crystal forms, closed and opened (pdb-code $CHMO_{closed}$ 3gwd, pdb-code $CHMO_{open}$ 3wdf). The enzyme has a three-domain architecture, a FAD binding domain a NADPH binding domain and, in between the two ones, a modulating domain which contains residues of the substrate binding poket. Moreover, numerous loops give to the enzyme a high conformational flexibility, in particular concerning the two different conformations of the nicotinamide cofactor domain. This suggests a large shift responsible for changing the size and accessibility of the substrate-binding pocket allowing the accommodation of a broad range of substrates [77].

In 2012 OTEMO, a type I BVMO catalyzing the lactonization of 2-oxo-3-4,5,5-trimethylcyclopentenylacetylcoenzymeA (CoA), a key intermediate in the metabolism of camphor by *Pseudomonas putida* ATCC 17453, has been solved [64]. The enzyme is a dimer in solution; each monomer consists of three domains (a FAD-binding domain, a NADPH binding domain and a "flap" domain) with bound flavin (pdb code 3UOV) as well as FAD and NADP (pdb code 3UOY). A comparison of various crystals of OTEMO bound to FAD and $NADP^+$ revealed the conformational plasticity of several loop regions, some of which have been implicated in contributing to the substrate specificity profile of structurally related BVMOs [33].

In the same year the crystallographic structure of steroid monooxygenase (STMO) from *Rhodococcus rhodochrous* (pdb code 4AOX), which exhibits the typical two-domain organization, was also solved . The active site resembles that of PAMO and the binding-site for $NADP^+$ with respect to the flavin is shifted compared with what observed in other type I BVMOs. This finding fully supports the idea that NADP(H) adopts various positions during the catalytic cycle to perform its multiple functions in catalysis [38].

Also the structure of a type II 3,6-diketocamphane monooxygenases (3,6-DKCMO) from *P. putida* NCIMB 10007, has been solved using a noncrystallo-

graphic symmetry (NCM) study (pdb code 2WGK). The enzyme complex is a trimer containing a NADH-dehydrogenase subunit and a homosymmeric FMN-binding monooxygenase component [25].

1.3.5 Enzymes and cofactor regeneration

Baeyer-Villiger monooxygenases and more in general oxydoriductases are cofactor-dependent enzymes because they require electrons to activate molecular oxygen. In such enzymes electrons are provided by nicotinamide coenzymes (NAD(P)H). The main drawback in the use of these cofactors is their high costs, limiting their use in stoichiometric amount, so they have to be regenerated *in situ* for large-scale applications. Many approaches have been developed to use this biomolecules in a catalytic amount facilitating the use of isolated enzymes [22, 21, 110]. Some important requirements are necessary for a good cofactor regeneration system. The total torn over number (TTN), for example must be high enough to make the regenerative process economically feasible (from 10^3 to 10^5) moreover, it reflects several important factors such as concentration and degradation of the cofactor over time, regioselectivity of regeneration, rate of catalysis, and time of reaction. In addition the regeneration step has to be favourable from a kinetic and thermodynamic point of view and, reagents and by-products, preferably stable in the process conditions, should not interfere with the whole system.

Enzymatic reactions are good candidates to be a regeneration system (Fig. 1.7). The most studied enzymatic methods have been developed for the regeneration of NAD(P)H, and involves mainly dehydrogenases such as formate dehydrogenase (FDH), glucose-6-phosphate dehydrogenase (G6PDH) [16], glucose dehydrogenase (GDH) [115] alcohol dehydrogenase (ADH)) and the phosphite dehydrogenase (PTDH) [112], the latter one was also used as fusion protein covalently linked to a BVMO to create a self-sufficient enzyme, able to perform the Baeyer-Villiger oxidation and cofactor regeneration (Fig. 1.8) [79].

Chemical regeneration of nicotidamide cofactors provides the use of organometallic-complex such as the rhodium complex $[Cp^*Rh(bpy)(H_2O)]^{2+}$ (Fig. 1.9) [48]. Ottolina and co-workers utilized this complex for FAD regeneration in BVMOs-mediated sulfoxidation but unfortunately the reaction conversion and

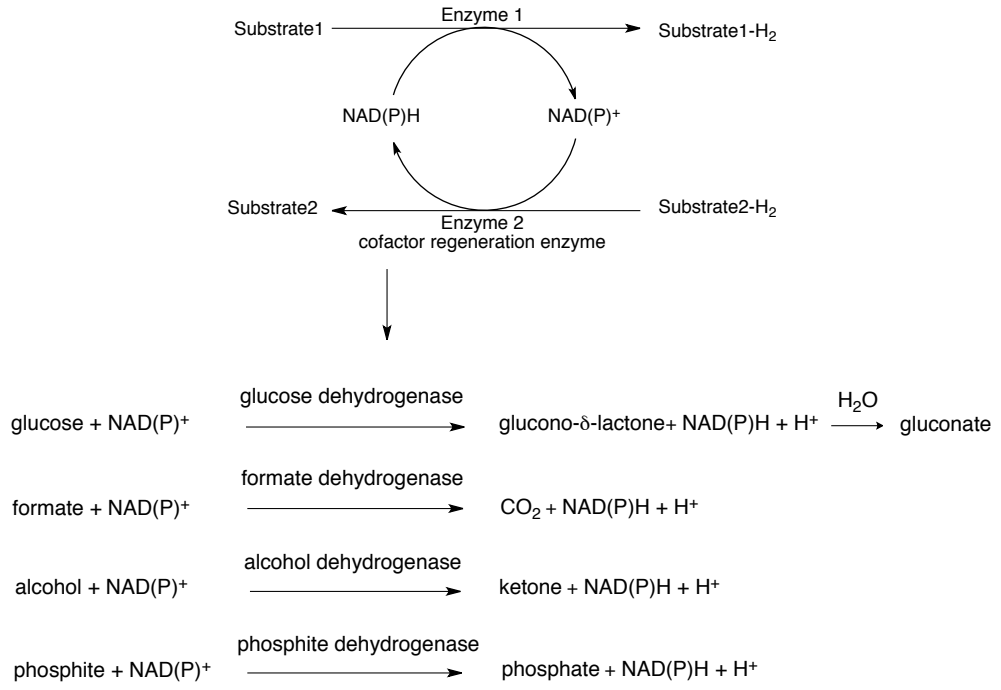


Figure 1.7: Enzymatic NAD(P)H cofactor recycling systems.

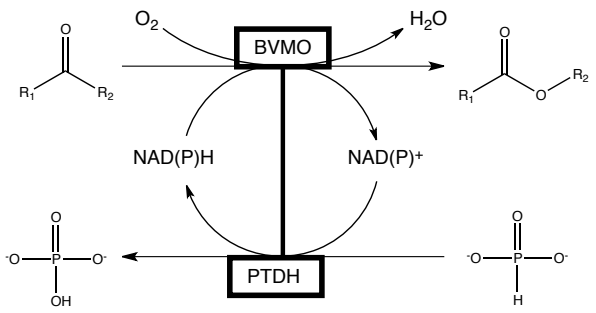


Figure 1.8: Coenzyme regeneration by fusion enzyme.

product *e.e.*'s were lower respect values obtained with the enzymatic regeneration system (G6PDH).

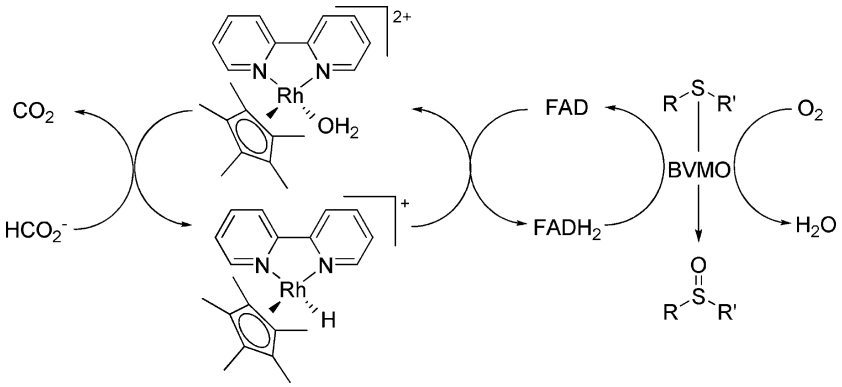


Figure 1.9: Organorhodium complex $[Cp^*Rh(bpy)(H_2O)]^{2+}$ regeneration in BVMO-catalyzed oxidation of sulfides.

The use of an electrode by direct or indirect reduction of $NADP^+$ using redox mediators is employed in electrochemical regeneration systems, additionally redox-active enzymes can be use to shuttle electrons from electrode to the coenzyme (Fig. 1.10) [46].

Photochemical regeneration is a light-driven regeneration system in which light can reduce FAD in the presence of an electron donor such as theethylenediaminetetraacetate (EDTA) or mercaptoethanol with a catalytic amount of NADPH for the substrate tournover (Fig. 1.11) [47].

1.3.6 Whole-cells and large scale application of BVMOs

To avoid in situ cofactor regeneration systems is possible perform whole-cell biooxidations using growing o resting cells. The organism expressing the BVMO provides natural recycling system for all coenzymes. However, there are also some disadvantages such as unwanted side reactions by competing enzymes or the fact that substrate and product can be used by others proteins present in the cells, mainly if the wild-type organism is utilized or if the enzyme is not overexpressed. Another limitation in whole-cell biotransformations is the possible substrate or product inhibition. Also for this reason SPRF (substrate feeding and product removal process) were developed (Fig. 1.12).

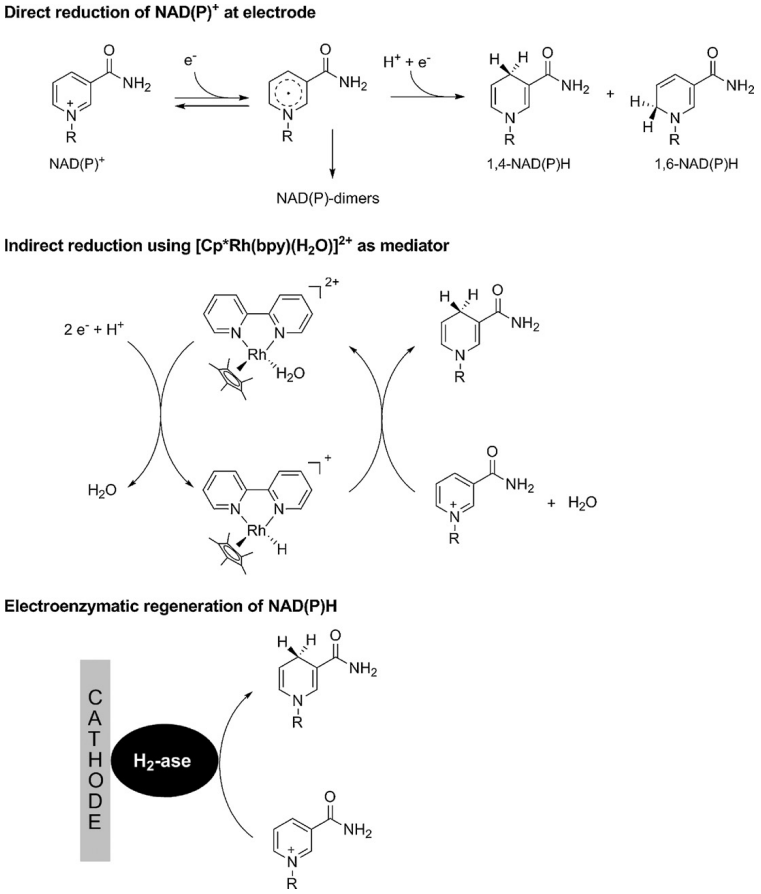


Figure 1.10: Available methods for electrochemical regeneration of NAD(P)H. (Image from http://dx.doi.org/10.1016/j.jbiotec.2010.01.021)

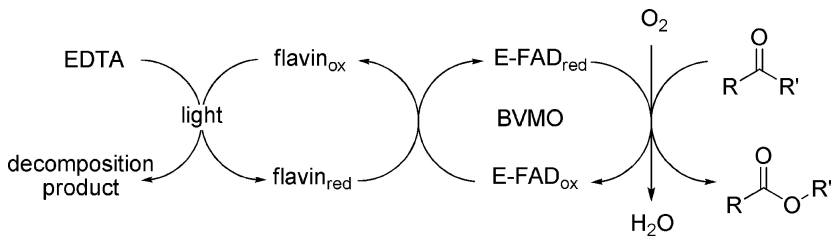


Figure 1.11: Light reduction of FAD in the presence of ethylenediaminetetraacetate.

The industrial applicability of the Baeyer-Villiger oxidation, was demonstrated with the first kilogram scale-up of asymmetric whole-cell bioconversion of the racemic bicyclo[3.2.0]hept-2-en-6-one, using a "resin-based in-situ substrate feeding and product removal" methodology (*in situ* SFPR). Substrate ketone and product lactone were adsorbed into a resin, which acts as a reservoir, allowing slow realise of substrate and at the same time continuous removal of formed product and maintaining both the substrate and product concentration below the inhibitory level. In such as system oxygen transfer is a critical factor. Using specially designed bubble column reactor, oxygenation of medium can be easily improved [44]. Down streaming process consists in a simply extraction of lactones from the resin, making it reusable. This technology gives the possibility to avoid some limitations such as substrate/product solubility and inhibition with a simplified downstream process.

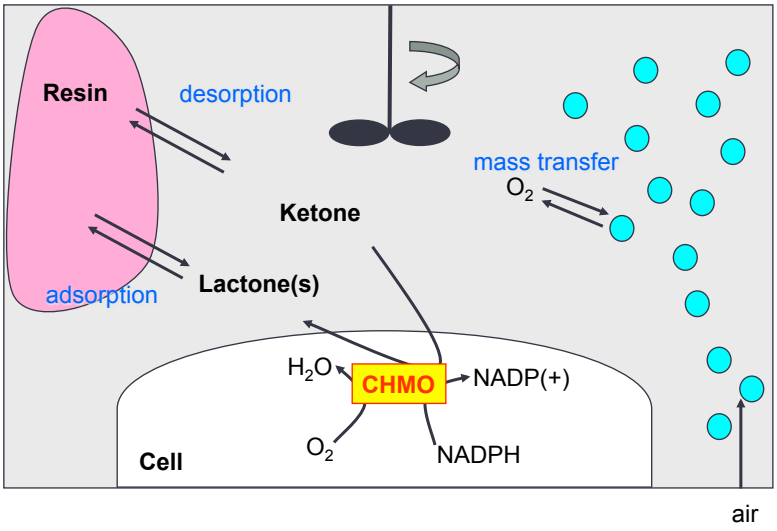


Figure 1.12: Principle of a microbial Baeyer-Villiger oxidation using a resin-based in situ SFPR strategy.

Using optimal conditions is possible to run a 900 g scale experiment, using a 50 L reactor at a substrate concentration of about 25 g/L. This reactor thus provided a volumetric productivity as high as 1.02 g lactones/L/h (136 U/L) [45].

1.3.7 Synthetic applications

Desymmetrization reactions are important applications of enzymatic Baeyer-Villiger oxidation in organic chemistry. Desymmetrization consists in the introduction of chirality into a symmetric element converting a prochiral precursor in a theoretically 100% yield of pure chiral product. Biooxidation of prochiral functionalized cyclohexanones and cyclobutanones is a good exemple of desymmetrization with high stereoselectivity. Also desymmetrization of fused bicycloketones were investigated, in particular Mihovilovic et al. have analysed the enantiocomplementary lactones produced by some representatives BVMOs observing a relevant clustering into two groups, CHMO-type and CPMO-type BVMOs, the first subclass give both lactones, the second one give the antipodal products with moderate to excellent enantiomeric excess [76]. An appealing application of BVMOs in desymmetrization processes is the synthesis of R-baclofen, a derivative of gamma-aminobutyric acid (GABA), acting as agonist for the $GABA_B$ receptors [71]. It can be prepared from the 3-(p-chlorophenyl)-γ-butyrolactone in excellent optical purity from biotransformations utilizing wild-type cells of *Cunninghamella*. Using the same strain 3-(benzyloxymethyl)-cyclobutanone can be oxidized to another active intermediate, the (R)-(–)-γ-buryrolactone, which is transformed via stereodivergent synthesis in (R,S)-β-proline [71]. All these compounds are interesting from a pharmaceutical point of view.

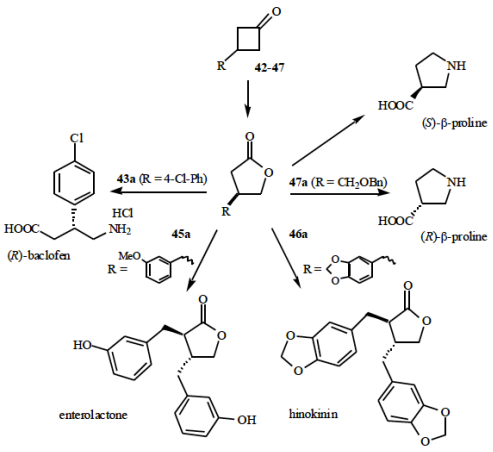


Figure 1.13: Biooxidation to butyrolactones of relevance in total synthesis.

Kinetic resolution is another relevant application of BVMOs. Starting from a racemate material and a chiral agent, a kinetic preference of one enantiomer over to the other is achieved, the 50% of the starting material is converted to the enantiopure product and the other half is lost. Several synthetically interesting substrates derived from functionalized cycloketones were used in kinetic resolution and the corresponding products were used as key-precursor for another transformation (Fig. 1.14).

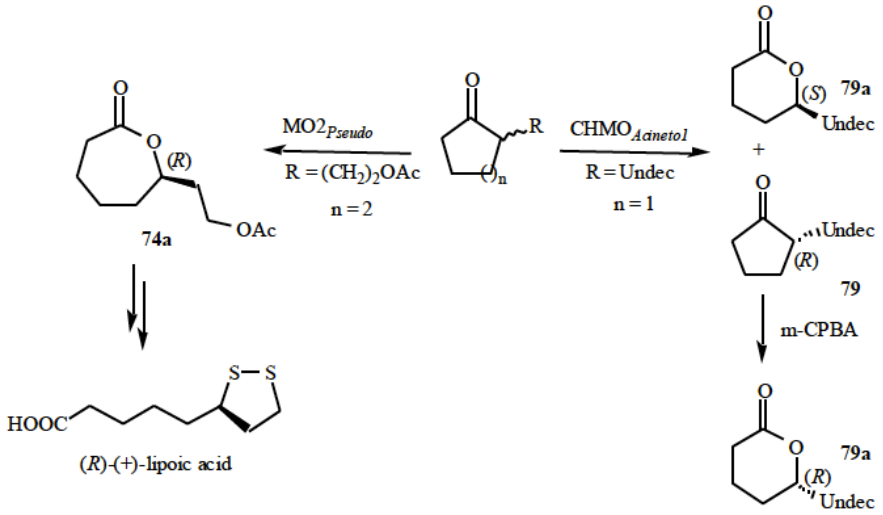


Figure 1.14: Application of chiral compounds obtained upon kinetic resolution of α-substituted ketones in bioactive compound synthesis.

An improvement to kinetic resolution in term of product yield is the development of a dynamic process. In dynamic kinetic resolution (DKR) epimerization and resolution occur simultaneously leading to 100% yield of the desired compound. The first example of DKR by BVMOs was for a benzyloxy-ketone in which the substrate was racemized at elevated pH within a whole-cell biotransformation giving 86% yield of (R)-lactone in 96% *e.e.* [11].

Another appealing synthetic application is the radiodivergent BVMO-mediated biotrasformation. In the conversion of racemic bicyclo[3.2.0]hept-2-en-6-one four possible products are obtainable (Fig. 1.16), two regioisomeric lactones, the expected "normal" and the unexpected "abnormal" according to the migratory preference of the higher substituted carbon center and two enantiomers of each.

The bioconversion of such racemic substrate by CHMO from *Acinetobacter*

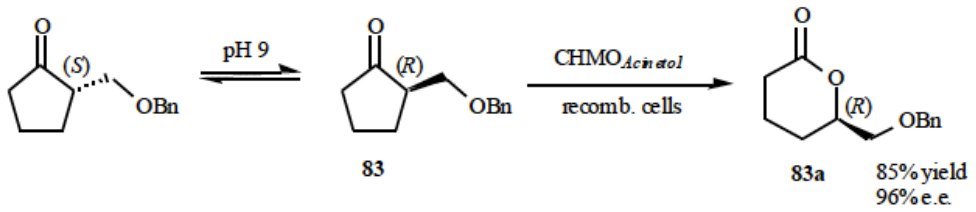


Figure 1.15: Dynamic kinetic resolution of ketone 83 by recombinant whole cells.

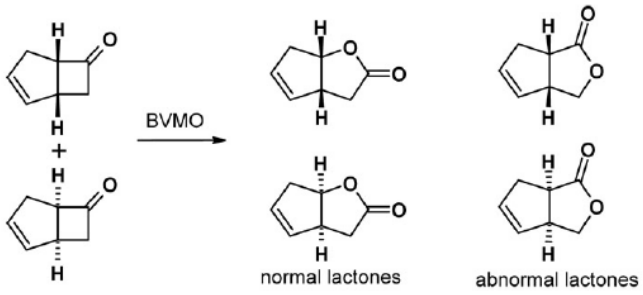


Figure 1.16: Radiodivergent biooxidation of racemic bicyclo[3.2.0]hept-2-en-6-one.

gives an equimolar mixture of "normal" as well as "abnormal" lactone in a excellent enantiomeric excess, the "normal" one is used as a precursor for a building block in the synthesis of prostaglandins [1]. The "abnormal" regioisomer is used as a starting material for the chemoenzymatic synthesis of a series of brown algal pheromones (multifidene, viridene, caudoxirene) and for the cytostatic drug (R)-(–)-sarkomycin. *C. echniulata* NRRL 3655 only produces the latter unexpected lactone (ratio 0:1) in a kinetic resolution process [61, 62]. The reason of these radiodivergence is probably due to the different position in the active site of each antipodal substrates and then different groups are situated antiperiplanar to the O-O bond in Criegee intermediate [73]. As for desymmetriation reactions several BVMOs were compared in the biooxidation of fused bicycles suggesting the same clustering; CHMO-type produces an equimolar amount of the two antipodal lactones in high enantioselectivity, CPMO-type yields mainly the expected "normal" lactones with low enantioselectivity.

Bayer-Villiger monooxygenases are also able to perform oxidation of heteroatoms (sulfur, nitrogen, phosphorus, boron, selenium) and even epoxidation reactions [84, 20, 15]. In particular asymmetric sulfoxidation seems to be the most appealing because chiral sulfoxides are valuable building block for the syn-

thesis of biologically active compounds. PAMO and others type I BVMOs (e.g. CHMO, HAPMO) have shown high enantioselectivity towards alkyl aryl sulphides, dialkyl sulphides and cyclic and acyclic 1,3 dithioacetals.

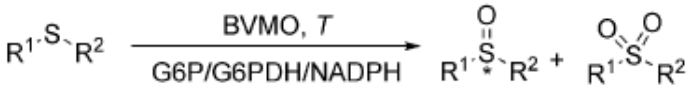


Figure 1.17: Sulfoxidations BVMO-catalyzed.

We have already mentioned some example of pharmaceutical active ingredients and intermediates available by Baeyer-Villiger monooxygenases, such as production of (R)-lipoic acid and or the synthesis of the cytostatic drug (R)-(–)-sarkomycin. Another important potential application of BVMOs is the synthesis of β-amino acids used as building block for the preparation of β-peptides, alkaloids, terpenoids and β-lactam antibiotics. The reaction involves is the oxidation of linear aliphatic ketones possessing an amino group in the β position. In a kinetic resolution process, amino β-ketones were used as racemic substrates leading to an enriched N-protected β-amino ester, then hydrolysed to obtain optically pure N-protected β-amino acid and N-protected β-amino alcohol depending on the regioselectivity of the biocatalyst [87, 88].

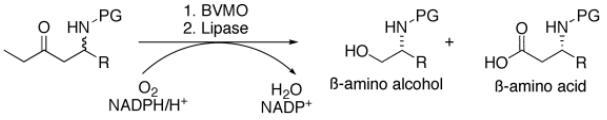


Figure 1.18: Enzymatic synthesis of β-leucine precursor.

1.4 Goals and Outline of the Thesis

The Baeyer-Villiger reaction, consisting in the production of esters or lactones starting from ketones, is considered a powerful oxidation in organic synthesis because of the cleavage of C-C bonds with the concomitant insertion of an oxygen atom, which is since a long time a big challenge for chemists. The classical approach to perform this reaction involves the use of hazardous oxidizing agents such as peroxyacids, very reactive compounds lacking the high chemo-,

region- and enantioselectivity expected in organic synthesis. Biocatalysis represents an alternative tool to carry out chemical transformations with an environmental friendly methodology. In this context, Baeyer-Villiger monooxygenases are "green" catalysts able to catalyse the Baeyer-Villiger reaction as natural equivalents of peroxyacids. BVMOs are flavin-dependent monooxygenases using NAD(P)H as electron donor and not-covalently bound FMN and FAD as prosthetic group. In particular Type I BVMOs, which use NADPH and FAD as cofactors and are the best widespread and characterized enzymes, seems to have great potential for synthetic applications. In general BVMOs have been identified in a large number of bacteria and fungi but only a restricted number of them are available in recombinant form.

The goal of the project was discover novel type I BVMOs to enlarge the toolbox of new recombinant biocatalysts able to perform oxidation reaction towards molecule of potential pharmaceutical interest. In order to discover novel type I BVMOs, new sources of enzymes have been explored.

A report on known Baeyer-Villiger Monooxygenases as recombinant biocatalysts is described in **Chapter 2**. This chapter provides an overview concerning the occurrence and metabolic role of known BVMOs in nature and the discovery of new type I BVMOs from unusual sources using genome-mining approach. In fact we have identified and selected five new putative BVMOs from different organisms: *Oryza sativa* (Os; plant), *Physcomitrella patens* (Pp; moss), *Cyanidioschyzon merolae* (Cm; red alga), *Trichodesmium erythraeum* (Te; cyanobacterium), *Haloterrigena turkmenica* (Ht; archeabacterium). In particular, basing on most updated literature, the photosynthetic eukaryotes *Oryza, Physcomitrella* and *Cyanidioschyzon* would result very uncommon sources for BVMOs.

In **Chapter 3** is reported a detailed description concerning the cloning and the heterologous protein expression in *E. coli* used for each putative sequence. The low solubility of the expressed proteins was the main bottleneck of the work; many strategies were adopted to obtain some soluble proteins. Addition of riboflavin to growth medium was successful to obtain soluble flavoproteins from *Physcomitrella patens* and *Cyanidioschyzon merolae*, showing the importance of the flavin cofactor in the folding process.

Chapter 4 illustrates the biocatalytic characterization of the two flavopro-

teins. It has been carried out at the Biotransformation and Biocatalysis group of Prof. Dr. M.W. Fraaije (University of Groningen, The Netherlands), a main expert concerning flavoenzymes and in particular in Baeyer-Villiger monooxygenases. Optimization of reaction conditions (pH and Temperature optimum and stability) was determined and after a substrate screening assays, a detailed steady-state study on the three enzymes was reported.

In **Chapter 5** the work described in this thesis is summarized, discussed and some future perspectives are mentioned.

Chapter 2

Discovering New BVMOs

Abstract:
Baeyer-Villiger monooxygenases (BVMOs) are an interesting and versatile class of flavoenzymes NADPH-dependent able to catalyse a remarkable wide variety of oxidative reactions, which are difficult to be obtained chemically [35, 73]. Up to the last ten years, a few BVMOs have been cloned and overexpressed. The number of recombinant BVMOs is considerably increased because of the vast genetic information available from genome sequencing. In particular, genome mining has proven to be an efficient approach to discover new biocatalysts. This chapter provides an overview concerning the occurrence and metabolic role of known BVMOs in nature and the discovery of new ones from unusual sources.

2.1 Occurrence and metabolic role of known BVMOs

The first BVMO purified was in 1965 by Conrad H.E. et al., while studying the camphor lactonization system occurring in *Pseudomonas* sp. [26]. However, the first indication of a Bayer-Villiger oxidation activity was found in the late 1940s: it was involved in the microbiological degradation of steroids, in particular the conversion of progesterone to testolactone in *Penicillum* species and *Aspergiullus flavus* [109]. Hereafter, it has been found that many BVMOs play an important role in the degradative metabolism of bacteria and fungi. The first discovered enzymes have been the cyclohexanone monooxygenases (CHMO) from *Acinetobacter calcoaceticu* and *Pseudomonas putida*. Within a few years several from then new type I BVMOs were overexpressed in *E. coli* and characterized (Tab. 2.1). They included several cyclohexanone monooxygenases (CHMOs), a steroid monooxygenase (STMO) acting on steroids, a cyclopentanone monooxygenase (CPMO) that modifies smaller ring systems, a cyclododecanone monooxygenase (CDMO) and cyclopendadecanone monooxygenases (CPDMO) acting on larger cyclic compounds. Additionally, 4-hydroxyacetophenone monooxygenases (HAPMO) and phenylacetone monooxygenases (PAMO) accept aromatic ketones and phenylacetone derivatives.

2.1.1 Cyclohexanone monooxygenase (CHMO)

In 1976 Trudgill et al. reported the isolation and characterization of the Type I BVMO cyclohexanone monooxygenase (CHMO) from *Acinetobacter* sp. NCBI 9871 (E.C. 1.14.13.22), responsible for the biodegradation of cyclohexanol to cyclohexanone (Fig. 2.1). Besides cyclohexanone, the enzyme was able to catalyze the oxidation of a vast variety of cyclic ketones making it potentially interesting for synthetic purposes [108]. The substrates list included cyclic, bicyclic and heterocyclic ketones with a lot of substituents [75]. The extraordinary characteristic of CHMO is its high chemo-, regio-, enantioselectivity toward more than 100 substrates; it also oxidizes enantioselectively a wide range of sulfides [**?**], tertiary, secondary and hydroxyl amines [24]. This enzyme was found as-

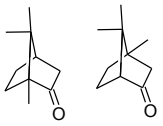
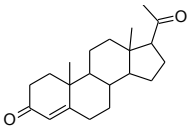
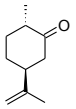
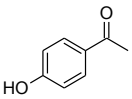
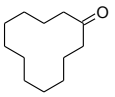
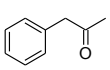
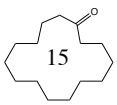
BVMO/Origin	Acronym	Primary substrate	Year of cloning
2,5- &3,6- Diketocamphane 1,2 –monooxygenase from P. *putida* ATCC 17453	2,5-&3,6-DKCMO		1986 (identified)
Cyclohexanone monooxygenase from A.*calcoaceticus* NCIMB 9871 (1) *Arthrobacter* BP2 (2),*Brachymonas petroleovorans* (3),*Rhodococcus* sp. Phi1 & Phi2 (4), *Brevibacterium* sp. HCU (5) and *Xanthobacter* sp. ZL5 (6)	CHMO		(1) 1988 (2) 2003 (3) 2003 (4) 2003 (5) 2000 (6) 2003
Steroid monooxygenase from *R.rhodochrous* IFO 3338 and *Cylindrocarpon radicicola* ATCC 11011	STMO		(1) 1999 (2) 1986 (identified)
Monocyclic monoterpene ketone Monooxygenase from *R.erythropolis* DCL14	MMKMO		2000 (identified)
4-Hydroxyacetophenone monooxygenase from and *P.flurescens* ACB	HAPMO		2001
Cyclododecanone monooxygenase from *R.ruber*	CDMO		2001
Cyclopentanone monooxygenase From *C.testosteronii* NCIMB 9872	CPMO		2002
Ethionamide monooxygenase from *Mycobacterium tuberculosis* H37Rv	EtaA		2004
Phenylacetone monooxygenase from *T. fusca*	PAMO		2005
Cyclopentadecanone monooxygenase from *Pseudomonas* sp. HI-70	CPDMO		2006

Table 2.1: Melting temperatures obtained at different pH value.

sociated with other enzymes involved in the fatty acid beta-oxidation pathway, where the Bayer-Villiger oxidation is the second step after the conversion of the cyclohexanol to cyclohexanone. *ChnC*, the gene flanking the one coding for CHMO, encodes a caprolactone hydrolase able to hydrolase the produced lactone. Such a monooxygenase-hydrolase gene localization (linked position) has been found also for others genome fragments containing BVMOs genes (i.e. cyclopentanone monooxygenases, pheylacetone monooxygenases, 4-hydroxyacetophenone monooxygenase and cyclopentadecanone genes.

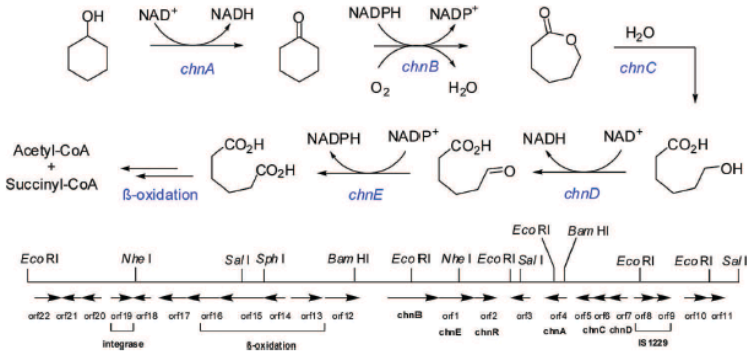


Figure 2.1: Role of CHMO from *Acinetobacter* sp. NCIMB 9871 in the catabolic degradation of cyclohexanol. (Image from Chem Rev. 2011 Jul 13;111(7):4165-222. doi: 10.1021/cr1003437)

2.1.2 Cyclopentanone monooxygenase (CPMO)

Cyclopentanone monooxygenase from *Pseudomonas* sp. NCIMB 9872, reclassified as *Comamonas* in the same year, is responsible for the degradation of cyclopentanol or cyclopentanone [50, 41]. The Baeyer-Villiger oxidation is the second step of the pathway and consists in the conversion of cyclopentanone to 5-valerolactone (Fig. 2.2).

Sequence analysis revealed a cluster of genes and six potential open reading frames grouped in 4 possible transcriptional units [50] The cpnABCDE genes encode for the five enzymes catalyzing the five steps of the conversion of cyclopentanol to glutaric acid (Fig. 2.3).

The protein sequence of CPMO is 37% identical to the CHMO of *Acinetobacter* strain NCIMB 9871, but the best score given by BLAST search was the

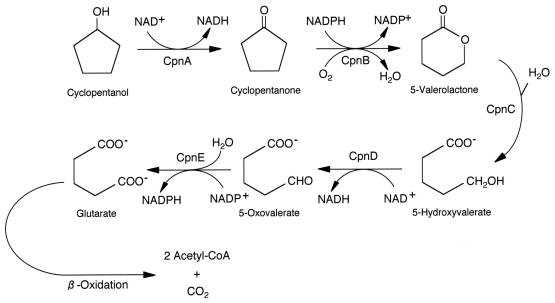


Figure 2.2: Degradation pathway for cyclopentanol by *Comamonas* sp. strain NCIMB 9872.

Figure 2.3: Cpn gene cluster organization in Comamonas sp. strain NCIMB 9872.

CHMO2 of *Brevibacterium* sp. Strain HCU, with 49,5% of identity [49]. The recombinant protein, available from 2002, is able to oxidize several cyclic ketones, but it may also transform aromatic and conjugated ketones not accepted by CHMO. Regarding its selectivity, CPMO results quite similar to CHMO, but many examples of enantio- and regiodivergent transformations were reported [113]. Comparative biocatalytic studies have revealed that CHMO from *Acinetbacter* and CPMO from *Comanonas* can be considered as models for two BVMO subclasses, the CHMO-type and CPMO-type groups [76].

2.1.3 4-Hydroxyacetophenone monooxygenase (HAPMO)

The Baeyer-Villiger oxidation has been detected also in the microbial degradation of aryl ketones, in particular ring-substituted acetophenones [29, 42]. *Pseudomonas fluorescens* ABC is able to grow on 4-hydroxyacetophenone as carbon source [43]: it produces a Baeyer-Villiger monooxygenases (HAPMO) that was isolated and purified in 1999,. The gene cluster hapCDEFGHIBA of this organism, involved in the 4-hydroxyacetophenone degradation, was cloned and analysed: it encodes five enzymes and others regulatory proteins. The first two reactions of the whole transformation are catalyzed by the monooxygenases (HapE) and an esterase (HapD); the formed hydroquinone is cleaved by a dioxygenase

(HapC) producing a 3-hydroxymuconic semialdehyde, oxidized to maleylacetate by HpA and then reduced to β-oxoadipate by HapB (Fig. 2.4).

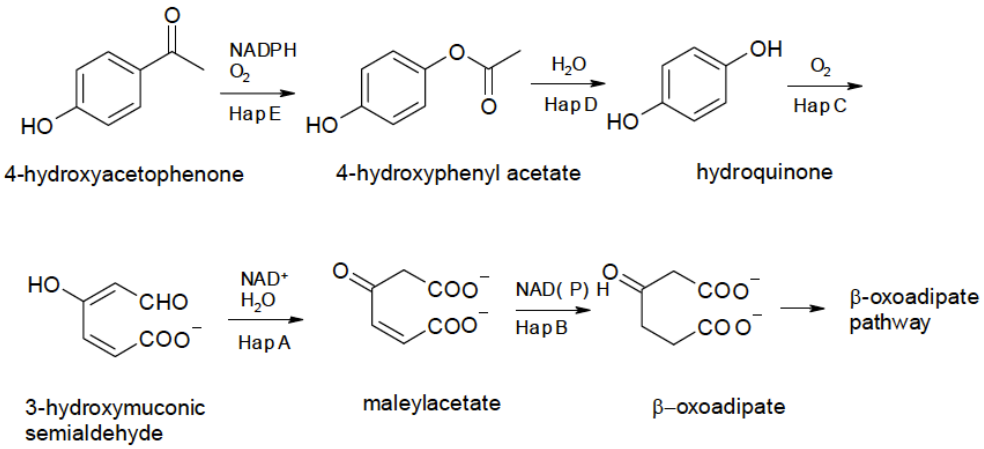


Figure 2.4: Baeyer-Villiger reaction observed during microbial degradation of aromatic compounds in *Pseudomonas fluorescens* ABC.

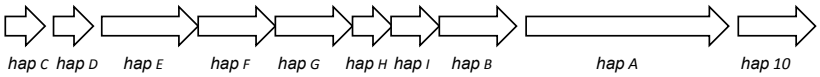


Figure 2.5: Organization of hap gene cluster of P. *fluorescens*.

4-hydroxyacetophenone monooxygenase (hapA), 4-hydroxyphenyl acetate hydrolase (hapB), HQDO (HapD), 4-hydroxymuconic semialdehyde dehydrogenase (hapE), and maleylacetate reductase (hapF) genes. The enzyme HAPMO, overexpressed in *E. coli* and characterized in 2001, has great substrate specificity and accepts heteroaromatic ketones, aliphatic ketones and sulfides [55]. Moreover, HPMO is a homodimer strictly dependent on NADPH and structurally different from others type I BVMO because of the presence of a 135 amino acid N-terminal extension playing an important role in stabilizing the HAMPO dimer. A HAPMO-homologue from *Pseudomonas putida* JD1, an organism able to grow on 4-etyilphenol which is degraded via 4-hydroxyacetophenone. was purified and characterized in 2000 [103].

2.1.4 Steroid monooxygenase (STMO)

As previously mentioned, the first indication of a Bayer-Villiger oxidation activity was found in the microbiological degradation of steroids (Fig 2.6). The steroid monooxygenases from the bacterium *Rhodococcus rhodochrous* catalyze the oxidation of progesterone side-chain to the testosterone acetate. The unique enzyme from the fungus *Cylindrocarpon ridicicola* ATCC 1101, is able to perform the bifunctional oxidation of progesterone to testosterone acetate and of androstenedione to testolactone.

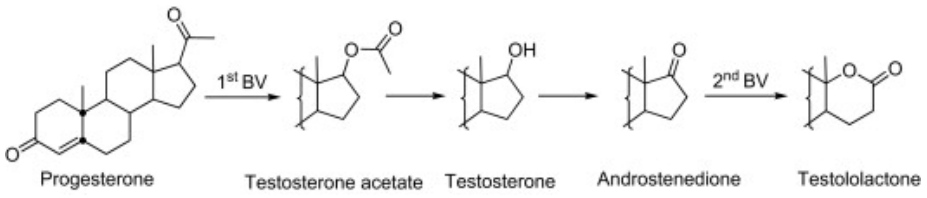


Figure 2.6: Metabolic pathway of progesterone in fungi.

This enzyme shows a high sequence homology with the well-known BVMOs, 53% sequence identity with PAMO from *Thermobifida fusca*, 45% with CHMO from *Rhodococcus* sp. strain HI-31, and 43% with OTEMO from *Pseudomonas putida*. It has been expressed in recombinant form in *E. coli* and has exhibited a diverse substrate preference ranging from a simple aromatic (phenylacetone) to a steroid molecule [38].

2.1.5 Cyclododecanone monooxygenase (CDMO)

The enzyme cyclododecanone monooxygenase was found in a strain able to degrade large cyclic hydrocarbons as Rhodococcus ruber CD4 and was purified and characterized. The coding gene (*cddA*) from *Rhodococcus ruber* SC1 was cloned and expressed in *E. coli*, resulting able to convert his natural substrate cyclododecanone into lauryl lactone and more in general C11-C15 cyclic ketones (Fig. 2.7).

The gene is grouped in a cluster with five other genes (Fig. 2.8): the first two genes, *cddA* and *cddB*, encode the enzymes catalysing the first two steps of the cyclododecanone oxidation. The last two genes, *cddC* and *cddD*, encode the two dehydrogenases necessary for oxidation of 12-hydroxylauric acid to generate

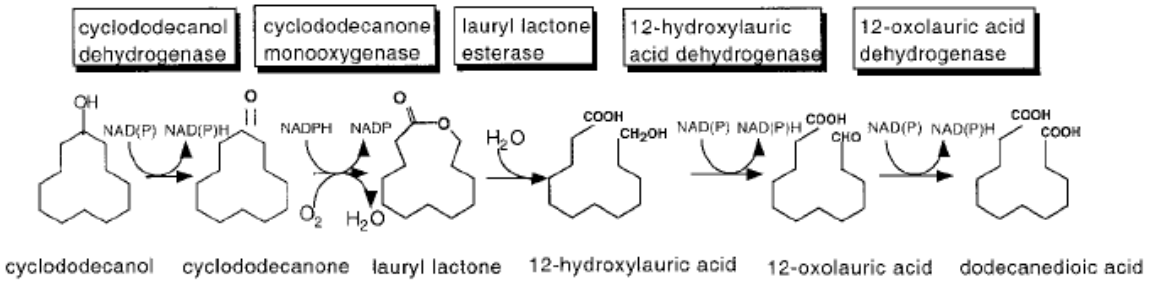


Figure 2.7: Biochemical degradation of cyclododecanone by *Rhodococcus ruber* SC1.

dodecanedionic acid. The two central open reading frames (ORFS), *cddX* and *cddY*, encode proteins of unknown function [58].

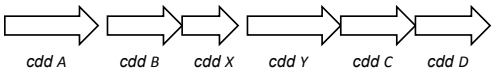


Figure 2.8: *cdd* gene cluster organization in *Rhodococcus ruber* SC1.

2.1.6 Phenylacetone monooxygenase (PAMO)

Differently from others BVMOs obtained by culturing a specific microorganism, the termostable phenylacetone monooxygenases (PAMO), from the termophilic actinomycete *Termobifida fusca*, was discovered by genome mining using the described "fingerprint" motif characterizing type I BVMOs. Two putative BVMO genes were identified, but only one of them shows relatively high sequence identities with known BVMOs, 53% with steroid monooxygenases and 40% with cyclohexanone monooxygenases. The purified recombinant enzyme was able to oxidize a wide range of carbonylic compounds in particular aromatic and aliphatic ketones and organic sulphides [37]. By genome analysis, genes flanking the *pamo* gene sequence were also investigated; they code for an alcohol dehydrogenase, a reductase and an esterase. The latter enzyme, in combination with the others simultaneously expressed, suggests their involvement in some degradation pathway, as for many other BVMOs. PAMO is also the first BVMO for which the X-ray structure has been solved [69].

2.1.7 Cyclopentadecanone monooxygenase (CPDMO)

Another enzyme able to accept large (C15) and small (C6) cyclic and bicyclic ketones is cyclopentadecanone monooxygenase. Naturally derived from *Pseudomonas* sp. strain HI-70, it has been expressed in *E. coli*. In 2009 Beneventi et al. showed that it was able to oxidize numerous 3 and 17-ketosteroids with a full control of the regiochemistry of the produced lactones, enlarging the substrate scope of the CPDMO to more demanding molecules such as steroids [10]. In the genome of *Pseudomonas* sp. strain HI-70, the gene coding for cyclopentadecanone monooxygenases (*cpdB*), is located in between two genes, encoding a transcriptional regulator (*cpdR*) and, a lactone hydrolase (*cpdC*) (Fig. 2.9).

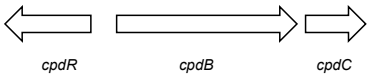


Figure 2.9: *cpdB* gene localization in *Pseudomonas* sp. strain HI-70.

2.1.8 Ethionamide monooxygenase (EtaA)

Ethionamide is a thioamide precursor of antitubercular prodrugs and having a lethal effect if oxidatively activated by the protein EtaA, a type I Baeyer-Villiger monooxygenase acting as a prodrug activator and found in *Mycobacterium tuberculosis*. The EtaA gene sequence contains the "fingerprint" motif typical of BVMOs and the closest homolog is 4-hydroxyacetophenone monooxygenase (HAPMO), with which it displays 24% of identity. The enzyme converts a wide range of ketones, in particular long-chain ketones (e.g. 2-hexanone and 2-dodecanone), but also aromatic ketones (phenylacetone and benzylacetone) are converted. EtaA is also able to catalyze enantioselective sulfoxidation of methyl-ptolylsulfide [36].

2.1.9 Camphor degradation (2,5-3,6-DKCMO and OTEMO)

The first BVMO identified in the bio oxidative degradation of camphor (Fig. 2.9) was the 2,5-diketocamphane 1,2-monooxygenase (2,5-DKCMO). After an oxida-

tive cascade drived by P450 complex, the (+)-camphor (+1) is transformed into the 2,5-diketocamphane (3), oxidized by the 2,5-diketocamphane monooxygenase (a) (2,5-DKCMO), an NADH dependent enzyme containing FMN [106]. (–)-Camphor (–1) is metabolized by the isofunctional enzyme 3,6-diketocamphane 1,6-monooxygenase (b) (3,6-DKCMO) [52]. Both proteins are type II BVMOs.

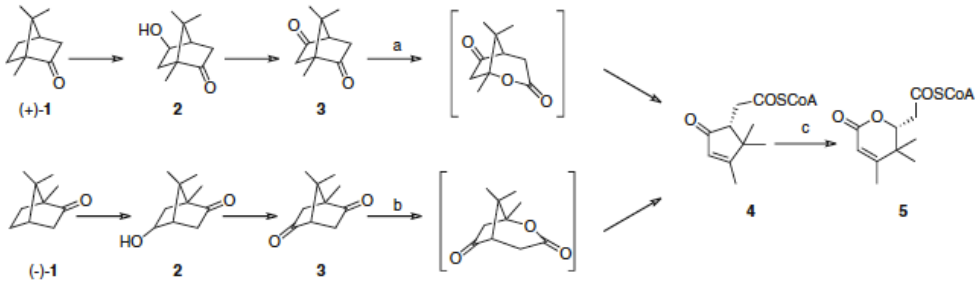


Figure 2.10: Degradative pathway of camphor.

Both resulting lactones formed are instable and undergo spontaneous formation of 2-oxo-Δ3-4,5,5-trimethylcylopentenylaceticacid, which Co-A activated (4) is the substrate of the 2-oxo-Δ3-4,5,5-trimethylcylopentenylacetyl-CoA monooxygenase (c)(OTEMO), a type I BVMO.

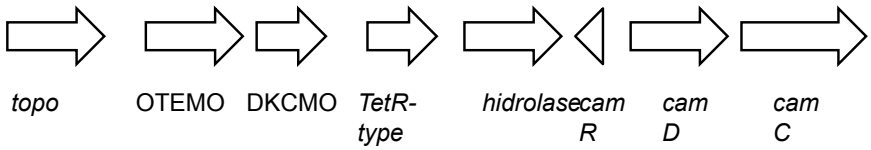


Figure 2.11: Localizations of the OTEMO-encoding gene in *P. putida* ATCC 17453. The identified open reading frames are as follows, from left to right: putative DNA topoisomerase III (topo), OTEMO, 2,5-diketocamphane monooxygenase (DKCMO), aTetR-type regulator, lactone hydrolase, and a camR repressor.

2.1.10 Monocyclic monoterpene ketone monooxygenase (MMKMO)

In contrast to the bicyclic monoterpene camphor degradation pathway, which involves three BVMOs, monocyclic monoterpene ketone monooxygenase (MMKMO)

is the only type I BVMO purified from *Rhodococcus erythropolis* DCL14, implicated in degradation pathways of many monocyclic monoterpenes, i.e. limonene, (dihydro)carveol and menthol [31].

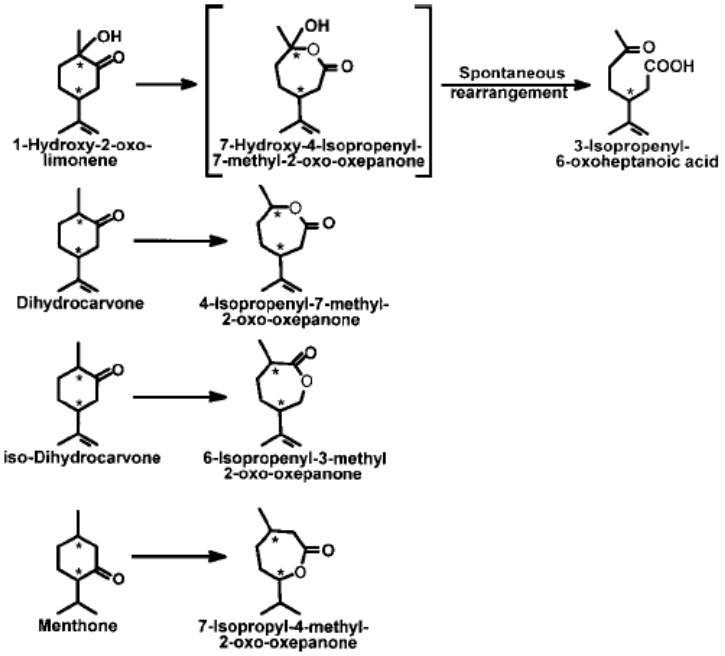


Figure 2.12: Reactions catalysed by MMKMO from *R. erythropolis* DCL14.

MMKMO has a broad substrate specificity, catalysing the lactonization of a large number of monocyclic monoterpene ketones and substituted cyclohexanones; aldehydes and acyclic ketones did not act as substrates. The regiospecificity of the enzyme depends on the diasteroisomer used as the substrate (compare dihydrocarvone with iso-dihydrocarvone) (Fig. 2.12).

2.1.11 Exploring novel putative type I BVMOs

The classical approach for discovering new enzymes has been the culture of microorganisms showing a specific activity, and then the rescue of the gene of interest through cloning. This time-consuming method was used for many years and has been successful to obtain most of the best-known BVMOs. However, up to 99% of bacteria in the environment are not readily cultured in the lab [**?**],

especially extremophile microorganisms. In this way, most of the novel potential enzyme activities result unavailable.

Advances in molecular biology techniques and genomics have laid the basis for the development of other interesting approaches for enzyme discovery. In fact, the rapid increasing of sequence information in public databases and the parallel development of bioinformatics have opened the route to the so called "*in silico* genome mining" approach. By homology alignment a BLAST-search can find conserved regions between sequences. Well-known enzymes with desired properties, preliminarily chosen as probes, are needed. Usually the sequences with moderate identity are selected as candidates [65]. In earlier studies searches performed by genome mining showed that type I BVMOs were present in a vast variety of bacteria and fungi, but not in Archea, plants or human genomes [54].

In order to enlarge the possibility to find out novel and promising biocatalysts, we used the genome mining approach to "hunt" BVMOs within specific and unusual organisms. Using as template PAMO and the "fingerprint" motif of type I BVMOs (FxGxxxHxxxWP/D) as discriminant, we have identified and selected five new putative BVMOs from five different organisms: *Oryza sativa* (Os; plant), *Physcomitrella patens* (Pp; moss), *Cyanidioschyzon merolae* (Cm; red alga), *Trichodesmium erythraeum* (Te; cyanobacterium), *Haloterrigena turkmenica* (Ht; archeabacterium). In particular, basing on most updated literature, the photosynthetic eukaryotes *Oryza*, *Physcomitrella* and *Cyanidioschyzon* would result very uncommon sources for BVMOs.

2.1.12 New putative type I BVMOs from prokaryotic organisms

It has been known for long time that many bacteria and fungi produce Baeyer-Villiger enzymes but until 1990s only few of these enzymes were isolated, purified and characterized. [101] Today, almost all of the available recombinant BVMOs are type I enzymes that require NADPH and FAD as cofactors. Thanks to the identification by genome mining, the collection of BVMOs is growing fast. In particular, we have identified two new putative BVMOs from recently sequenced prokaryotic organisms, *Trichodesmium erythraeum* and *Haloterrigena*

turkmenica.

Trichodesmium erythraeum IMS101 is a filamentous cyanobacterium and is a species of Trichodesmium genus also called "sea sawdust", founded in nutrient poor tropical and subtropical ocean waters. It fixes atmospheric nitrogen into ammonium using photosynthesis [19]. The complete sequence of its genome appeared in 2006 (http://www.jgi.doe.gov). The primary structure of the protein annotated as "probable flavin-binding monooxygenase" (protein_id=ABG52868.1, locus_tag=Tery_3824) is consists of 493 residues with 43% sequence identity to PAMO and 37% to CHMO.

Haloterrigena turkmenica DSM 5511 is an extremely halophilic archeon isolated from a sulphate saline soil in Turkmenistan. The plasmid pHTUR01 encodes a protein "similar to cyclohexanone monooxygenase" (protein_id=ADB62698.1, locus_tag=Htur_3838): its 554 residues long sequence is 50% identical to PAMO and 43% to CHMO from *Acinetobacter*. No BVMOs have been identified so far in archea genomes.

2.1.13 New putative type I BVMOs from eukaryotic organisms

Up to last year, all the available recombinant BVMOs were from prokaryotic origin. The first eukaryotic BVMO cloned and expressed in *E. coli* is cycloalkalone monooxygenases from the ascomycete *Clyndrocarpon radicicola* ATCC 11011 (identical to *Cylindrocarpon destructans* DSM 837). This enzyme is able to convert progesterone *via* androstenedione towards Δ1-dehydrotestololactone [63]. Before the publication of this very recent discovery, we had begun our research with the aim of identifying new BVMOs. In particular, we had identified by genome mining and then selected new putative enzymes from three photosynthetic eukaryotes: *Cyanidioschizon merolae*, *Phiscomitrella patens* and *Oryza sativa*. The complete genomic sequences of the two former organisms were respectively released in 2004 [70] and 2008 [89]. For the latter organism, a very wide collection of cDNA sequences was deposited in 2003 [39]. The primitive red algae *Cyanidioschizon merolae* is a unicellular organism living in acidic environments (like hot springs), even at pH < 2 and temperature of 45°C. The

putative enzyme is coded by an open reading frame present in chromosome 12 of the organism genome and its primary structure has 43% sequence identity to PAMO.

Phiscomitrella patens is a moss, classified as a no vascular plant because it differs from "higher" plants by not having internal vessels. From an evolutionary point of view, *P. patens* is situated between water and land plants: thus, it has become a model organism for studies on plant evolution and developmental physiology. In addition, it is increasingly used in biotechnology since it can be efficiently transformed and is able to integrate exogenous DNA by homologous recombination [94]. The 536-residue primary structure of the identified hypothetical protein from *P. patens* reveals 54% sequence identity to PAMO and 37% to CHMO *Rhodococcus* sp. Phi1. Differently from those, however, the BVMO "fingerprint" motif is not perfecltly respected in the protein from the moss, being FxGxxxYxxxWP instead of FxGxxxHxxxWP. It is worth noting that the central histidine of the motif (His-173 in PAMO) is crucial for both FAD binding and catalysis.

Oryza sativa is a plant commonly known as Asian rice. It is relatively easy to be genetically modified and it is used as model organism for cereal biology. The putative enzyme that we selected correspond to the translation of a full-length cDNA clone from *Oryza sativa* Japonica Group rice (GenBank: AK107233.1): its amino acid sequence reveals 39% identity to PAMO and 32% to CHMO.

Chapter 3

Cloning and Production of New Putative Type I BVMOs

Abstract:
Type I Bayer-Villiger monooxygenases are FAD-containing, NADPH dependent oxidative enzymes sharing, besides common biochemical properties, the conserved "fingerprint" motif FxGxxxHxxxWP/D. This BVMO-specific sequence was used to compare BVMO sequences in genomic databases trying to found out novel BVMOs. Till now, all identified BVMO genes come from bacteria or fungi, but not from Archea, plant or human genomes. In order to enlarge the possibility to find out new and promising biocatalysts, we used the genome mining approach to "hunt" BVMOs within genomes of unusual organisms. By this way we have identified and selected five new putative BVMOs from five different organisms: *Oryza sativa* (Os; plant), *Physcomitrella patens* (Pp; moss), *Cyanidioschyzon merolae* (Cm; red alga), *Trichodesmium erythraeum* (Te; cyanobacterium), *Haloterrigena turkmenica* (Ht; archibacterium). In particular, basing on most updated literature, the photosynthetic eukaryotes *Oryza*, *Physcomitrella* and *Cyanidioschyzon* would result very uncommon sources for BVMOs. This chapter gives a detailed description concerning cloning and heterologous protein expression in *E. coli* used for each putative enzyme.

3.1 Cloning and expression strategy

To obtain the desired putative recombinant proteins, a suitable expression system had to be generated. To facilitate this work, we decided to use the same cloning strategy for all the five sequences. The expression vectors were created by digestion of pET-28a(+) with *NcoI/NotI* and ligation of the amplified Ht-BVMO, Te-BVMO, Os-BVMO, Cm-BVMO, Pp-BVMO sequences, cut by the same enzymes. These sequences were obtained by two subsequent PCR reactions, the first one producing a preliminary amplimer, which was then used for the second mutagenic amplification. The latter PCR introduce proper restriction sites through nested primers. The only difference among the five sequences was the template material: genomic DNA for Te-BVMO, Cm-BVMO, Pp-BVMO and plasmid DNA for Ht-BVMO and Os-BVMO (Fig. 3.1). Plasmid pET28a (+), which carries the T7 promoter, the kanamycin resistance gene, and permits to add a hexahistidine coding sequence tail at the cloned ORF was used as expression vector (Fig. 3.2). To transcribe the cloned sequence the T7 RNA polymerase must be present. The host cells for this expression system are usually those of the *E. coli* strain BL21 (DE3), having the gene for T7 RNA polymerase stably integrated in its genome and controlled by the inducible promoter lac. Thus, the obtained constructs were transformed in *E. coli* BL21 (DE3) and expression tests were performed: in order to select better conditions for obtaining soluble proteins, parameters such as temperature for expression, IPTG concentration and suitable strain were varied.

3.2 Os-BVMO: Putative type I BVMO from *Oryza sativa*

Oryza sativa is a plant specie (Fig. 3.3 and Fig. 3.4) commonly known as Asian rice. Is a monocot and it has become a model system of genome research for cereal plants, being favourable to genetic modification. Two sub-species of *O. sativa* have been sequenced; the sticky, short-grained japonica group usually cultivated in dry fields of East, Southeast Asia and Japan and, the nonsticky, long-grained indica group, cultivated mostly submerged in tropical Asia and

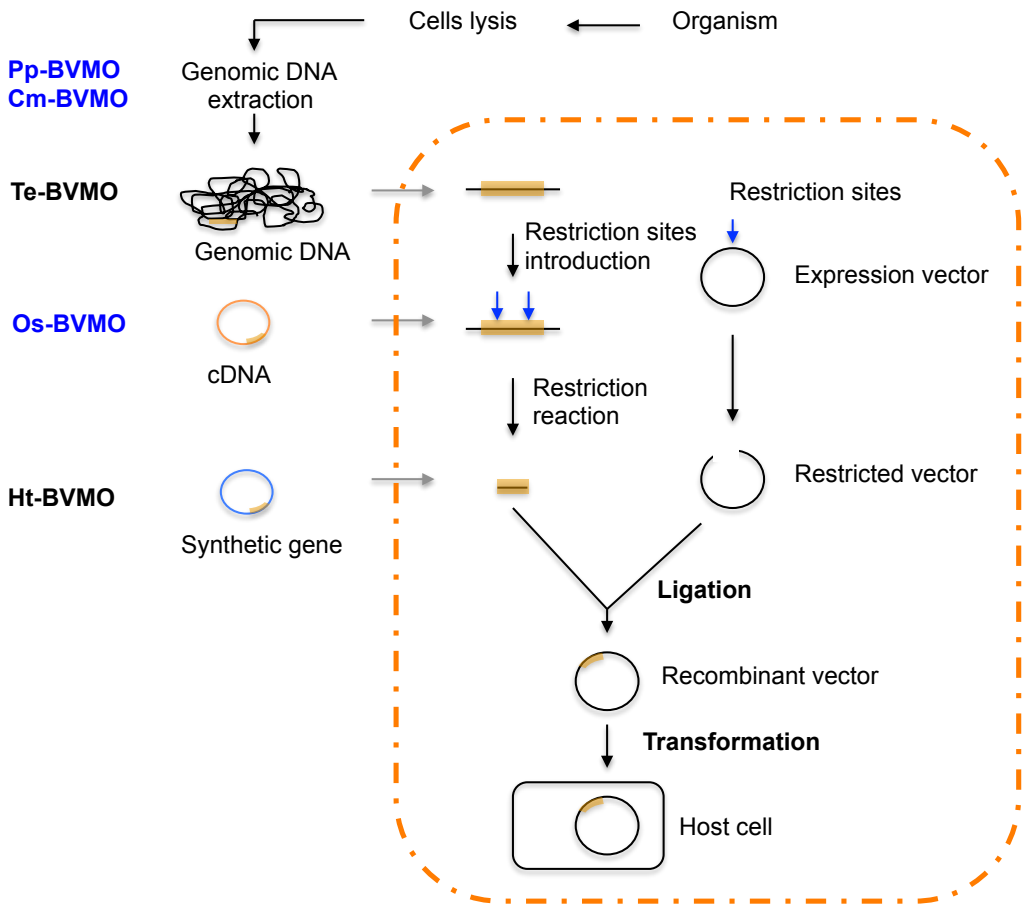


Figure 3.1: Cloning strategy scheme.

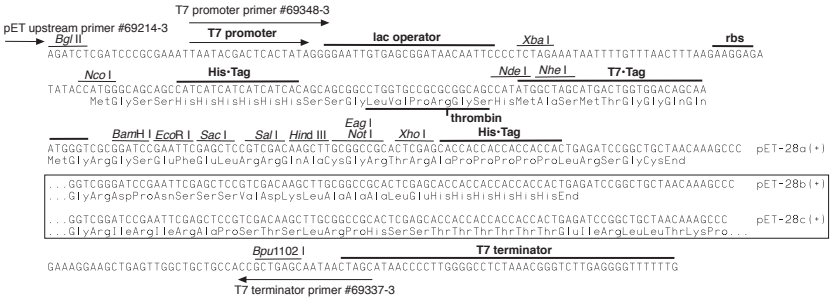


Figure 3.2: Cloning region of plasmid pET28a(+).

Philippines [82]. Another sub specie is the broad-grained javanica growing in tropical areas. The variety used in this work was the spp. japonica. The *O. sativa* L. ssp. japonica genome was sequenced in 2002 and published in 2005 by International Rice Genome Sequencing Consortium (IRGSP) because of its major importance in the world food industry. Its genome is the smallest one, estimated at just 430Mb across 12 chromosomes (IRGSP 2005).

Figure 3.3: *Oryza sativa* spp japonica (Image from Genoscope, http://www.genoscope.cns.fr).

Kingdom:	Plantae
Division:	Angiosperms
Class:	Monocots
Order:	Poales
Family:	Poaceae
Genus:	*Oryza*
Species:	*O. sativa*

Figure 3.4: Taxonomic classification of *Oryza sativa*.

3.2.1 Cloning of Os-BVMO

The cDNA of the putative gene Os-BVMO was present in a clone library (Microbial Culture Collection-NIES) of over 28,000 cDNA clones from japonica rice, collected by a research team of the National Institute of Agrobiological Sciences, Japan (Rice Full-Length cDNA Consortium, 2003). The purchased plasmid pCMVFL3 is a pCMVSport6 vector containing the cDNA of the putative BVMO. We used as DNA template a fragment of the plasmid obtained from a

first PCR reaction using the universal sp6 e T7 primers, annealing to the sequences of the plasmid flanking the Os-BVMO coding region. The 2000pb long amplicon obtained was used for a second mutagenic PCR amplification driven by the nested Os1 forward and Os2 reverse primers (Fig. 3.5). The two amplification steps and accurate optimization of annealing temperature was needed because of the number of subsequent mismatches present in the Os1 primer. To find the optimum temperature, PCR reaction tests with a temperature gradient were performed. The annealing temperature found was 52°C, quite low if compared to the theoretical one of 68°C. The PCR product obtained was the correct DNA fragment of 1641 pb (Fig. 3.5).

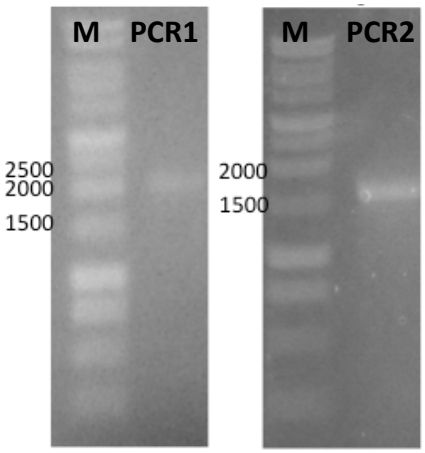


Figure 3.5: 1% agarose gel analysis of the preliminary 2000 bp and final 1641 bp PCR products (PCR1 and PCR2). M: standard marker (Promega).

Afterwards, a restriction digestion using both *NcoI/NotI* was performed to prepare the insert for the ligation into pET28a. The resulting ligation product was used to transform chemically competent XL1-Blue cells. Successful ligation was examined by colony PCR and analytical restriction digestion (Fig. 3.6). .

3.2.2 Expression of Os-BVMO

After transformation into the suitable BL21(DE3) cells, pre-cultures were grown overnight and used to inoculate a main culture for expression. Overexpression tests were performed in order to find out better conditions for expression. Different growth temperatures (18, 30 and 37°C) and different times for induction

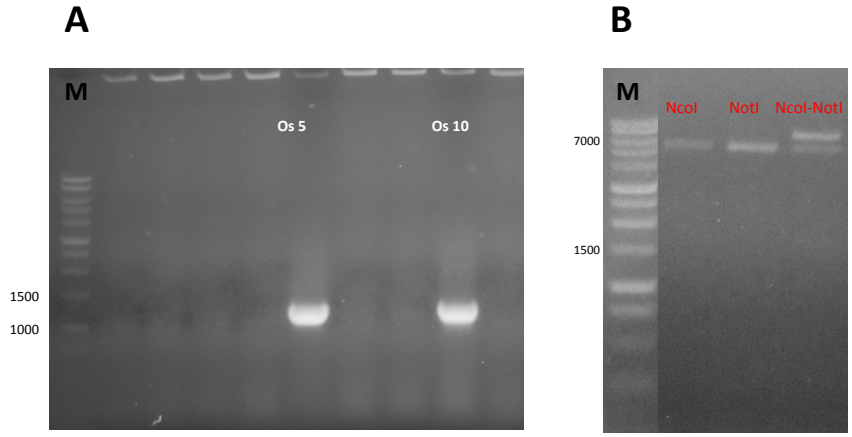


Figure 3.6: (A) 1% agarose gel analysis of amplimers produced by colony PCR. (B) 1% agarose gel analysis of a positive recombinant plasmid pET28a-Os-BVMO by restriction digestion using *NcoI/NotI*. M: standard marker (Promega).

were checked. Final cell samples were also disrupted by French press and respective homogenates were separated by centrifugation into pellet and surnatant fractions. Results of these tests were examined by SDS-PAGE, showing that the expression of the protein was easily obtained at all different induction times and different temperatures (gels not shown). The expression of the target protein increased during the time after IPTG induction, resulting in huge amount. However, at all tested condition (even at 18°C), the enzyme was produced in form of insoluble inclusion bodies, which were detected in the *E. coli* cell debris (lines p). To confirm the identity of the target protein, an immunoblotting assay has been performed using anti His-tagged antibody (Fig. 3.7).

The SDS-PAGE gel clearly points out a high basal expression of the target protein, as noticeable in the lanes referring to non-induced cell. In fact pET vectors, respect other *E. coli* based expression systems, suffer from genetic instability [85]. In the system based on strain BL21(DE3), the coding sequence for T7 RNA polymerase is located in the bacterial chromosome under the control of the inducible *LacUV5*-promoter while the coding sequence for the target protein is cloned in a pET vector under control of T7-lac promoter. In the absence of the inducer, typically isopropyl-β-D-thiogalactoside (IPTG), no (or a very small quantity of) target protein should be present. Upon addition of IPTG, the T7 RNA polymerase is expressed and transcribes the DNA downstream from the

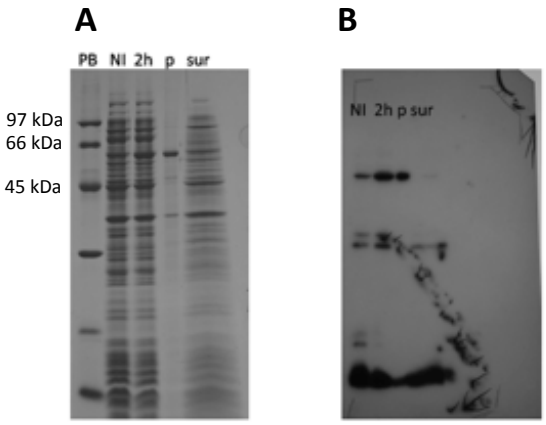


Figure 3.7: 12% SDS-PAGE analysis (A) and Western blot (B) of Os-BVMO protein expressed at 18°C. PB: protein standard marker; NI: noninduced sample; 2h: 2 hours induced sample; p: pellet fraction of induced sample; sur: soluble fraction of induced sample.

T7-lac promoter. This enzyme is so effective that a basal expression level of the desired protein can be present even in the absence of IPTG. The level of basal expression can be partially reduced using an *E. coli* strain carrying the pLysS plasmid and encoding for the T7 lysozyme an inhibitor of the T7 RNA polymerase [102]. Some attempts were made using BL21 (DE3) pLysS strain but without significant results. Many classical strategies are available to improve the solubility of an expressed protein. Decreasing the rate of protein synthesis by lowering the temperature cultivation was already considered. However, low temperature also reduces the activity of bacterial chaperonins. In order to overcome this drawback an engineered strain was used, i.e. the *E. coli* ArcticExpress strain. These cells co-express the cold-adapted chaperonins Cpn10 and Cpn60 from the psychrophilic bacterium *Oleispira Antarctica*, improving protein folding process at lower temperatures. *E. coli* ArcticExpress cultures transformed with Os-BVMO gene were grown at two different temperature, 18 and 4°C, using 0.2 mM of IPTG concentration (Fig. 3.8).

As displayed from gels shown in the figure (Fig. 3.8), neither expression in the specific host strain ArcticExpress was helpful to solve the problem since the protein remained insoluble. Changing growth medium can be useful too for protein solubility. Terrific Broth (TB) is a nutritionally rich medium [104]

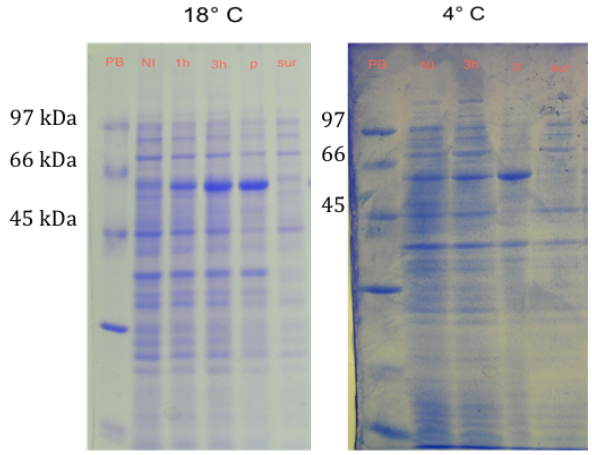


Figure 3.8: 12% SDS-PAGE analysis of cell extracts from *E. coli* ArcticExpress cells transformed with Os-BVMO and grown at 18 and 4°C, induced with IPTG 0.2mM. PB: protein standard marker (kDa); NI: noninduced sample; 1h: 1 hours induced sample; 3h: 3 hours induced sample p: pellet fraction of induced sample; sur: soluble fraction of induced sample.

commonly used for cultures in the laboratory scale. Its formulation contains increased concentrations of peptone, yeast extract, and glycerol as a carbon source providing more essential element for protein folding and stability. A culture of cells transformed with plasmid pOs-BVMO grown in TB medium was prepared and after cell disruption no soluble protein was obtained (data not showed). A final attempt in modifying growth conditions was performed adding riboflavin, the natural precursor of FAD to the medium. Also this attempt was unsuccessful. Recombinant expression in the form of fusion product is another possibility to improve the solubility of a heterologous protein. Fusion partners are characterized by their ability to enhance protein expression, reduce proteolytic degradation of the recombinant protein, and improve both protein folding and solubility. One of such fusion partners is SUMO (small ubiquitin-related modifier), a protein consisting of about 100 residues; it modulates protein structure and function by covalent modification of target proteins in eukaryotes [51, 72, 105]. Recently, a system fit for performing an efficient expression of recombinant proteins fused to SUMO in *E. coli* has been described [68]. One important feature of the SUMO-system is that the expressed fusion protein can be cleaved by the highly specific SUMO-protease, capable of generating a recombinant protein of interest with

native N-terminal amino acids. We decided to clone the SUMO gene at 5' of the monooxygenase gene for obtaining the expression of a protein having SUMO at its N-terminus. The starting material for the cloning was the genomic DNA of the yeast *Saccharomyces cerevisiae*, containing a single SUMO gene (SMT3). A first PCR reaction was performed using two specific primers (Smt3_est_forward and Smt3_est_reverse), giving a PCR product 554 pb long. The obtained DNA fragment was used as template for a second mutagenic PCR reaction, introducing the *NcoI* restriction site at both the extremities of the gene sequence (Fig. 3.9). Afterwards, a restriction digestion using *NcoI* was performed to prepare the SUMO insert (308 pb) for the ligation into the pre-digested pET28a-Os-BVMO, cut by the same restriction enzyme.

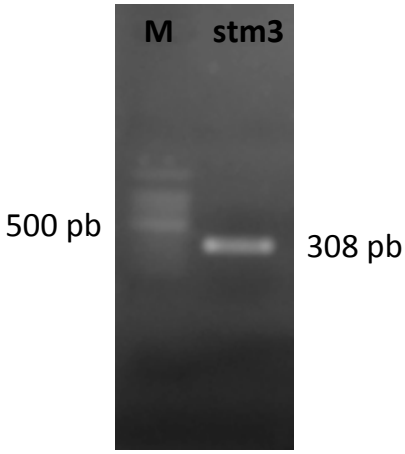


Figure 3.9: 1% agarose gel analysis of the amplimer containing the stm3 gene from *Saccharomices cerevisiae*, obtained by two-step amplification. M: standard marker (Promega).

The resulting ligation product was used to transform chemically competent XL1-Blue cells. Clones were analysed by restriction digestion and the correct ones were transformed into the suitable BL21(DE3) cells; pre-cultures were grown overnight and, on the following day, used to inoculate a main culture for expression. Also in the form of fusions, the proteins were very well expressed. Unfortunately, however, their solubility was not improved and they remained in the cell derbies after disruption.

3.3 Pp-BVMO: Putative type I BVMO from *Physcomitrella patens*

The moss *Physcomitrella patens* (Fig. 3.10) belongs to the family of Funariaceae in the order Funariales within the class Broypsida [60] (Fig. 3.11). It grows in late summer to autumn in temperate zones in an open distributed environment, often close to the waterline [60]. This species has received increasing attention in the recent years as a model organism, in particular for the study of plant evolution. Indeed, it stands in an important phylogenetic position for illuminating the evolutionary transition from the aquatic environment of algae to the terrestrial one of higher plants. Moreover, it is the only plant that undergoes homologous recombination with a frequency that allows easy targeting of genes for replacement and elimination. This, in turn, allows handy studying of gene function and helps in predicting the physiological role of genes in higher organisms. *Physcomitrella* can be easily cultivated in laboratory conditions both in liquid and solid media containing only inorganic salts. Cultivation on agar plates with rich medium (PpNH4) was used only to obtain a starting culture for genomic DNA extraction. Dr. Tomas Morosinotto (Department of Biology, University of Padova) kindly provided the starting culture.

Figure 3.10: The moss *Phiscomitrella patens* (image from Laboratoire de Génétique et Biophysique des Plantes (LGBP), www.lgbp.univ-mrs.fr).

3.3.1 Cloning of Pp-BVMO

The genomic DNA extracted from a culture of *Physcomitrella* was used as original template to obtain the putative Pp-BVMO gene sequence. A first PCR reaction,

Kingdom:	Plantae
Division:	Bryophyta
Class:	Bryopsida
Order:	Funariales
Family:	Funariaceae
Genus:	*Physcomitrella*
Species:	*P. patens*

Figure 3.11: Taxonomic classification of *Phiscomitrella patens*.

using Pp3 forward and Pp4 reverse primers, was performed to obtain a 2188 pb DNA fragment, which was employed as template for the second mutagenic PCR reaction, using Pp1 forward and Pp2 reverse primers. The desired amplimer was obtained: it was 1625 bp long and exactly contained the ORF bounded by the chosen restriction sites (Fig. 3.12).

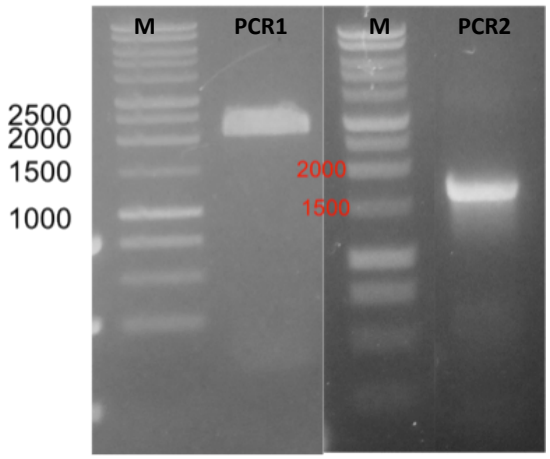


Figure 3.12: 1% agarose gel analysis of of the preliminary 2188 bp and final 1625 bp long PCR products (PCR1 and PCR2). M: standard marker (Promega).

Afterwards, a restriction digestion using both *NcoI/NotI* was performed to prepare the insert for the ligation into pET28a. The resulting ligation product was used to transform chemically competent XL1-Blue cells. Clones were analysed by PCR colony and analytical restriction digestion with *NcoI/NotI*. The Pp3 primer, internal to the target sequence, and the universal T7 were used in

the PCR colony screening getting a DNA fragment 1140 pb long (Fig. 3.13).

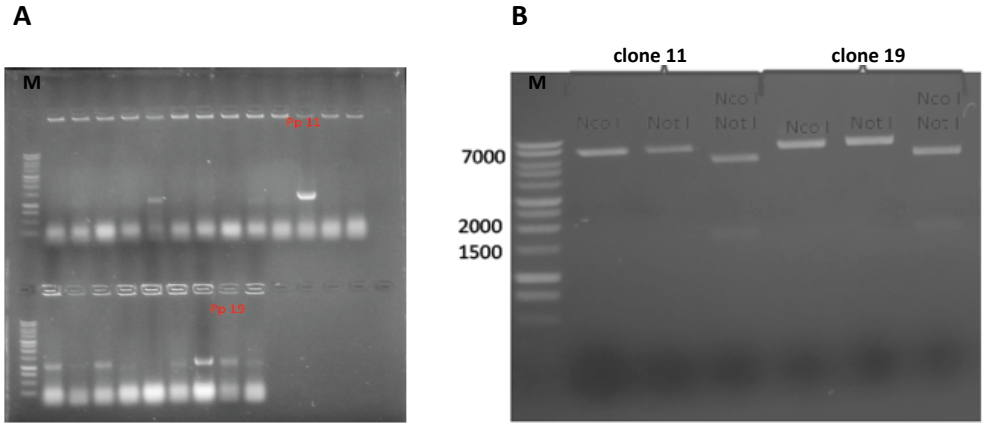


Figure 3.13: (A) 1% agarose gel analysis of bands produced by colony PCR. (B) 1% agarose gel analysis of a positive plasmids pET28a-Pp-BVMO (from clone 11 and clone 19) by restriction digestion using *NcoI/NotI*. M: standard marker (Promega).

3.3.2 Expression of Pp-BVMO

After transformation into BL21(DE3) strain, pre-cultures were grown overnight and on the following day used to inoculate cultures for expression. Overexpression tests were performed in order to define the better conditions for expression. Different growth temperatures were checked, 18, 30 and 37°C (gel not shown). An SDS-PAGE analysis was used to verify the expression of the protein at different induction time at the three temperatures and its solubility behavior after cell disruption. We observed that the expression of the target protein increased after the induction during the time at 18°C and that the enzyme was partially produced in soluble form, being detected in the protein soluble fraction too. To confirm the identity of the target protein and its partial solubility, an immunoblotting assay was performed using an anti His-tag antibody (Fig. 3.14). A first purification attempt by IMAC cromatography was performed without any success; the obtained eluted solution contained completely uncolored and aggregated protein. Because of the production of a likely mis-folded protein we decide to utilize the ArcticExpress strain, trying to improve the folding process As explained in paragraph 3.2.2 for the expression of the putative Os-BVMO,

transformed ArcticExpress cells were grown at two different temperature, 18 and 4°C and induced by different concentrations of IPTG (0.2, 0.02 and 0 mM) (Fig. 3.15). In fact, lowering temperature and inducer concentration can be helpful in reducing the speed of protein synthesis, favoring not only a correct folding process but also an improved solubility

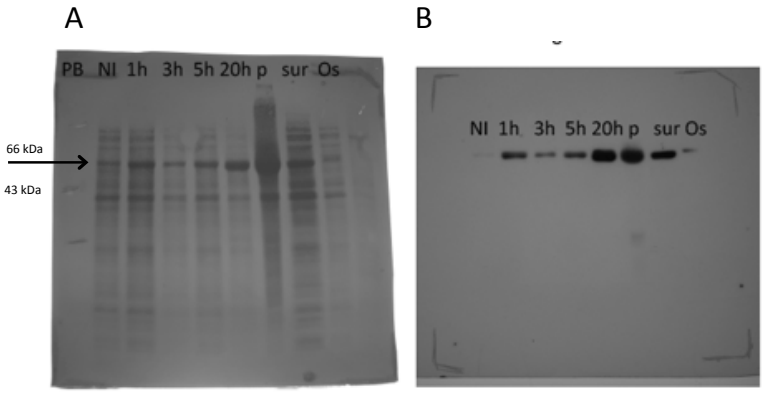


Figure 3.14: 12% SDS-PAGE analysis (A) and Western blot (B) of total extracts of bacterial cells expressing Pp-BVMO at 18°C. PB: protein standard marker; NI: noninduced sample; 1h: 1 hour induced sample; 3h: 3 hours induced sample; 5h: 5 hours induced sample; 20h: 20 hours induced sample; p: pellet fraction of induced sample; sur: soluble fraction of induced sample; Os: induced sample of Os (control sample).

As displayed from both gels (Fig. 3.15) the expressed protein partially aggregates in inclusion bodies, as shown by its presence in the pellet fractions, but is also present in soluble fraction samples. Thus, a second purification attempt was performed but during elution, non-colored fractions showing protein aggregation were collected.

After this second unsuccessful attempt we concluded that something different could be responsible for the incorrect folding of the recombinant protein. Considering how highly and rapidly the target protein was expressed (even in the absence of IPTG, Fig. 3.15), an insufficient amount of its flavin cofactor might be a possible cause. So, as we did for the putative Os-BVMO, we changed the composition of the growth medium by adding the natural precursor of FAD, riboflavin. In this case, the adjustment was successful, leading to a higher proportion of correctly folded, soluble versus unsoluble protein. Results of purification

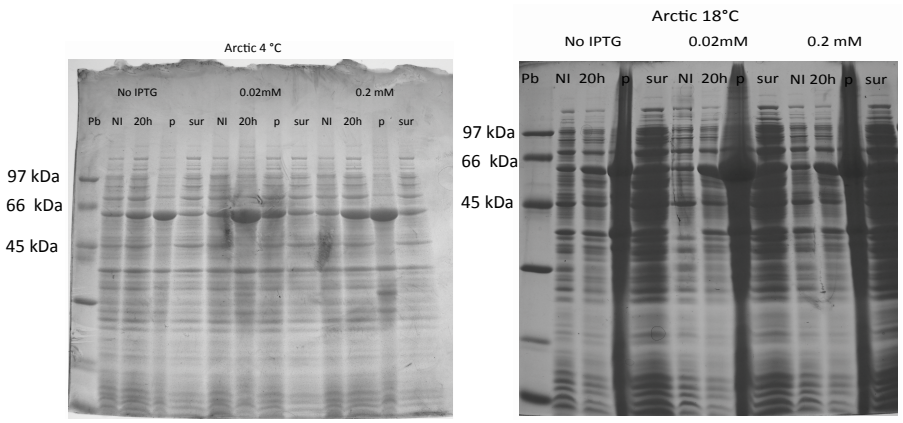


Figure 3.15: 12% SDS-PAGE analysis of cell extracts from *E.coli* ArcticExpress cells transformed with Pp-BVMO, grown at 18 and 4°C and induced at different IPTG concentrations (0, 0.2 and 0.02 mM). PB: protein standard marker; NI: noninduced sample; 20h: 20 hours induced sample; p: pellet fraction of induced sample; sur: soluble fraction of induced sample.

of soluble Pp-BVMO are reported in the next paragraph.

3.3.3 Purification of Pp-BVMO

Since the expressed recombinant protein displays a C-terminal His6-tag, immobilized metal affinity chromatography (IMAC) followed by desalting was chosen as purification method. The purification was started from the soluble crude extract, obtained by centrifugation after cell disruption by French Press. All buffers used were pre-cooled at 4°C and the fractions (crude extract or purified enzyme) were kept in ice, since the stability of the recombinant enzyme at room temperature was not known. Purification was performed by Ni-NTA chromatography on a 5 ml column using a high concentration of imidazole, as elution agent. During elution, positive fractions were identified by the bright yellow coloration (Fig. 3.16), indicating that the flavin cofactor (FAD) was bound to the enzyme [37]. The SDS-PAGE gel confirmed the identity and the presence of the putative Pp-BVMO in different fractions, before and after purification. Only elution fractions E2 and E3 were yellow colored, indicating the presence of correctly folded recombinant Pp-BVMO (Fig. 3.17). The purified solution was then desalted and finally concentrated by centrifuge filters. The evaluated concentration of the

purified protein was 158 mM.

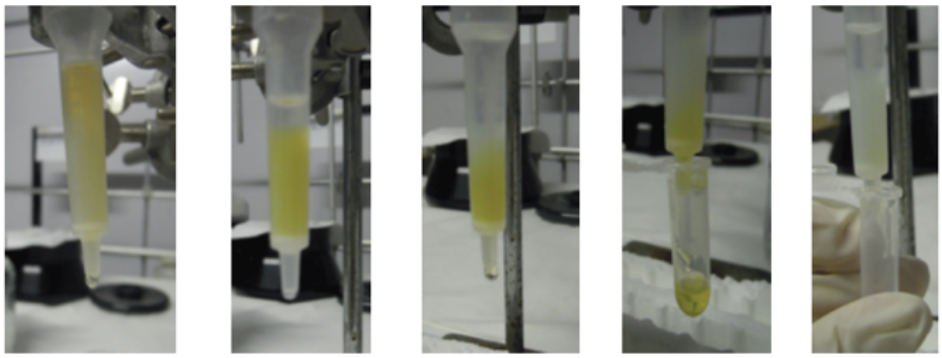


Figure 3.16: Purification steps of Pp-BVMO via immobilized metal affinity chromatography (IMAC).

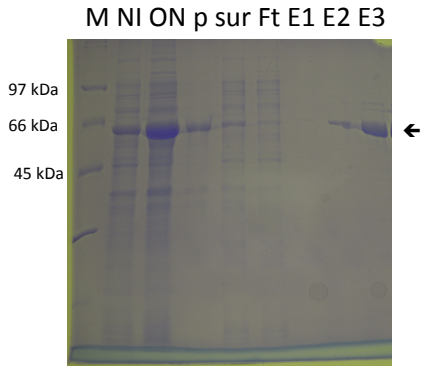


Figure 3.17: 12% SDS-PAGE analysis of samples from expression and IMAC purification of Pp-BVMO. M: protein standard marker; NI: noninduced sample; ON: over night induced sample; p: pellet fraction of induced sample; sur: soluble fraction of induced sample; Ft: flowthrough; E1,E2, E3: eluition fractions.

3.4 Cm-BVMO: Putative type I BVMO from *Cyanidioschizon merolae*

Cyanidioschyzon merolae 10D is a primitive unicellular red alga (Fig. 3.18) from the phylum Rhodophyta (Fig. 3.19), which lives in acidic environments (like hot springs), even at a pH < 2 and temperature of 45°C. It's one of the

photosynthetic eukaryotes with very simple cell architecture offering advantages for studies on mitochondrion and chloroplast [59, 80]. The cell does not have a rigid cell wall and contains a single nucleus, a single mitochondrion and a single chloroplast, divisions of which can be highly synchronized by light/dark cycles [107]. Its genome is really small, like 16 Mbp, (which is about 1/7 of that of *Arabidopsis thaliana* and about 1/27 of that of rice *Oryza sativa*).

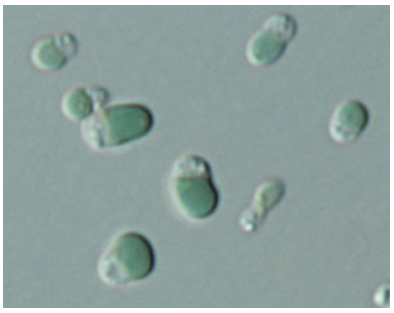


Figure 3.18: The unicellular red alga *C. merolae* 10D. The dividing cell contains a nucleus, a V-shaped mitochondrion, a dumb-bell-shaped plastid, a microbody and a Golgi apparatus, divisions of which can be highly synchronized by light/dark cycles (images from NIES-1332 (http://mcc.nies.go.jp/) and from Matsuzaki et al. 2004 [7]).

Domain:	Eukaryota
Phylum:	Rhodophyta
Subphylum:	Cyanidophytina
Class:	Cyanidiophyceae
Order:	Cyanidiales
Family:	Cyanidiaceae
Genus	Cyanidioschizon

Figure 3.19: Taxonomic classification of *Cyanidioschizon merolae*10D.

The genome sequence of *C. merolae* 10D was published in 2004 [7], and all sequence data can be downloaded from the Cyanidioschyzon merolae Genome Project Web site (http://merolae.biol.s.u-tokyo.ac.jp/), whereas sequencing of the chloroplast genome was accomplished in 1999 [81] and has revealed that the chloroplast genome of this organism retains many genes that are not present in

the chloroplast genomes of other algae and plants; they probably represent one of the most ancestral plastid genomes. So, from an evolutionary perspective, *C. merolae*, can be considered a model system for studying the origin, evolution and fundamental mechanisms of eukaryotic cells. The strain NIES-1804, which we purchased from the Microbial Culture Collection-Nies, (Japan), was identified by Mio Ohnuma at Sardinia Island, Italy. The alga was cultivated in M-Allen liquid medium and the starting culture was directly used for extraction of genomic DNA. M-Allen is a medium for freshwater, terrestrial, hot spring and salt-water algae; it contains organic salts, EDTA, metal traces with a pH of 2.5, using sulfuric acid.

3.4.1 Cloning of Cm-BVMO

The genomic DNA extracted from the organism was used as starting material to obtain the coding sequence of the putative Cm-BVMO. A first PCR reaction, using Cm3 forward and Cm4 reverse primers, was performed to obtain a 2198 pb DNA fragment, which was employed as template for the second mutagenic PCR reaction, using Cm1 forward and Cm2 reverse primers. The final amplimer had the attended size of 1627 bp (Fig. 3.20).

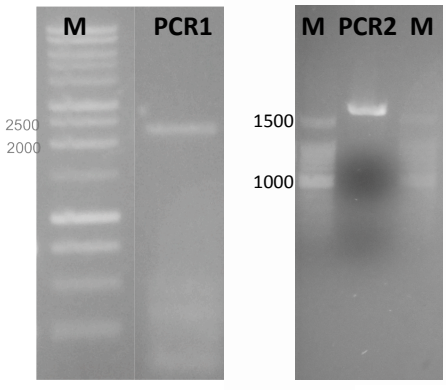


Figure 3.20: 1% agarose gel analysis of preliminary 2198 bp and final 1627 bp PCR products (PCR1 and PCR2). M: standard marker (Promega).

A restriction digestion using both enzymes *NcoI* and *NotI* was performed to prepare the insert for the ligation into pET28a. The resulting ligation product

was used to transform chemically competent XL1-Blue cells. Recombinant plasmids were examined by analytical restriction digestion with *EcoRV*. This analysis led to the formation of three bands, whose size was compatible with the lengths of 4326 bp, 2429 bp and 105 pb, as calculated for the digestion of the attended plasmid pET28a-Cm-BVMO (Fig. 3.21). The smaller band is not visible on the agarose gel, due to its low fluorescence (Fig. 3.22). One of the positive plasmids was then analysed by *NcoI/NotI* restriction to confirm the presence of the 1627 bp DNA fragment.

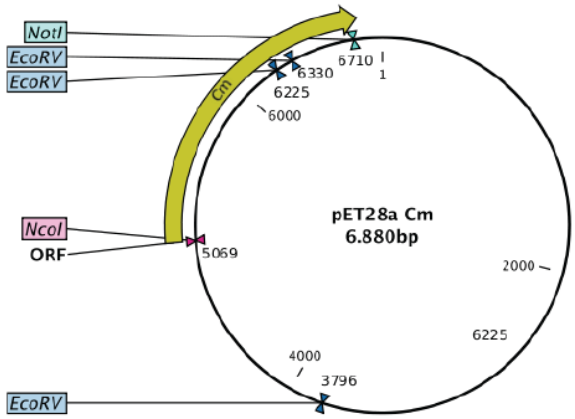


Figure 3.21: Map of the plasmid pET28a-Cm-BVMO.

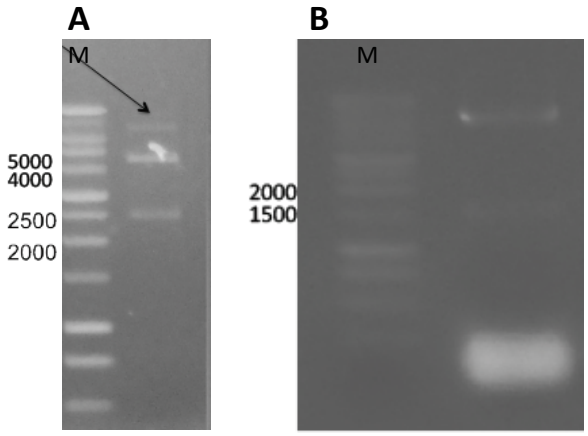


Figure 3.22: 1% agarose gel analysis of *NcoI/NotI* (A) and *EcoRV* (B) restriction pattern of plasmid pET28a-Cm-BVMO ; M: standard marker (Promega).

3.4.2 Expression of Cm-BVMO

After transformation of chemically competent BL21 (DE3) cells by plasmid pET28a-Cm-BVMO, a pre-culture was grown overnight and used the day after to inoculate a main culture for expression. Basing on the previous expression tests for Os-BVMO and Pp-BVMO proteins, we decided to use 18°C as growth temperature for the culture. Moreover, in order to avoid a possible flavin cofactor deficiency and a consequent protein misfolding, 100 μM riboflavin was added to the culture medium. As shown by the SDS-PAGE of Fig. 3.23, the target protein Cm-BVMO is well expressed after overnight induction by IPTG 0.2 mM. As in the case of the recombinant protein from *Physcomitrella*, it was partially produced in form of insoluble inclusion bodies and detected in the pellet fraction, but a considerable quantity was also present in the soluble fraction.

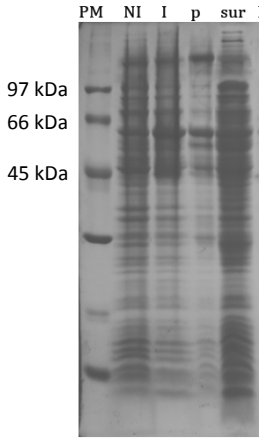


Figure 3.23: 12 % SDS-PAGE analysis of total extracts of cells transformed by plasmid pET28a-Cm-BVMO and induced at 18°C: the recombinant Pp-BVMO. protein is indicated by the arrow. PM: protein standard marker; NI: noninduced sample; I: over night induced sample; p: pellet fraction of induced sample; sur: soluble fraction of induced sample.

3.4.3 Purification of Cm-BVMO

Recombinant Cm-BVMO, supplied of His-tag, was purified by Ni^{2+}-chelating affinity chromatography starting from a soluble crude extract, obtained by centrifugation after cell disruption by French Press. The resulting yellow protein

fractions e1, e2, e3 and e4 were eluted and collected (Fig. 3.24). The presence of the target protein during in these fractions as well as in other fractions collected throughout the purification was examined by SDS-PAGE gel (Fig. 3.25). All buffers used for the purification were pre-cooled at 4°C and all fractions were kept in ice. Then, fractions containing the purified enzyme were desalted and finally concentrated by centrifuge filters. The evaluated concentration of the purified protein was 356 mM.

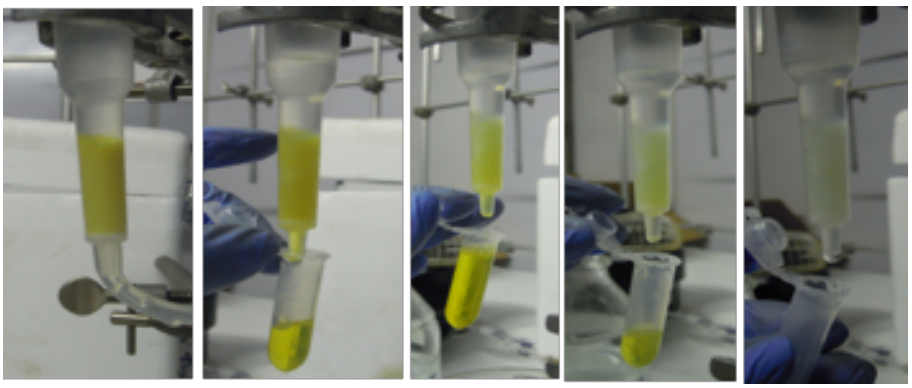


Figure 3.24: Purification steps of Cm-BVMO via immobilized metal affinity chromatography (IMAC).

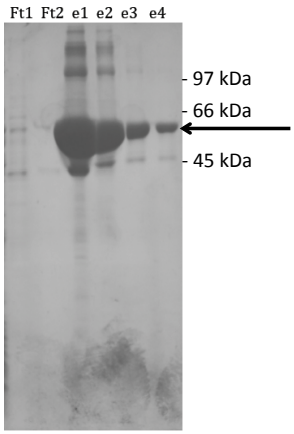


Figure 3.25: 12% SDS-PAGE analysis of fraction from IMAC purification of Cm-BVMO. Ft1, Ft2: flowthroughs; e1,e2,e3,e4: elution fractions.

3.5 Te-BVMO: Putative type I BVMO from *Trichodesmium erythraeum*

Trichodesmium erythraeum IMS101 is a marine filamentous cyanobacteria (Fig. 3.27) living in tropical and subtropical oceans. It is commonly called "sea sawdust" because its colonies and large brown blooms can be mistaken as sandbars by ships in the ocean (Fig. 3.26). It contain many gas vesicles aligned so that their longitudinal axis is oriented perpendicular to the direction of the filament [7] allowing its flotation according to changes in the concentration of carbohydrates within the cell [19]. Photosynthetic lamallae can be seen throughout the cell, but tend to be more populous toward the center of the cell [7]. *Trichodesmium erythraeum* plays an important role in the nitrogen fixation using photosynthesis [56]. It converts inorganic nitrogen gas to organic nitrogen that can be utilized by other organisms. The complete genome of *Trichodesmium erythraeum* has been sequenced by the DOE Joint Genome Institute (University of California) and has been deposited in 2006.

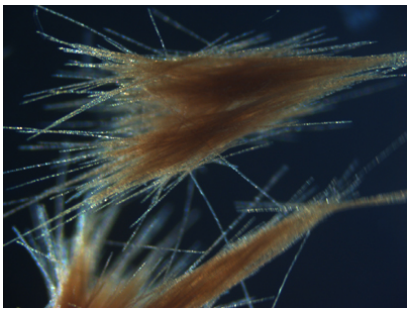


Figure 3.26: *Trichodesmium erythraeum* IMS101(image from the Alfred Wegener Institute (AWI) http://www.awi.de).

3.5.1 Cloning of Te-BVMO

The starting material for the cloning was genomic DNA provided by the lab of prof. Ilana Berman-Frank, Faculty of Life Sciences of the Bar Ilan University, Israel. Te5 forward and Te6 reverse primers were used in a first PCR reaction to replicate a preliminary large DNA fragment of 1934 pb containing the gene coding for the putative monooxygenase (Fig. 3.28, part A). Starting from the $A_{(1)}$TG of

Domain:	Prokaryotes
Kingdom:	Bacteria
Phylum:	Cyanobacteria
Order:	Oscillatoriales
Genus:	*Trichodesmium;*

Figure 3.27: Taxonomic classification of *Trichodesmium erythraeum* IMS101.

the open reading frame a *NcoI* restriction site is present at position 651. In order to remove this internal *NcoI* restriction site (which would have made following cloning difficult) we choose to suppress it by performing amplification in two steps. In the first step, two separate mutagenic PCR reactions were carried out using the same template, that is the purified Te5-Te6 amplimer: primers Te1 forward was combined with Te9 reverse primer in the first PCR tube; primers Te8 forward and Te2 reverse were mixed together in the second tube (Fig. 3.28, part B). As previously described for the cloning of genes from other organisms, the most external primers (Te1 and Te2) were designed to introduce *NcoI* and *NotI* restriction sites respectively at the ATG and downstream of the stop codon. Te8 and Te9 were complementary oligonucleotides differing from the wild-type sequence by a single mis-match, able to suppress the undesired *NcoI* site without altering the corresponding translated amino acid. From these mutagenic PCR reactions we obtained two DNA fragments, 653 pb and 847 pb long, containing complementary overhangs. Both were employed as DNA templates to perform a last PCR amplification using Te1 forward and Te2 reverse as primers and amplifying the desired 1487 pb long coding sequence (Fig. 3.28, part C).

Afterwards, the obtained ligation products were used to transform chemically competent XL1-Blue cells. Recombinant clones were picked up and the corresponding plasmid DNAs were isolated. Those containing the correct insert were identified by restriction enzyme analysis using *NcoI* and *NotI* (Fig. 3.29). As shown by the agarose gels displayed in Fig. 3.29, three positives colonies (Te2, Te4, Te6) were identified in the screening, by the releasing of an insert of the correct size (1487 pb). The plasmid purified from clone Te2 was sequenced and

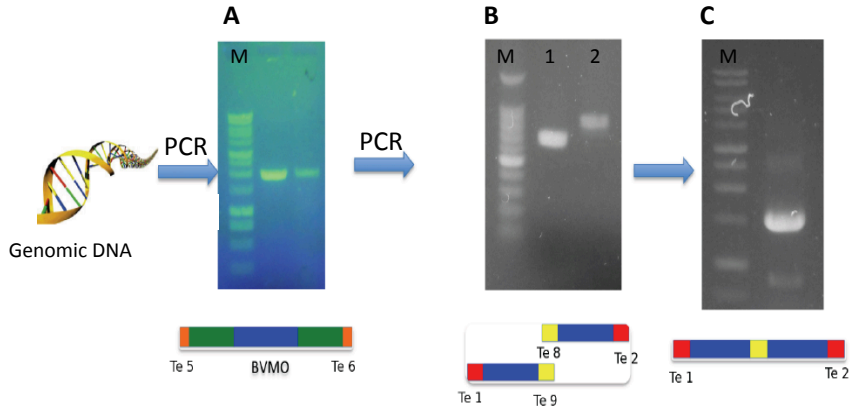


Figure 3.28: 1% agarose gel analysis of PCR amplification of the sequence coding for the putative Te-BVMO. M: standard marker.

an aliquot of the corresponding culture was stocked in glycerol.

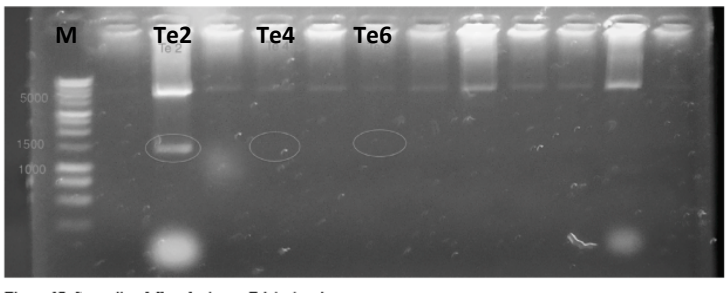


Figure 3.29: 1% agarose gel analysis of restriction by *NcoI/NotI* of recombinant plasmids from clones obtained from bacterial transformation. M: standard marker.

3.5.2 Expression of Te-BVMO

As done for all other putative BVMO proteins, after transformation of the appropriate BL21(DE3) strain, pre-cultures were grown overnight and used to inoculate cultures for expression. Overexpression tests were performed in order to define the better conditions for expression. Different growth temperatures, 18, 30 and 37°C, and different induction times were checked. SDS-PAGE analysis was used to verify the expression and the solubility behavior of the protein after cell disruption. We observed that, apparently, there was no expression of the

target protein at all conditions tested (Fig. 3.30). An immunoblotting assay, performed using anti-His tag antibody, confirmed the result of the expression tests on two different clones, Te2 and Te4 (Fig. 3.31). The anti-His tag antibody detected some bands at low apparent molecular weight: these were possibly due to proteolytic cut/degradation of the recombinant protein or aspecific recognition of endogenous bacterial proteins, since similar bands had been detected by the same antibody in analysis of recombinant putative BVMOs from *Oryza* and *Haloterrigena* (see next paragraph). Differently from those cases, however, no trace of the fully translated product was revealed.

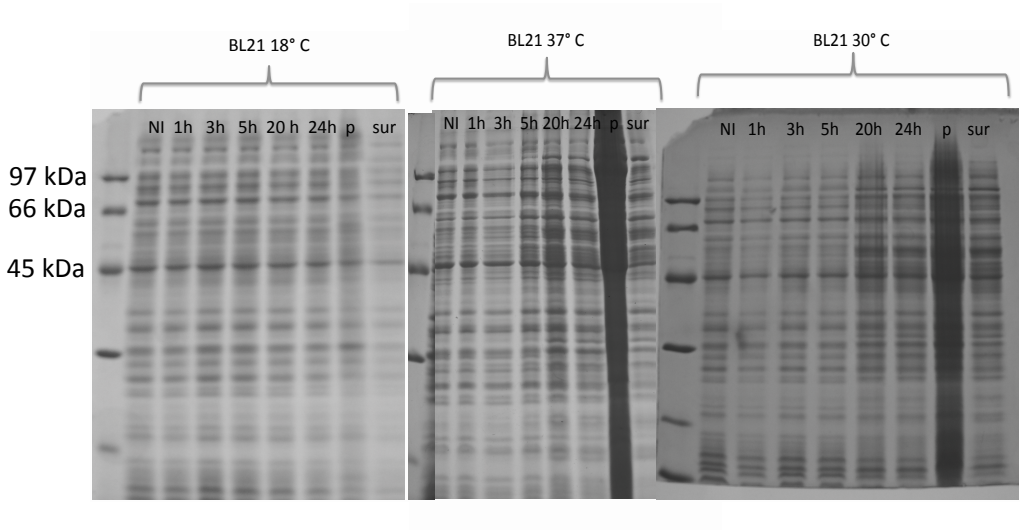


Figure 3.30: 12% SDS-PAGE analysis of total extracts from cells expressing Te-BVMO, grown at 18, 30 and 37°C and induced with IPTG 0.2 mM. M: protein standard marker; NI: noninduced sample; 1h: 1 hour induced sample; 3h: 3 hours induced sample; 5h: 5 hours induced sample; 20h: 20 hours induced sample; 24h: 24 hours induced sample; p: pellet fraction of induced sample; sur: soluble fraction of induced sample.

Re-sequencing of inserts and corresponding flanking regions on plasmids pET28a-Te-BVMO Te2 and Te4 (i.e. promoter and terminator sequences) ruled out the hypothesis that incidental mutation(s) could be responsible for switching off the expression. We then considered the possibility that secondary structures of the Te-BVMO transcript could block the beginning of the translation, preventing recognition of the ribosome-binding site (RBS) or the initial codon AUG, or could hamper elongation of the growing polypeptide (Fig. 3.32). As a matter of

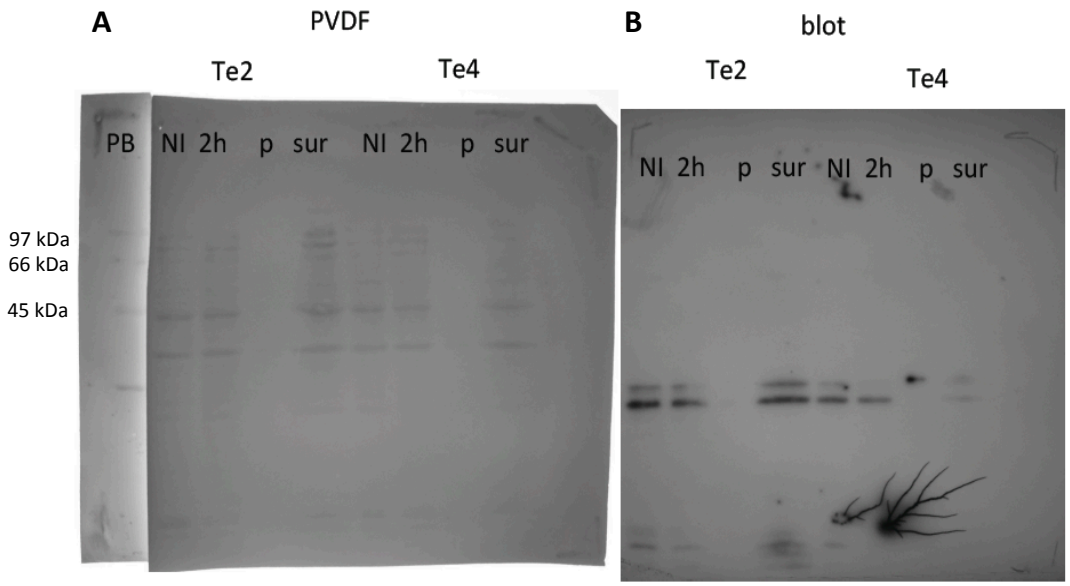


Figure 3.31: Ponceau Red stained PVDF membrane (A) and Western Blot (B) of cell extracts from clones Te2 and Te4. M: protein standard marker; NI: noninduced sample; 2h: 2 hour induced sample; p: pellet fraction of induced sample; sur: soluble fraction of induced sample.

fact, the bioinformatics analysis of the Te-BVMO mRNA predicted possible secondary structures, one example of which is shown in Fig. 3.32. In this particular secondary structure, i.e the one corresponding to the minimal free energy, we observed the presence of a hairpin at the 5'-terminus of the mRNA. We wonder if this hairpin could interfere with translation.

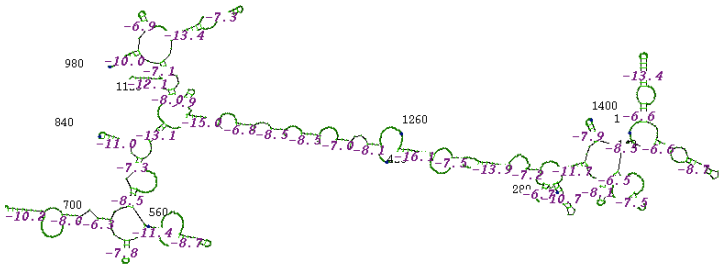


Figure 3.32: mRNA secondary structure of Te-BVMO

To investigate this possible hypothesis, we decided to engineer the transcript by changing the 5'-terminal part of the Te-BVMO gene with the 5'-terminal part

of the Pp-BVMO gene, very well-expressed in our system. In order to realize the chimeric construct two mutagenic PCR reactions were carried out. The first one was performed using Pp1 forward e PpTe1 reverse primers starting from the 2188 bp long ORF of *P. patens* as template DNA and obtaining the 5'-terminal part of the chimeric gene. The second mutagenic PCR reaction was carried out using PpTe2 forward e Te2 reverse primers and the 1934 bp long ORF of *T. erythraeum* as template, in order to obtain the 3'-terminus of the designed construct. PpTe1 and PpTe2 primers present a complementary portion of 20 pb respectively to *P. patens* and *T. erythraeum* and complementary 13 bp long tales for blending the two obtained fragments; Pp1 and Te2 respectively introduce the restriction sites *NcoI* and *NotI* at its two extremities. By a further final PCR reaction, using Pp1 forwards and Te2 revers primers, the hybrid gene (1456 bp long) was amplified (Fig. 3.33).

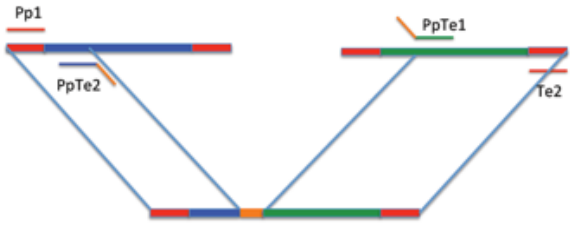


Figure 3.33: Cloning strategy used for production of the chimeric (Pp-Te-)gene.

Once digested by restriction enzymes, the hybrid ORF was used in a ligation with pET, the products of which were transformed in chemically competent XL1-Blue cells. Recombinant clones were picked up and the corresponding plasmid DNAs were isolated: those containing the new chimeric insert were identified by restriction enzyme analysis using *NcoI/NotI*. Hybrid plasmid Te2 was sequenced and the corresponding colony was conserved in glycerol stock. BL21 (DE3) cells were transformed by the new plasmid Te2 and overexpression tests were performed. Two different growth temperatures were checked, 18 and 37°C. In both cases the result was the same of the preceding experiment: no expression of the chimerical protein could be observed (data not shown). This result brought us to conclude that the secondary structure of the mRNA at its 5'-terminus could not be the only reason for suppression of Te-BVMO expression. Other investigations are needed to solve the question.

3.6 Ht-BVMO: Putative type I-BVMO from *Haloterrigena turkmenica*

Haloterrigena turkmenica DSM 5511 is the type strain of the species *Haloterrigena turkmenica*, which is the type species of the genus *Haloterrigena* (Fig. 3.35). It is an extreme halophilic archeum isolated from a sulphate saline soil in Turkmenistan. On the basis of 16S rRNA analysis, it belongs to the Euryarchaeota phylum that, in addition to extreme halophiles, includes strict anaerobic methanogens (procaryotes that produce methane), differing on their physiology. *Haloterrigena turkmenica* is a relatively fast growing, chemoorganotrophic, carotenoid containing organism, requiring at least 2 M NaCl for growth. Its cells occur mostly as single cells (Fig. 3.34); colonies are pigmented red or light pink due of the presence of C_50-carotenoids. It's a chemo-organotrophic and aerobic organism [111].

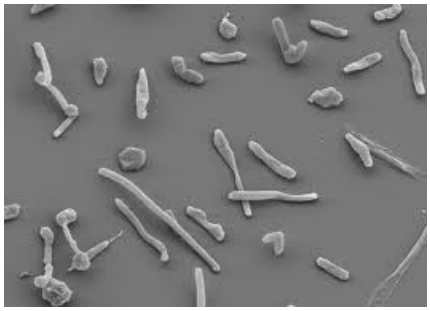


Figure 3.34: Scanning electron micrograph of *H. turkmenica* DSM 5511 (image from Saunders et al. 2010 [93].

3.6.1 Cloning of Ht-BVMO

The gene coding for the putative Ht-BVMO (locus tag=Htur_3838) has been identified on plasmid pHTUR01 of *H. turkmenica* DSM 5511. On the base of observed sequence similarity, the gene product is annotated in data-banks as "cyclohexanone monooxygenase". The organism *Haloterrigena turkmenica* is available at the German Collection of Microorganisms and Cell Cultures (DSMZ) but the presence of the plasmid is not assured. For this reason, a synthetic gene encoding the putative CHMO was purchased from GeneArt®, cloned into the

Domain:	Archaea
Kingdom:	Euryarchaeota
Phylum:	Euryarchaeota
Class:	Halobacteria
Order:	Halobacteriales
Family:	Halobacteriaceae
Genus:	*Haloterrigena*

Figure 3.35: Taxonomic classification of *Haloterrigena turkmenica* DSM 5511.

plasmid pMK-RQ (kanR) (Fig. 3.36). The gene was drawn and synthesized with the flanking restriction sites *NcoI* and *NotI*, required for cloning, and a codon usage optimized for translation in *E. coli*.

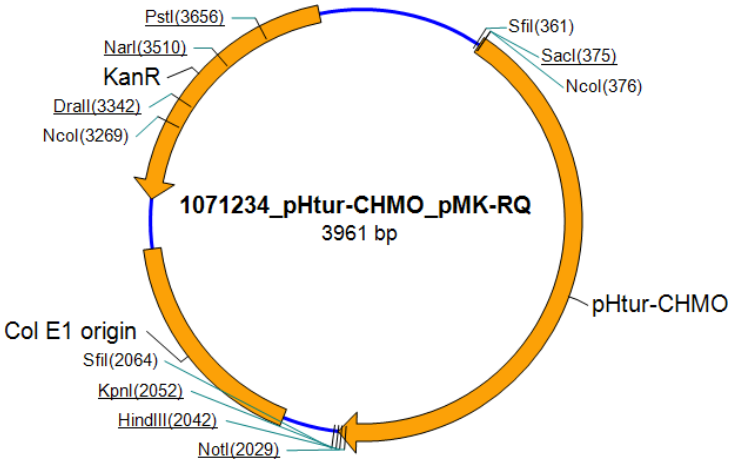


Figure 3.36: Map of the plasmid pHtur-CHMO-pMK-RQ.

The existing *NcoI/NotI* restriction sites in the pHtur-CHMO_pMK-RQ plasmid would have facilitate the cloning step if it had not been. Digestion of the purchased plasmid by enzymes *NcoI* and *NotI* produced three DNA fragments 1653, 1240 and 1068 pb long, because of the presence of another *NcoI* restriction site in the carrier vector pMK-RQ (Fig. 3.37). We therefore choose a "shot-gun" cloning approach in which all DNA fragments produced by *NcoI/NotI* restriction digestion were mixed in the same ligation reaction with pre-digested pET28a. Being possible different ligation products, a special attention was given to the

screening of recombinant colonies.

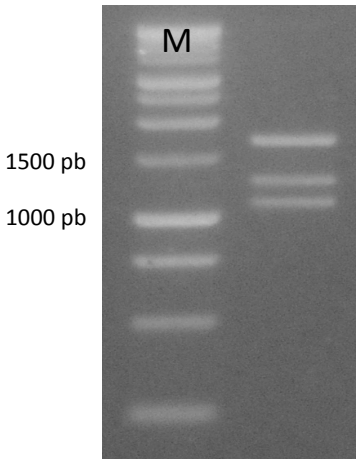


Figure 3.37: 1% agarose gel analysis of *NcoI/NotI* digestion pattern of pHtur-CHMO_pMK-RQ; M: standard marker (Promega).

Transformed XL1-Blue clones were picked up and the corresponding plasmid DNA was isolated: those containing the correct insert were identified by restriction enzyme analysis using *NcoI/NotI*. As shown by the agarose gels, two positives colonies (Ht5 and Ht35) were identified by the presence of released inserts of the correct size (1653 pb), corresponding to the open reading frame for Ht-BVMO. One of the two selected plasmids was sequenced and the corresponding colony was conserved in glycerol stock (Fig. 3.38).

3.6.2 Expression of Ht-BVMO

After transformation of the BL21(DE3) expression strain by plasmid pHt5, precultures were grown overnight and used to inoculate cultures for expression the following day. Tests were performed in order to define the better conditions for expression. Different growth temperatures were checked, 18, 30 and 37°C. SDS-PAGE analysis was used to verify the expression of the protein at different induction time at the three temperatures and its solubility behavior after cell disruption. We observed that from the third hour after IPTG induction a large amount of protein was produced at both temperatures tested, even at the lowest one of 18°C. First of all, we noticed that the apparent molecular weight of the expressed protein (80 kDa) was very different from the theoretical one (62 kDa).

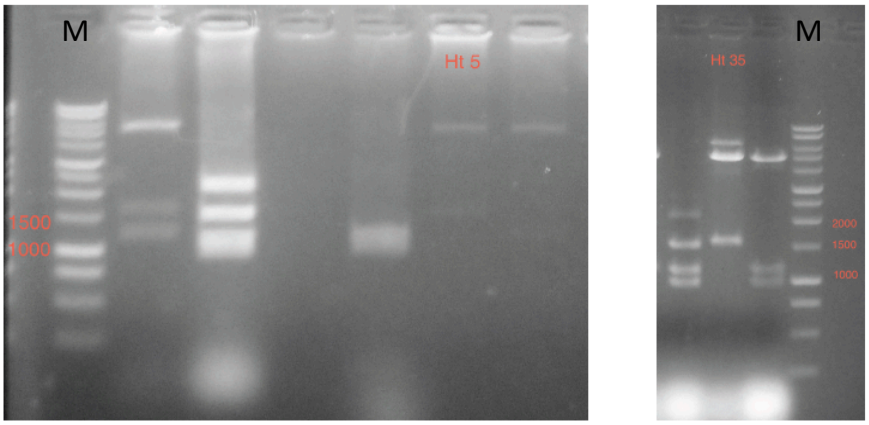


Figure 3.38: 1% agarose gel analysis of the digestion of the obtained plasmids using *NcoI/NotI*. M: standard marker.

Furthermore, the overexpressed protein was possibly insoluble and collected in inclusion bodies, being detected in the pellet fraction (Fig. 3.39). The identity of the target protein was confirmed by an immunoblotting assay, performed using anti-His tagged antibody (Fig. 3.40).

A possible explication for the unusual electrophoretic migration of the recombinant protein was the presence of a post-translational covalent modification of the target protein, since all samples for SDS-PAGE analysis had been prepared in a denaturing loading buffer and subjected to boiling. However, it is known from the literature that proteins containing high quantity of negatively charged amino acids can have an abnormal migration in SDS-PAGE. Indeed, Ht-BVMO contains 169 charged amino acids, 119 of which negatively charged. The actual reason for the apparent molecular weight of the protein, higher than the calculated value, was not further investigated.

Concluding remarks:

Using bioinformatics tools we have identified and selected five new putative BVMOs from five different organisms, *Oryza sativa* (Os; land plant), *Physcomitrella patens* (Pp; moss), *Cyanidioschyzon merolae* (Cm; red alga), *Trichodesmium erythraeum* (Te; cyanobacterium), *Haloterrigena turkmenica* (Ht; archibacterium). In particular, basing on most updated literature, the photosynthetic eukaryotes *Oryza, Physcomitrella* and *Cyanidioschyzon* would result very uncommon

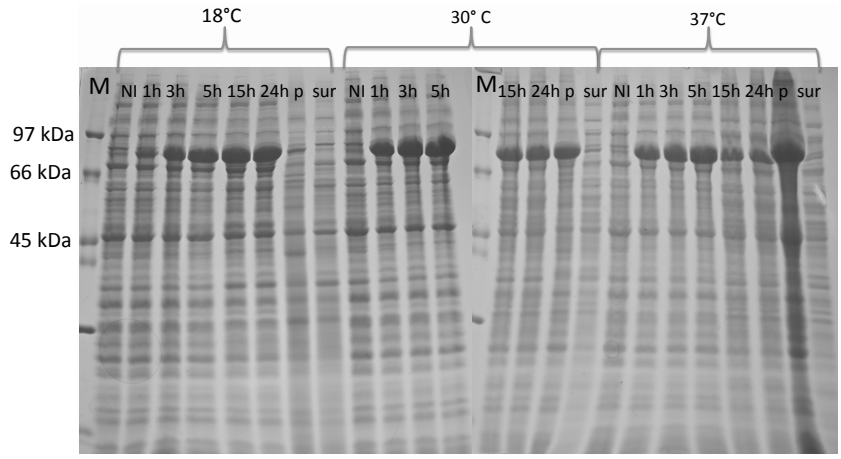


Figure 3.39: 12% SDS-PAGE analysis of total extracts from *E.coli* cells expressing Ht-BVMO, grown at 18, 30 and 37°Cand induced by IPTG 0.2 mM. M: protein standard marker; NI: noninduced sample; 1h: 1 hour induced sample; 3h: 3 hours induced sample; 5h: 5 hours induced sample; 15h: 15 hours induced sample; 24h: 24 hours induced sample; p: pellet fraction of induced sample; sur: soluble fraction of induced sample.

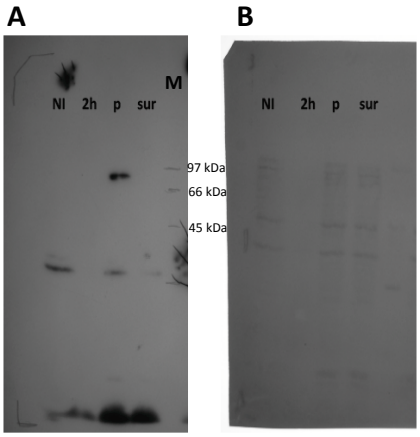


Figure 3.40: Western Blot (A) and PVDF membrane stained by Ponceau Red (B) of total extracts from *E. coli* cells expressing Ht-BVMO. M: protein standard marker; NI: noninduced sample; 1h: 1 hour induced sample; p: pellet fraction of induced sample; sur: soluble fraction of induced sample.

sources for BVMOs. Although the starting material was different, the cloning strategy used was the same for all the five putative enzymes. Overexpression of enzymes was made in *E. coli* BL21 (DE3) and tested at 18, 30, and 37°Cand concentration of IPTG ranging from 0.2 to 1 mM. Whole cells and partially purified extracts were analysed by SDS-PAGE. All the putative enzymes, apart from Te-BVMO, were very well expressed, but they were present in the cell debries after cell homogenization. Thus, solubility resulted to be the main difficulty. To overcome the problem of correct folding and confinement in inclusion bodies, we tried many known strategies, like expression at low temperature in the *E.coli* ArcticExpress strain containing cold-adapted chaperones or fusion to suitable partners. Adding an excess of riboflavin, the precursor of the FAD, to the culture resulted in successful folding for the proteins from *Physcomitrella patens* and *Cyanidioschyzon merolae*, showing the importance of the flavin cofactor in the process. Overall result concerning cloning, expression and purification of the five putative type I BVMOs is summarized in Fig. 3.41.

	Cloning	Expression	Purification
Pp-BVMO	pET28a	BL21(D3)	✔
Os-BVMO	pET28a	BL21(D3)	✗
Cm-BVMO	pET28a	BL21(D3)	✔
Ht-BVMO	pET28a	BL21(D3)	✗
Te-BVMO	pET28a	BL21(D3)	✗

Figure 3.41: Cloning, expression and purification of the five putative type I BVMOs identified.

Chapter 4

Characterization of Type I BVMOs from Photosynthetic Eukaryotes: Cm-BVMO and Pp-BVMO

Abstract:
After cloning and expression of five putative genes encoding type I Baeyer-Villiger monooxygenases, only two of them were successfully purified by IMAC chromatography. One BVMO was from *Phiscomitrella patens* (Pp-BVMO) and the other one from *Cyanidioschizon merolae* (Cm-BVMO). Because the BVMO-identifying "fingerprint" motif in Pp-BVMO (FxGxxx**Y**xxxWP) differs in one amino acid with respect to the typical (FxGxxx**H**xxxWP) sequence motif of type I BVMOs, a mutant of the flavoenzyme was prepared by site-directed mutagenesis, in order to repair the Tyr (Y) residue by introducing a His (H). A biocatalytic characterization of the three generated enzymes, Cm-BVMO, Pp-BVMO and Pp-BVMO-Y173H was carried out at the Molecular Enzymology group of the Groningen Biomolecular Science and Biotechnology Institute (GBB), University of Groningen (The Netherlands) headed by Prof. Dr. Marco Fraaije. After verification of their Baeyer-Villiger monooxygenase activity, conditions optimization (pH, temperature and stability) has been performed. After that, the enzymes were tested to investigate their substrate specificity. A steady-state kinetic analysis was carried out in order to obtain the kinetics parameters and some conversions were made in order to determine the selectivity profile of the newly discovered BVMOs.

4.1 Enzyme activity and concentration

The flavin cofactor present in flavoproteins is essential as adaptable redox catalyst for oxidation-reduction reactions. The flavin cofactor is derived from the riboflavin precursor (vitamin B2) which consists of two parts: an isoalloxazine ring system, responsible for the yellow colour of the oxidized form, and a ribose moiety. Flavin mononucleotide, FMN consists of a molecule of riboflavin linked to a phosphate group on the 5'-hydroxyl group of the ribose chain. Phosphorylation of riboflavin is achieved by a riboflavin kinase. Flavin adenine dinucleotide (FAD), the most common flavin cofactor, is formed from a molecule of FMN and a molecule of adenosine triphosphate (ATP) joined together by action of FAD synthetase. The result is a riboflavin-diphospho-adenosine molecule, with the diphosphate group attached to the 5'-hydroxyl group of the linearized ribose on one side and a hydroxyl group of the adosine cyclic ribose on the other side (Fig. 4.1). Both riboflavin derivatives, flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD), function as cofactors in a wide variety of redox enzymes. In fact they can act as oxidizing agents because of the ability of the isoalloxazine ring to accept a pair of hydrogen atoms yielding the reduced forms of the flavoproteins ($FMNH_2$ and $FADH_2$) (Fig. 4.2). It is well known that flavins are capable of accepting both one- and two-electrons and, as a result, they can exist in three oxidation states, the oxidized state, the one-electron reduced semiquinone state (flavosemiquinone) and the two-electron reduced hydroquinone state (flavohydroquinone). Depending on the protonation state there are many forms of the flavin for each oxidation state (Fig. 4.3).

An interesting aspect of flavins is that each of these redox states has specific spectral properties, making it relatively easy to follow redox transitions by spectral changes. In fact, in the oxidized form the flavin exhibits two absorption peaks at $\approx$ 370 and $\approx$ 450 nm. The exact shape, position and intensity of these absorbance peaks dependent upon the protein environment. Reduction of the flavin to its two-electron reduced form results in the loss of the typical two absorbance peaks in the UV-visible region. The purified BVMOs all exhibited typical flavin absorption spectra indicating that they indeed contain a flavin cofactor, presumably FAD. To verify their activity towards the coenzyme

NADPH, an excess of reducing equivalents was added and UV-vis spectra were detected before and after NADPH addition (Fig. 4.4). The typical flavin spectrum was nearly completely lost upon adding NADPH confirming that NADPH is indeed an effective electron donor for these BVMOs. By knowing the extinction coefficient for free FAD ($\varepsilon_{\mathrm{FAD}}$, 450 nm = 11,300 $\mathrm{M^{-1}\ cm^{-1}}$) it is possible to quantify the exact amount of the flavoprotein by calculating the concentration of free FAD of the enzyme after denaturation. By this approach, the extinction coefficient of the flavin-containing enzymes could be calculated by comparing the absorption spectra before and after denaturing with 0.1% SDS. Assuming that the absorption spectrum of the unfolded protein can be compared to that of free FAD ($\varepsilon_{\mathrm{FAD}}$, 450 nm = 11.3 $\mathrm{mM^{-1}\ cm^{-1}}$), a molar extinction coefficient of 11.7, 13.7 and 11.9 $\mathrm{mM^{-1}\ cm^{-1}}$ (all at 450 nm), can be calculated for the native forms of Cm, Pp and PpY173H, respectively.

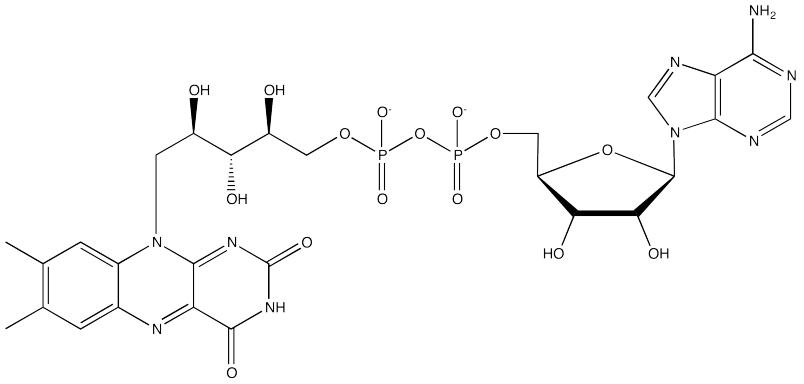


Figure 4.1: Structure of riboflavin, FMN and FAD.

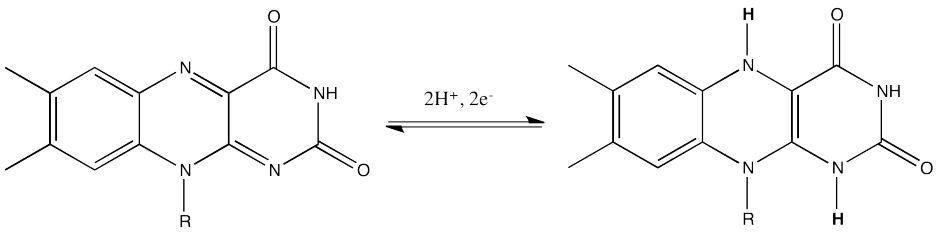


Figure 4.2: 2-Electron reduction of the isoalloxazine ring in flavin cofactors.

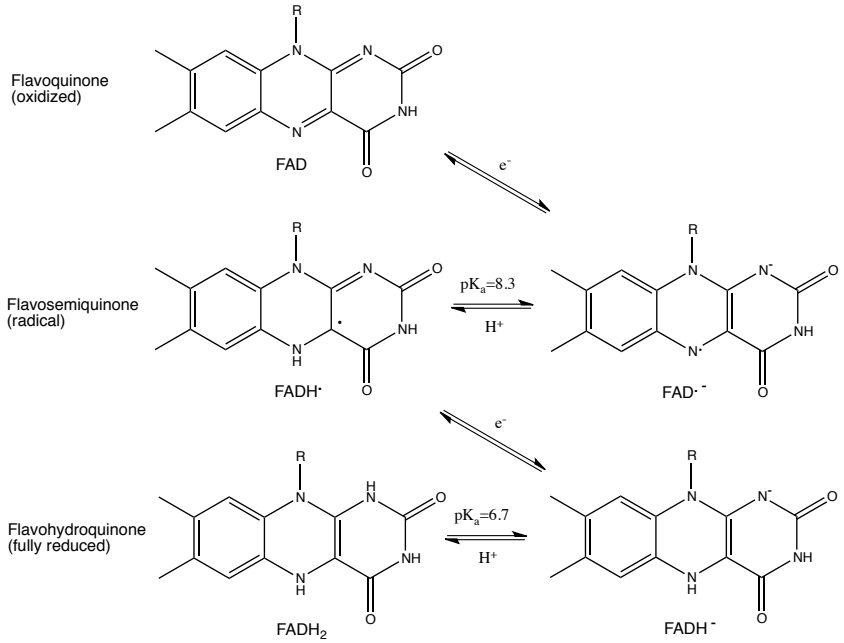


Figure 4.3: Different redox and protonation states of FAD.

4.2 pH optimum and temperature stability

All enzymes have an optimal pH and temperature at which reaction rates go fastest without denaturing the enzyme [18]. To determine the pH and temperature optima, the NADPH depletion assay was employed while varying the pH and temperature. In the NADPH depletion assay, the consumption of NADPH by a BVMO can be conveniently monitored by measuring the absorbance of NADPH at 340 nm at which the reduced coenzyme absorbs light.

The type I BVMOs all showed a bell-shaped pH profile with maximum activity for bicyclo[3.2.0]hept-2-en-6-one at pH 8.5 for Cm-BVMO and pH 8.0 for Pp-BVMO and PpY173H-BVMO. An optimum pH around 8 has also been found for PAMO, whereas for HAPMO and CHMO pH 7.5 and 9, respectively, were found to be the pH for maximum activity (Fig. 4.5).

The temperature optima were measured between 17°C and 70°C. For Cm-BVMO this revealed an increase of enzymatic activity with an increase of temperature until 70°C. At the latter temperature the initial enzymatic activity was

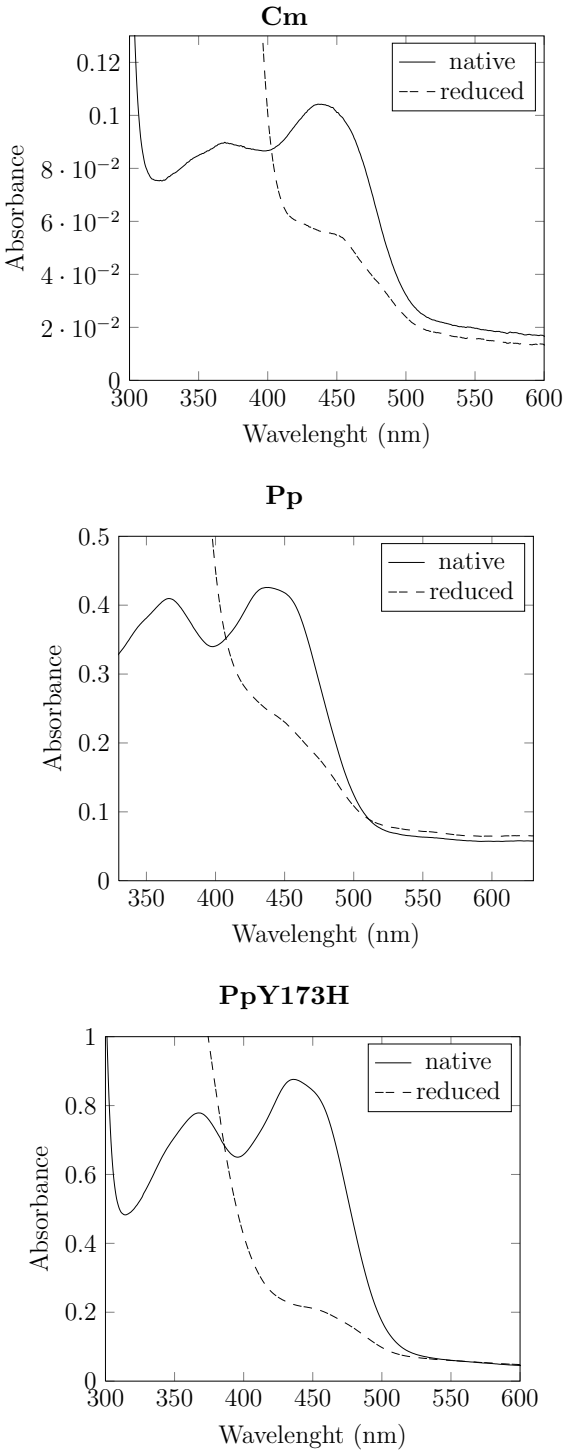


Figure 4.4: Uv-vis spectra of Cm, Pp and PpY173H before and after addition of NADPH. Spectra of native protein (solid line) and after addition of an excess of NADPH (dashed line).

five-fold higher when compared to the activity at 30°C (Fig. 4.5). Incubation of the enzyme for 1.5 h resulted in full retention of enzymatic activity between 20°C and 42°C. Higher incubation temperatures led to inactivation of the enzyme. Cm-BVMO seems to be a reasonably thermostable enzyme. Considering that *Cyanidioschizon merolae* typically grows at 45°C, this result was not so unexpected. The only thermostable type I BVMO found so far was discovered by Fraaije et al. [37] in *Thermobifida fusca* exhibiting an activity half-life of 24 h at 52°C. The optimal temperature for Pp-BVMO and Pp-BVM-Y173H was probed between 17°C and 40°C and it proofed to be around 25°C and 30°C for Pp-BVMO and Pp-BVMO-Y173H, respectively. Concerning stability, after 1.5 h of incubation both enzymes retained their activity between 17°C and 25°C.

4.3 Determining unfolding temperature (Tm)

Protein stability can be also analysed as the tendency of a protein to unfold. One of the methods to measure stability by determining the apparent unfolding or melting temperature of a protein (Tm). Unfolding experiments on Cm-BVMO, Pp-BVMO and Pp-BVMO-Y173H were performed using the *Thermo*FAD method, a *Thermofluor*®-adapted flavin *ad hoc* detection system for measuring the stability of flavoproteins. Using the *Thermofluor*® technique it is possible to measure the unfolding temperature of a protein through evaluation of the fluorescence of a dye in a Real-Time PCR instrument. As soon as the protein unfolds, the fluorescent dye will report on this. In the *Thermo*FAD method the flavin cofactor is used as an internal probe releasing fluorescence as soon as the protein denatures [34]. Usually a temperature gradient is applied, from 20°C to 90°C, detecting the fluorescence signal of the cofactor during heating. A sigmoidal curve is obtained by plotting the fluorescence intensity against the temperature. By determining the derivative of this sigmoidal curve, the melting temperature (Tm), of the protein can be determined.

Unfolding experiments using the *Thermo*FAD method revealed that the Tm of Pp-BVMO is 44°C and is comparable with that of the mutant PpY173H-BVMO (43.5°C). Apparently, the introduction of the "correct" histidine residue in the "fingerprint" motif of the mutant does not affect the melting temperature

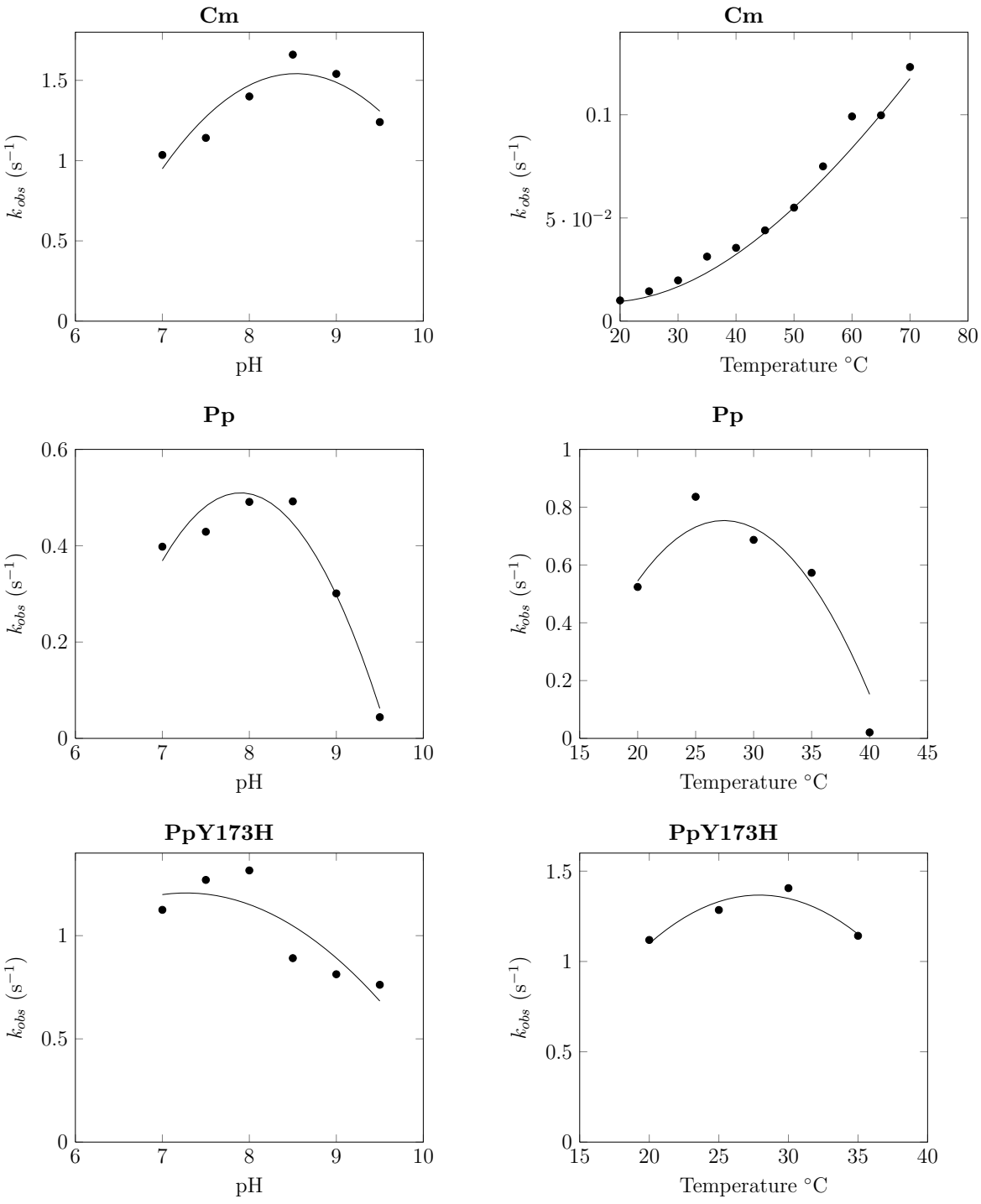


Figure 4.5: Optimum pH and temperature values for Cm, Pp and PpY173H.

of the enzyme. The *Tm* of Cm-BVMO was shown to be 56°C, which is considerably higher when compared with most known BVMOs and it confirms the measured thermostability of the enzyme (Fig. 4.6).

Moreover, a pH screening was performed with Cm-BVMO in order to evaluate changes in *Tm* at different pH values. In Tabble 4.1 the *Tm* values obtained for a set of different pH values are reported. The results coincide with those obtained from the optimum pH experiments. The best *Tm* value is at pH 8.5 which is also the optimum pH found for the enzyme (Table 4.1).

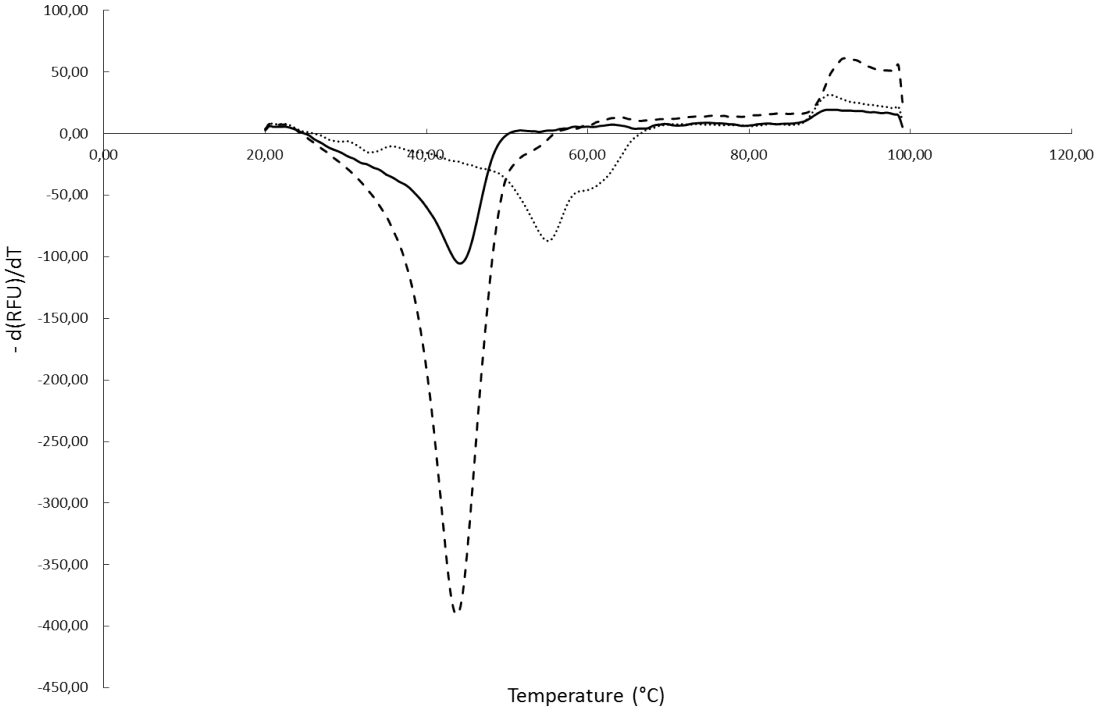


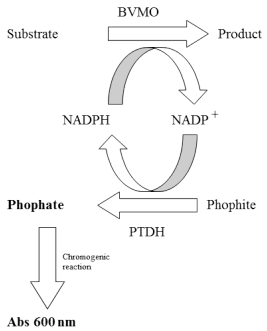
Figure 4.6: Melting curve derivatives of Cm (dot line), Pp (solid line), PpY173H (dashed line).

pH	*Tm* (°C)
7.0	53,00
7.5	54,50
8.0	55,50
8.5	56,00
9.0	55,50
9.5	53,50

Table 4.1: Melting temperatures obtained at different pH value.

4.4 Substrate profiling

To explore the biocatalytic potential of these new BVMOs a broad range of potential substrates was used in a so-called phosphate assay. In this indirect assay the Baeyer-Villiger oxidation is coupled with a regeneration reaction catalysed by a phosphite dehydrogenase (PTDH). PTDH is able to generate one molecule of phosphate, starting from a molecule of phosphite, for the regeneration of one NADPH molecule. The formed phosphate can be quantified by using a chromogenic reaction allowing spectrophotometric detection of the degree of conversion.



A collection of 46 substrates, that have been found to be converted by BVMOs, were tested. They belong to different types of molecule classes, e.g. linear and cyclic aliphatic ketones, aryl and aromatic ketones, steroids, aromatic amines, and aromatic sulfides. The substrate scope obtained for each enzyme is

summarized in Tab 4.2.

In general all the three BVMOs seem to have a broad substrate scope. As expected, they exhibit significant activity against bicycloheptenone, a typical substrate for type I BVMOs, and used as substrate for the NADPH depletion assays mentioned above. Good activity was also observed for linear ketones, in particular long-chain alkanones. Phenylacetone and some derivatives also seem to be good substrates, as well as some cyclic ketones like cyclopentanone and some sulfides. By comparing Pp-BVMO and the corresponding mutant Pp-BVMO-Y173H, the latter one seems to display a larger substrate profile suggesting a role of the mutated residue for substrate acceptance. Some other potential substrates such as acetone, indanone, phenindione, cyclopentadecanone, benzoin and steroids were poor or not accepted as substrates. To validate the results obtained from this preliminary substrate-screening assay, a set of selected substrates was chosen for further kinetics analysis and to evaluate some conversions by GC and GC-MS analysis.

4.5 Steady-state kinetics

Steady-state kinetic parameters of the three isolated enzymes, Cm-BVMO, Pp-BVMO and Pp-BVMO-Y173H were determined in 50 mM Tris/HCl using the optimal conditions for each enzyme, that are pH 8.0, at 25°C for Pp-BVMO and Pp-BVMO-H173Y and, pH 8.5, at 35°C for Cm-BVMO. Bicyclo[3.2.0]hept-2-en-6-one and NADPH were utilized as substrates. All the three enzymes showed a typical Michaelis-Menten behaviour and the kinetics parameters k_{cat} and K_{M} were calculated for a set of identified substrates in the preliminary screening assay. Selected substrates were linear aliphatic ketones, cyclic aliphatic ketones, aromatic ketones, and some sulfides. In particular for the linear ketone containing eight carbon atoms, the position of the keto group on the molecule was investigated. In Table 4.3 and Table 4.4 the results are summarized obtained for Cm, Pp and PpY173H, respectively.

In general, by comparing Michaelis constants for all the substrates and for all the three enzymes, it is worth noting that the K_{M} values are in the micromolar range, which are very low values for non-natural substrates. More variation was

	Substrate	Cm	Pp	PpY173H
1	Acetone	+		
2	Methylketone	+	+	
3	Methyl vinyl ketone	+	+	+
4	2-Octanone	++	++	+++
5	3-Octanone	+	++	+
6	4-Octanone	+		
7	2-Dodecanone	+++	++	+++
8	Butyl levulinate	+++	+	+++
11	3-Methyl-2,4-pentanedione	+		+
12	Cyclobutanone	+	+	+
13	Cyclopentanone	++	+	++
14	Cyclohexanone	+		+
15	Cyclopentadecanone			+
16	2-Oxocyclohexanecarbonitrile		++	+++
17	4-Methylcyclohexanone	+	+	+
18	2-Propylcyclohexanone		+	+
21	Dihydrocarvone			+
22	Cyclopropylmethylketone	+		++
23	Norcamphor		+	+++
24	Bicycloheptenone	++	++	+
25	Progesterone			+
26	Androstenedione			+++
27	4-Dimethylaminobenzaldehyde		+	+
28	Nicotin			++
31	Thioanisole	+		
32	Benzylethyl sulfide	++	+++	
33	Benzylphenyl sulfide			+++
34	Ethionamide		+	
35	Diphenylmethylthioacetamide			+
36	Thiacetazone			+
37	Indole	+		+++
38	3-Acetylindole			++
41	5-Mathylfurfural	+	++	+
42	Benzaldehyde		+	+
43	Acetophenone		+	++
44	4-Hydroxyacetophenone		+	+
45	2,6-Dihydroacetophenone			+++
46	3-Phenylpentane-2,4-dione	++	++	+++
47	Phenylacetone	+++	++	+++
48	4-(4-Hydroxyphenyl)-2-butanone	+++	+++	+
51	2-Phenylcyclohexanone			+
52	Benzoin			
53	Phenindione			
54	2-Indanone		+	+
55	1-Indanone		+	+
56	6-Hydroxy-1-indanone	+		

Table 4.2: Substrate scope. Observed activities measured by phosphate formation are shown for Cm-BVMO, Pp-BVMO and Pp-BVMO-Y173H. Activities are indicated as +, ++ or +++ and reflect a 1.2-, 2- or 5-fold increase in phosphate formation, respectively, when compared with uncoupling reactions, which lack the tested compound. Substrate structure are listed in Chapter 6-Material and Methods.

found for the k_{cat} values suggesting different rates of catalysis depending on the substrate considered. Unfortunately, also the k_{cat} values are considerably low. Cm exhibits the best catalytic efficiency against 2-dodecanone (k_{cat}/K_M 97.000 s^{-1} mM^{-1}), which seems one of the best substrates found for this enzymes. Octanone appears to be the second best substrate. In particular, if we look at the K_M and k_{cat}/K_M, the position of the keto group at C3 seems to be the preferred one. By comparing all the positions of the carbonyl function in octanone it can be noticed that the Michaelis constant is the main parameters influencing the catalytic efficiency because the k_{cat} values were found to be similar for all the positions investigated. Also for Pp and PpY173H linear ketones seem to be very well accepted substrates; 2-dodecanone is the one with the highest catalytic efficiency. Different from the position of the keto group in octanone for Cm, Pp prefers the keto group at position 2. 4-octanone is not well accepted by both Pp enzymes and the data are in agreement also with the previous substrate-screening assay (Table 4.1). As for Cm the k_{cat} values, for all the four linear C8-ketones tested, are approximately in the same range. Comparing turnover numbers of Pp and PpY173H it can be concluded that the mutant appears to be faster.

Cm, Pp and PpY173H are able to oxidize some substituted aryl ketones like phenylacetone and 4-(4-hydroxyphenyl)-2-butanone. The latter compound is a well-accepted substrate, especially for Pp and PpY173H. Moreover the kinetic analysis revealed a poor activity against butyl levulinate and 3-phenylpentane-2,4-dione which is in contrast with the preliminary results of the substrate screening assay. Thioanisole is accepted only by PpY173H whereas benzyl ethyl sulfide is converted by both Pp and PpY173H. The affinity for the coenzyme NADPH was found to be very high, as a K_M value < 5 μM was determined. Due to the limited solubility, other substrates such as 4-dimethylaminobenzaldehyde and 3-acetylindole could not be investigated for kinetic analysis.

As previously mentioned the general observation from the kinetic analysis is that the enzymes display rather low catalytic efficiencies, while micromolar range K_M values were found, no more than 0,4 s^{-1} for k_{cat} and around 10^3 s^{-1} M^{-1} for k_{cat}/K_M were found. Usually a k_{cat} for an enzyme is > 1 s^{-1}, (most k_{cat} values are in the range of 1-100 s^{-1}) and, the k_{cat}/K_M is ˜10^2 s^{-1} M^{-1} while k_{cat}/K_M values are normally in the range of 10^3-10^6 M^{-1} s^{-1}. It is known that

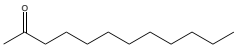
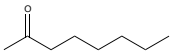
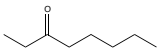
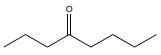
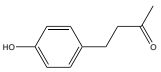
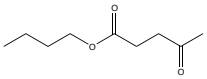
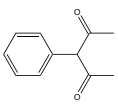
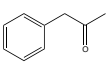
		Cm		
Substrate		**K_M [mM]**	**k_{cat} [s^{-1}]**	**k_{cat}/K_M [s^{-1} mM]**
2-Dodecanone		0,004 ± 0,001	0,383 ± 0,018	97,664
Octanal		0,010 ± 0,008	0,252 ± 0,043	25,325
2-Octanone		0,011 ± 0,006	0,250 ± 0,031	23,271
3-Octanone		0,003 ± 0,0005	0,210 ± 0,006	70,036
4-Octanone		0,004 ± 0,0009	0,202 ± 0,006	44,141
4-(4-Hydroxyphenyl)-2-butanone		0,007 ± 0,0007	0,079 ± 0,002	11,071
Bicyclo[3.2.0]hept-2-en-6-one		0,004 ± 0,0007	0,085 ± 0,002	20,721
Butyl levulinate		0,205 ± 0,091	0,181 ± 0,030	0,884
3-Phenylpentane-2,4-dione		0,099 ± 0,026	0,015 ± 0,001	0,153
Phenylacetone		0,097 ± 0,071	0,032 ± 0,006	0,330

Table 4.3: Steady-state kinetic analysis of Cm.

	Pp			PpY173H		
Substrate	K_M [mM]	k_{cat} [s^{-1}]	k_{cat}/K_M [s^{-1} mM^{-1}]	K_M [mM]	k_{cat} [s^{-1}]	k_{cat}/K_M [s^{-1} mM^{-1}]
2-Dodecanone	0,0004 ± 8,7 10^{-5}	0,096 ± 0,002	215,455	0,002 ± 0,0004	0,252 ± 0,011	160,145
Octanal	0,019 ± 0,009	0,061 ± 0,004	3,129	0,003 ± 0,0008	0,125 ± 0,006	43,792
2-Octanone	0,003 ± 0,0008	0,060 ± 0,002	12,727	0,003 ± 0,0006	0,204 ± 0,007	73,784
3-Octanone	0,010 ± 0,005	0,075 ± 0,008	7,582	0,0162 ± 0,007	0,123 ± 0,011	7,599
4-Octanone	//	//	//	//	//	//
4-(4-Hydroxyphenyl)-2-butanone	0,002 ± 0,001	0,068 ± 0,004	32,152	0,003 ± 0,0007	0,152 ± 0,006	48,011
Bicyclo[3.2.0]hept-2-en-6-one	0,006 ± 0,00008	0,058 ± 0,0001	9,864	0,016 ± 0,002	0,120 ± 0,005	7,531
Butyl levulinate	//	//	//	0,036 ± 0,016	0,095 ± 0,012	2,631
Phenylacetone	0,038 ± 0,016	0,077 ± 0,009	2,036	0,032 ± 0,0092	0,146 ± 0,0113	4,569
Thioanisole	//	//	//	0,026 ± 0,016	0,030 ± 0,0031	1,171
Benzylethyl sulfide	0,011 ± 0,005	0,090 ± 0,009	8,451	0,005 ± 0,0006	0,188 ± 0,004	39,310

Table 4.4: Steady-state kinetic analysis of Pp and PpY173H.

natural selection shapes the regulation of metabolic enzymes and plays a role in the evolution of enzymatic kinetic parameters, including the turnover number (k_{cat}) and the Michaelis constant (K_{M}). Kinetic parameters of enzymes belonging to the central metabolism typically are more efficient with respect to these of the intermediate and secondary ones. Such variations suggest a strong selective pressure towards primary metabolism enzymes, which are forced to increase their rates. However, for enzymes operating in secondary metabolism, which work only under specific conditions or for short periods of time, the selective pressure is weaker resulting in moderate/low k_{cat} and $k_{\mathrm{cat}}/K_{\mathrm{M}}$ values. Also for this reason identification of their natural substrate and reaction for secondary metabolism enzymes remains a demanding task. The photosynthetic eukaryotes Cm, Pp and the mutant PpY173H seem to fit in this hypothesis. Their substrate preferences point towards linear alkanones, structurally similar to side chains of chlorophylls and pheophytins, and could suggest some involvement in modifications of photosynthetic pigments. Furthermore, low $k_{\mathrm{cat}}/K_{\mathrm{M}}$ values in the range of 10^3 M^{-1} s^{-1}, or even lower, are commonly seen in plant secondary metabolism enzymes [86, 40].

4.6 Conversions

Biotransformations using purified enzymes and different types of substrates were investigated using GC and GC-MS analysis in order to detect both substrates and products. PTDH and phosphite were added for cofactor regeneration.

Conversions of various bicyclic ketones have been studied with CHMO [91]. In particular, bicyclo[3.2.0]hept-2-en-6-one is used to probe the bocatalytic potential of type I BVMOs. This conversion is a good example of regiodivergent parallel kinetic resolution (PKR) because the racemic substrate (+)-(1R,5S)/(–)-(1S,5R) leads to two regioisomeric compounds (–)-(1S,5R) and (–)-(1R,5S) (Fig. 4.6). In fact, incubation of this racemic bicyclic ketone, with *Acinetobacter* strains yields the two regioisomeric lactones with almost total stereoselectivity; one originates from a "normal" Baeyer-Villiger-type oxygen insertion between the more substituted carbon atom and the carbonyl group, the second is formed with the chemically disfavoured regiochemistry [2]. The chiral lactones pro-

duced are relevant building blocks for the chemical synthesis of prostaglandins. Thus, CHMO is able to produce equimolar amount of the "normal" (–)-(1S,5R) and "abnormal" (–)-(1R,5S) lactones in nearly enantiomerically pure form (*ee* 95%). CPMO instead yields mainly "normal" lactones (–)-(1S,5R) and (+)-(1R,5S) with low enantioselectivity. The conversion of this racemic substrate by HAPMO generates (–)-(1S,5R) and (+)-(1S,5R) lactones in equal amounts starting from (–)-(1R,5S) enantiomer, whereas it has shown more regioselective against the other enantiomer (+)-(1S,5R). The final enantioselectivity is not high enough, resulting in 77% *ee* for (–)-(1S,5R) and 34% *ee* for (–)-(1R,5S) lactones [55].

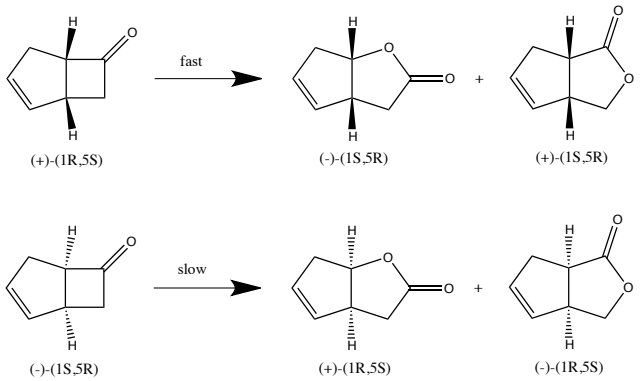


Figure 4.7: Possible structures of bicyclo[3.2.0]hept-2-en-6-one isomers and lactones formed by type I BVMO activity.

The behaviour of Cm towards the same racemic substrate (Fig. 4.8) displays, as for HAPMO, a significant preference for (+)-(1R,5S) enantiomer. But in contrast to CHMO, the regioselectivity is not achieved in a good percentage, 73% for the expected (–)-(1S,5R) and 27% for the unexpected (+)-(1S,5R), and the same is for the other enantiomer getting 88% for (+)-(1R,5S) and 12% for (–)-(1R,5S), finally obtaining the formation of all the four possible lactone products in different amounts. "Normal" lactones for both, the fast and slow reactions, are mainly produced. The *ee* values of the product mixture resulting from complete conversion are 12% for the "normal" (+)-(1R,5S) and 37% for the "abnormal" (+)-(1S,5R). After 5 hours the racemic substrate is fully converted. Pp and PpY173H show approximately the same behaviour (Fig. 4.8). The favoured substrate is the (+)-(1R,5S) enantiomer, mainly producing 76%

and 72% of the expected (-)-(1S,5R) and 24% and 28% of the (+)-(1S,5R), respectively. PpY173H exhibits much higher regioselectivity in the conversion of the other (–)-(1S,5R) substrate yielding 95% for the normal lactone and 5% for the abnormal one, whereas for Pp the percentages are 62% and 4% for the expected and unexpected lactones produced considering that the conversion of such slow reaction after 16 hours is not complete but is 66% . Also in this case *ee* values for the "normal" and "abnormal" enantiomeric products are too low to be of practical significance (Table 4.5).

Reaction	Cm	Pp	PpY173H
Fast	Conv(%) = 100% Ratio 73:27	Conv(%) = 100% Ratio 76:24	Conv(%) = 100% Ratio 72:28
Slow	Conv(%) = 100% Ratio 88:12	Conv(%) = 100% Ratio 73:27	Conv(%) = 100% Ratio 73:27
	eeN%(1R,5S) 12% eeA%(1S,5R) 37%	eeN%(1S,5R) 7% eeA%(1S,5R) 69%	eeN%(1R,5S) 14% eeA%(1S,5R) 68%

Table 4.5: Regiodivergent biooxidation by Cm, Pp and PpY173H.
Conv (%) = conversion
Ratio between normal and abnormal lactone

Cm, Pp and PpY173H are able to oxidize other compounds like long chain alkanones, phenylacetone and cyclic ketones (Table 4.6). Conversion of these ketones resulted in the formation of the expected ester product, obeying the stereochemical rule of Baeyer-Villiger reactions where the migrating group is the most substituted. Pp and PpY173H give poor conversion of cyclopentanone and almost no conversion of cyclohexanone. Cm is able to convert in a good amount the latter compound but it cannot transform cyclopentanone to the corresponding ε-caprolactone. In the case of 2-octanone, 3-octanone and 4-octanone, all of them are fully converted by Cm to the expected esters. PpY173H performs full conversion of 2-octanone and 3-octanone and poor transformation of 4-octanone whereas Pp is able to oxidize in good amount only the first two linear ketones. This result is found to be in line with the substrate profiling study (Table 4.1)

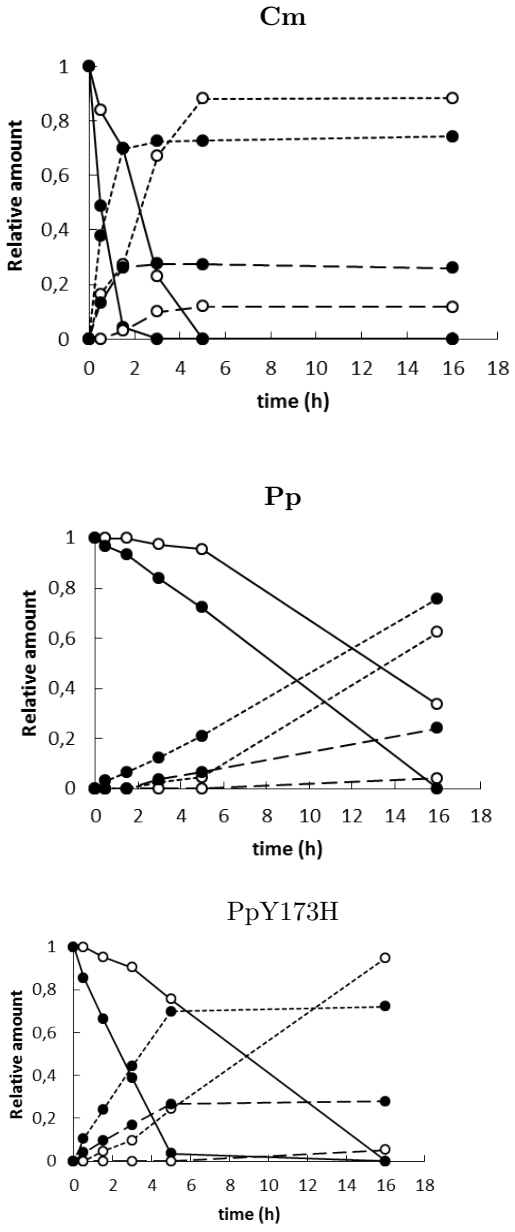


Figure 4.8: Time course of Cm (A), Pp (B) and PpY173H (C) catalyzed oxidation of racemic bicyclo[3.2.0]hept-2-en-6-one (–)-(1R,5S)/(+)-(1S,5R) to 2-oxabicyclo[3.3.0]oct-6-en-3-one (–)-(1S,5R)/(+)-(1R,5S) and 3-oxabicyclo[3.3.0]oct-6-en-2-one (+)-(1S,5R)/(–)-(1R,5S).

and also with the steady-state kinetic analysis (Table 4.2 and Table 4.3). Also octanal has found to be a well accepted substrate for these enzymes but we were not able to detect neither substrate nor the corresponding product by GC analysis. Finally phenylacetone is fully converted by all the three enzymes.

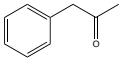
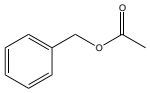
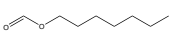
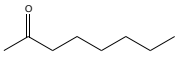
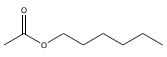
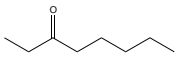
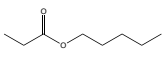
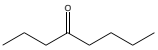
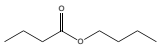
Substrate	Product	Conv (%)		
		Cm	Pp	PpY173H
		//	17%	27%
		80%	1%	3%
		100%	100%	100%
		ND	ND	ND
		100%	90%	100%
		100%	67%	100%
		100%	//	17%

Table 4.6: Conversion of some identified substrates and corresponding products.

4.7 UV-visible spectrum of Cm-BVMO

As described in paragraph 4.1, flavins and flavoproteins in their oxidized state are typically yellow with an absorption spectrum in the UV and visble spectra range characterized by two resolved bands at around 450 and 370 nm. In

particular, flavoproteins exhibit a pronounced spectral variability in term of position of the absorbance maxima, as well as their extinction coefficients and peak ratio, reflecting different environments of the flavin binding-site of the protein [78]. On the other hand, the overall shape of their UV-visible spectra (oxidized form) remains basically the same. By carefully observing the UV-vis spectrum of native Cm-BVMO we noticed a significant variation in its general shape (Fig. 4.4), which shows a higher peak around 450 nm and a nearly complete lack of minimum between both peaks resulting in a quite unusual and distinctive UV-visible spectrum. As just mentioned, the protein moiety perturbs the flavin spectrum and such variations can also be ascribed to interactions in the protein microenvironment; in fact the flavin cofactor can interact with the protein in many different ways. Usually it is tightly but noncovalently bound to the protein, however in some cases the cofactor is covalently linked to the polypeptide chain perturbing spectral features of the flavoprotein. Trying to investigate if some covalently bond could affect such spectrum perturbation, protein denaturation with TCA (trichloroacetic acid) was performed. Treatment using 5% TCA results in apoprotein precipitation with the removal of the noncovalently linked flavin remaining in the soluble fraction of the sample (Fig. 4.8).

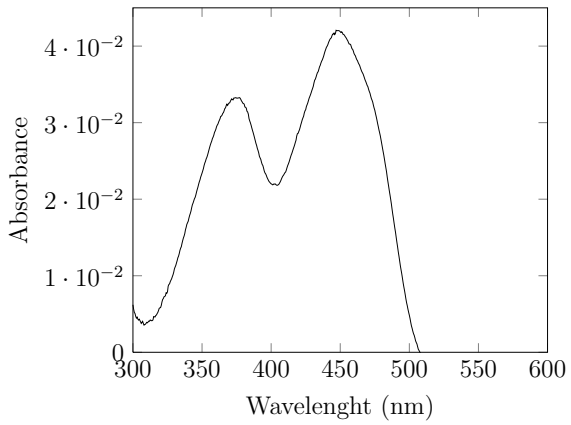


Figure 4.9: Uv-visible spectrum of the soluble fraction after protein denaturation with TCA (trichloroacetic acid) and containing the noncovalently linked flavin cofactor.

In the case of covalently bound cofactor, the flavin cofactor would coprecipitate together with the protein during the denaturation step, resulting in a yellow protein precipitate. This was clearly not the case for the studied BVMOs.

Variations in the flavin or in its protein environments can be also due to the binding of substrates, substrate analogues and inhibitors [67]. Titrations with the compound of interest can yield valuable information concerning changes in the active site environment.

As elucidated by Sheng et al. for CHMO from *Acinetobacter* sp. [100] and by Torres Pazmiño for PAMO [79], spectral changes of the flavin are observed upon binding of $NADP^+$. Thus, titration of 12 μM Cm-BVMO with 0-100 μM $NADP^+$ was investigated by monitoring changes in the UV-visible region of the flavin absorption spectrum (Fig. 4.9). To determine at which wavelengths big differences were presented, difference spectra obtained by subtraction of the wild-type Cm-BVMO spectrum and the recorded ones at different $NADP^+$ concentrations were calculated and plotted (Fig. 4.10). The largest decrease in absorbance was observed at 395 nm. In fact, experimental data indicate that the addition of 5 μM $NADP^+$ is sufficient for a more pronounced depression between to the two typical peaks (Fig 4.9) resulting in a better-resolved flavin spectrum. Moreover, we also noticed that the variation of the spectrum is achieved by just adding 5 μM $NADP^+$ which is around the half of the enzyme concentration (12 μM). By considering that spectra referred to other $NADP^+$ concentrations do not show relevant variations, it might be hypothesized that the oxidized Cm-BVMO could still contain some bound $NADP^+$.

Nevertheless, by analysing this preliminary study we can surely conclude that the affinity of the enzyme for $NADP^+$ is relatively high. This is in line with determined low K_M value for NADPH. More a detailed kinetic analysis it would be interesting to better understand if there is any role of the oxidized coenzyme concerning changes around the flavin or in the active site environment.

Concluding remarks

In the reported study, a biocatalytic characterization of the newly discovered type I BVMOs Cm, Pp from photosynthetic eukaryotes and the corresponding mutant PpY173H was carried out. Optimization of reaction conditions in terms of optimal pH and temperature values were performed. Also the thermodynamic stability measured as protein melting temperature (Tm) was investigated. *Thermo*FAD assay was used to determine Tm values for all the three enzymes revealing that Cm represent a thermostable type I BVMO. The substrate pro-

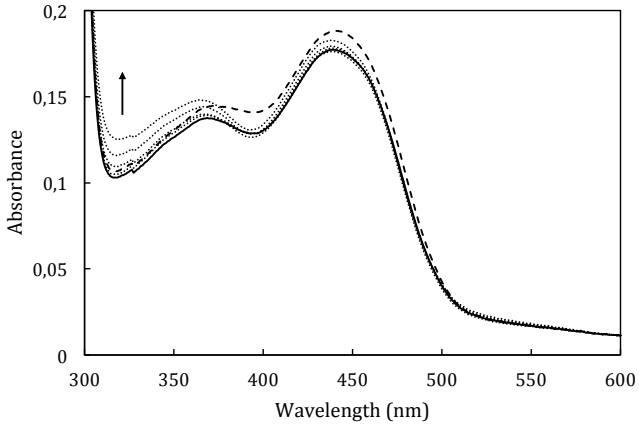


Figure 4.10: Flavin absorbance spectra of 12 μM Cm-BVMO upon titration with $NADP^+$. The concentration of $NADP^+$ is varied from 0 to 100 μM.

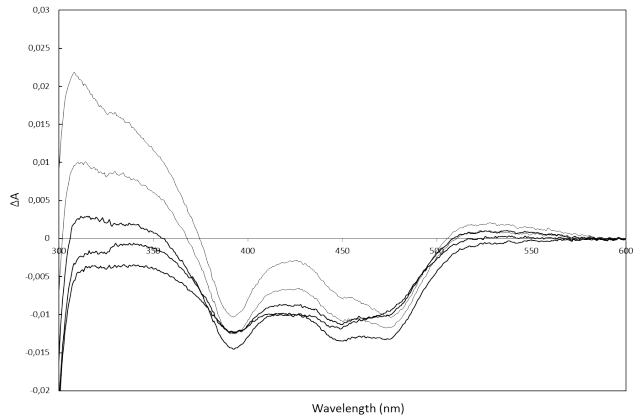


Figure 4.11: Difference spectra of Cm-BVMO upon titration with $NADP^+$ using the spectrum of Cm-BVMO with no $NADP^+$ as reference.

files were found to be quite versatile. The BVMOs can be used to catalyse the Baeyer-Villiger reaction of aliphatic, arylaliphatic and bicyclic ketones. In particular linear aliphatic ketones (C9 and C12), carring the keto functionality in different positions, have been the best substrates in steady state kinetic analysis and together with phenylacetone the best converted in enzymatic assay and bioconversions. Biooxidation of the well-known bicyclo[3.2.0]hept-2-en-6-one led to the production of all four possible lactone products for the three BVMOs, and the observed *ee* values suggest poor enantioselectivity towards this racemic substrate. The "normal" lactone was the main product for both, fast and slow reactions. The studies enzymes were found to be rather slow. Such moderate/low k_{cat} and $k_{\mathrm{cat}}/K_{\mathrm{M}}$ values are typical of enzymes operating in secondary metabolism subjected to a weak selective pressure. The enzymes may be involved in modifying some special secondary metabolites. Alternatively, the current study did not include compounds that resemble the physiological substrate for the discovered BVMOs.

Chapter 5

Summary and Conclusions

5.1 Summing up

Biocatalysis utilizes enzymes or whole cells as natural catalysts in chemical transformations. Since several decades it has played a central role in turning chemical processes green, to realize a more sustainable chemistry. In fact, the aim of "Green Chemistry" is the design of products and processes that minimize the use and generation of hazardous substances. Compared to chemical processes, biocatalysis offers important advantages such as mild reaction conditions, which minimizes problems of undesired side-reactions, and complete degradation in the environment. The most significant advantage is the high selectivity of enzymes, a keyword in organic synthesis because of the necessity to obtain enantiopure compounds. For these reasons chemical industry has progressively revealed its interest in biocatalysis, which is becoming a useful tool to design new processes for the production of fine chemicals, more compatible and respectful for the natural environment.

Monooxygenases are a class of oxidizing enzymes considerably appealing for industrial uses because of the reaction they are able to catalyze. It consists in the transfer of one oxygen atom from molecular oxygen to an organic substrate, reducing the other oxygen atom to water. In particular, Baeyer-Villiger monooxygenases (BVMOs) can perform oxidation of a ketone to an ester (Baeyer-Villiger oxidation) but also sulfoxidations and other oxidation reactions, often with good chemo-, regio- and enantio-selectivity. The latter feature is impossible to be obtained by the equivalent chemical methods. From an industrial point of view, the great potential of these enzymes is the possibility to produce chiral lactones and sulfoxides by asymmetric synthesis, resulting attractive for organic synthesis. Baeyer-Villiger monooxygenases (BVMOs) contain a flavin cofactor, FAD or FMN, which is tightly bound to the enzyme. The main described and characterized enzymes are type I BVMOs, containing FAD and using NADPH as electron donor. They consist in a single polypeptide chain having two Rossman-fold motifs responsible for the binding of FAD and NADPH. In addition, type I BVMOs contain a precise motif (FxGxxxHxxxWP/D) that is a conserved BVMO-specific sequence used for the identification of novel type I BVMO genes by the genome-mining approach.

Because of the development of bioinformatics tools and the increase of sequenced genomes in public database, it is possible to discover new enzymes by simply performing an in silico search using the program BLAST, which finds conserved regions among sequences. Till last year, all identified putative BVMO genes originated from bacteria or fungi. The first eukaryotic BVMO cloned and expressed has been cycloalkalone monooxygenases from the ascomycete *Clyndrocarpon radicicola* ATCC 11011, in 2012 [63]. Up to now, no BVMO has been found in archea or higher eukaryotes. By genome mining using PAMO, which is the best-known and characterized type I BVMOs, as bait sequence and the "fingerprint" motif FxGxxxHxxxWP/D as discriminant, we have identified and selected five new putative BVMOs from different organisms: *Oryza sativa* (plant), *Physcomitrella patens* (moss), *Cyanidioschyzon merolae* (red alga), *Trichodesmium erythraeum* (cyanobacterium), *Haloterrigena turkmenica* (archeabacterium). In particular, basing on most updated literature, the photosynthetic eukaryotes *Oryza, Physcomitrella* and *Cyanidioschyzon* would result very uncommon sources for BVMOs.

To obtain the identified putative type I BVMOs in recombinant form, a cloning strategy was set up and was the same for all the five sequences. The expression vectors were created starting from the same pET-28a(+) plasmid, in which the different Ht-BVMO, Te-BVMO, Os-BVMO, Cm-BVMO, Pp-BVMO coding sequences were inserted. Such sequences were obtained by two subsequent PCR reactions, the first one producing a preliminary amplimer, used for the second mutagenic amplification introducing proper restriction sites. The only difference between the five sequences was the starting material, genomic DNA for Te-BVMO, Cm-BVMO, Pp-BVMO, and purchased plasmid DNA for Ht-BVMO and Os-BVMO. The obtained constructs were transformed in *E. coli* BL21 (DE3) and overexpression tests were performed: in order to obtain soluble proteins, parameters such as expression temperature, IPTG concentration and suitable strains were varied. Paradoxically, the very high level of expression obtained for all apart one BVMOs resulted in a disadvantage. In fact, it exacerbated the insolubility of the recombinant proteins, which was the main bottleneck encountered in the whole work. Different arrangements were tried to slow down translation, the classical strategy adopted to favor protein folding, but no one

was fully effective. On the contrary, addition in the culture medium of an excess of riboflavin, the precursor of the FAD, was successful in order to obtain soluble recombinant proteins, at least for Pp-BVMO (*Physcomitrella patens*) and Cm-BVMO (*Cyanidioschyzon merolae*). Both of them were successfully purified by IMAC chromatography. Moreover, a mutant form of Pp-BVMO with a restored histidine (H) in the "fingerprint" motif FxGxxxHxxxWP/D (a tyrosine in the wild-type sequence) was obtained by site directed mutagenesis.

A biocatalytic characterization of the new photosynthetic eukaryotic type I BVMOs, Cm-BVMO, Pp-BVMO and the mutant PpY173H-BVMO, was carried out. The optimal pH found is 8.5 for Cm- and 8.0 for Pp- and PpY173H- BVMOs. While the latter two enzymes have a temperature optimum around 25°C, Cm-BVMO shows high resistance to temperature. The *Termo*FAD assay used to determine *Tm* values revealed that the recombinant Cm-enzyme represents a considerable thermostable type I BVMO. To evaluate the synthetic potential of these new type I BVMOs a broad range of substrates was used in a preliminary phosphate assay. Substrate profiles seems to be versatile; the enzymes can oxidize linear and cyclic aliphatic ketones, aryl and aromatic ketones. Linear ketones (C8 and C12) have proven to be the best substrates and steady-state analysis and conversions confirmed this result. Racemic bicyclo[3.2.0]hept-2-en-6-one, typically used to probe the biocatalytic potential of BVMOs, was oxidized by Cm-, Pp- and PpY173H- BVMOs giving the four possible lactone products with poor enantioselectivity. By steady-state analysis, it was observed that kinetic parameters such as K_{M}, k_{cat} and $k_{\mathrm{cat}}/K_{\mathrm{M}}$ displayed considerable low values suggesting a potential role as secondary metabolism enzymes. Probably the selective pressure applied on them is not enough strong, resulting in moderate/low k_{cat} and $k_{\mathrm{cat}}/K_{\mathrm{M}}$ values. Also for this reason their natural substrates remain yet to be identified.

Overall, Baeyer-Villiger monooxygenases (BVMOs) are useful oxidizing enzymes important for organic synthesis. We have discovered new type I BVMOs from very unusual sources, contributing to increase the number of available recombinant type I BVMOs and to enlarge the toolbox for their synthetic utility.

5.2 Outlooks

The presented work has brought to the discovery of new type I BVMOs able to perform oxidation reactions of chemical interest. Being the first discovered type I BVMOs from photosynthetic eukaryotes, natural substrates and reactions are yet unknown. Nevertheless, the performed biocatalytic characterization has given us some indications about their substrate profile and efficiency as biocatalysts. To gain more information about their structure-to-function relationships, 3D protein structures would be very helpful. Obtaining crystals from the obtained recombinant proteins would be a fundamental goal, since the structure of proteins gives much more insight in their function: it reveals how they are built and how they work, and allows prediction for their reactivity and for other properties. Knowledge of positions and interactions between amino acids residues, for example in the active site of an enzyme, would be a good starting point for protein engineering, in order to improve substrate acceptance, selectivity and catalytic efficiency. Indeed, native enzymes hardly fit the demanding requests imposed by synthetic applications; enzyme engineering is usually the solution adopted to make them suitable for required properties. Many approaches have been developed to allow the production of engineered enzymes having useful qualities such as increased activity, modified specificity or selectivity [14]. Rational design was one of the first approaches put forward to obtain desired changes: it consists in site-directed mutagenesis of chosen amino acids and obviously requires a detailed knowledge of the structure and function of the target protein. During the last two decades directed evolution has became a more attractive strategy for the modification of enzyme properties. It initially consists in generating molecular diversity by subsequent rounds of random mutagenesis and in identifying improved enzyme variants by high-throughput screening or selection after each round. The advantage of directed evolution is that it does not require knowledge of the protein. However, actual exploitation of directed evolution is strictly dependent on considerable instrumental and energy resources, to achieve the most stable, selective and productive biocatalyst. Recently, a semi-rational method to evolve enzymes has been developed consisting in a rational approach combined with stochastic elements [23, 66]. By utilizing information on protein sequence and structure,

it is possible to preselect promising target sites and to exhaustively modify only identified residues. There are many examples of protein-engineering for BVMOs demonstrating that activity, substrate range, stability and selectivity can be successfully improved [53].

Moreover, exploitation of the genetic properties of *Phiscomitrella patens* could be helpful for better understanding the functional role of the native Pp-BVMO. As previously reported, *Phiscomitrella patens* is a model organism used for studies in plant evolution, development and physiology. It displays a highly efficient homologous recombination, meaning that an exogenous DNA sequence can be targeted to a specific genomic position (by gene targeting) to create knockout mosses. This approach represents a powerful method to study the function of a gene [95, 94].

Finally, obtaining the crystallographic structure of the new Cm-, Pp- and PpY173H- type I BVMOs could enable a better understanding of structural, functional and phylogenetic relationships. More information concerning all these aspects could help to elucidate their natural role and facilitate their enhancements by protein engineering. In general, we think that all mentioned efforts could help in characterizing the new Baeyer-Villiger monooxygenases and in planning their transformation into more versatile enzymes, able to meet the strict demands of specific applications in organic synthesis.

Chapter 6

Materials and Methods

6.1 Materials

6.1.1 Primers

P. patens	
	sequence 5' - 3'
Pp3 for	ACAGGCCACGGGGGTAGTTCTGTTG
Pp4 rev	CAACCCTGGACAGCATCGGAAGCCT
Pp1 for	AAGTATGTCCAATTCCATGGCTGAGTTCGATGCTGTTATAGTCGGAG
Pp2 rev	AAACAATGCCCGCGGCCGCCAGCTTGAATCCC
Pp5 seq for	AAGGCTCATCGTACTACACGGG

Oryza sativa	
	sequence 5' - 3'
sp6 for	GATTTAGGTGACACTATAG
T7 term	GCTAGTTATTGCTCAGCGG
Os 1 for	CACACCCACCATGGCCCAGCATTACGACG
Os 2 rev	CTTGTTCAGGCGGCCGCGAGGTCGTATCC
Os3 seq for	TTTGTTCGCACGCCCAACCT

C. merolae	
	sequence 5' - 3'
Cm3 for	AGTGATGCGCGTGGCCGGCA
Cm4 rev	AGGTGTCTGCACCTCGCCAGCG
Cm1 for	TTTGACCGGCCATGGGAGCGGAGCTCAAC
Cm2 rev	GCATCCACCGCGCCGGCGTACAGCGAAG
Cm 5 seq for	ATCTTCCACACCGCTCGTTGG

T. erythraeum.	
	sequence 5' - 3'
Te5 for	AGCCTAGTTCCTGATTGGGTTGGTT
Te6 rev	GGCCTGAGGAGAACACATACAAGCT
Te1 for	TTGTTAGAGCAATTACCATGGATAATCAAATTTCTAGTAACCC
Te2 rev	AAAGTCTTTGGCGGCCGCTTGAGAAGTTGTGAG
Te8 for	CGTAGCGCGCCGTGGATCGCACC
Te9 rev	GGTGCGATCCACGGCGCGCTACG
Te7 seq for	TGGCGTATTTATAGTGCCTCTGGG

H. turkmenica	
	sequence 5' - 3'
Ht1 seq for	CGTGTTGGTCTGGTTGGTAC

6.1.2 Nutrition media

Nutrition media were autoclaved for 20 min at 121°C and agar (1.6% w/v) was added in order to prepare solid media. Antibiotics and others non autoclavable additives were filtrated (0.2 μm, Minisart, Sartorius).

E. coli

LB-medium (pH 7.0)
10 g tryptone
5 g yeast extract
10 g NaCl

TB-medium (pH 7.0)
12 g peptone
24 g yeast extract
5 g glycerol
filled up to 900 ml with water

Buffer for TB-Medium
125.4 g K_2HPO_4 (0.72 M)
23.1 g KH_2PO_4 (0.17 M)

Cyanidioschizon merolae

M-Allen

$(NH_4)_2SO_4$	262 mg
KH_2PO_4	54 mg
$MgSO_4 \cdot 7H_2O$	50 mg
$CaCl_2 \cdot 2H_2O$	14 mg
A2 trace element stock solution	0.2 mL
Distilled water	99.4 mL
	pH 2.5 1)

After autoclaving, add 0.4 mL of A2 Fe stock solution (filter-sterilized).
1) Adjust to pH 2.5 using 0.5mol/L H_2SO_4.
A2 trace element stock solution

H_3BO_3	285 mg
$MnCl_2 \cdot 4H_2O$	180 mg
ZnCl2	10.5 mg
$Na_2MoO4 \cdot 2H_2O$	39 mg
$CoCl_2 \cdot 6H_2O$	4 mg
$CuCl_2 \cdot 2H_2O$	4.3 mg
Distilled water	100 mL

A2 Fe stock solution

EDTA · 2Na	700 mg
$FeCl_3 \cdot 6H_2O$	400 mg
Distilled water	100 mL

Phiscomitrella patens

Nutrition media were autoclaved for 20 min at 121°C and plant agar (7 g/l) was added in order to prepare solid media. Buffer Phosphate (1000X): 25 g di

$PpNH_4$ rich medium (1liter in milliQ water):	
$CaNO_3$. $4H_2O$	0.8g/l
$MgSO_4$. $7H_2O$	0.25g/l
$FeSO_4$. $7H_2O$	0.0125g/l
Phosphate buffer (1000X)	1 ml
Microelements stoks (1000X)	1 ml
NH4 tartrate	500mg/l
Glucose	5g/l
Microelements stock (1000X):	
$CuSO_4$. $5H_2O$	55 mg/l
$ZnSO_4$. $7H_2O$	55 mg/l
H_3BO_3	614 mg/l
$MnCl_2$. $4H_2O$	389 mg/l
$CoCl_2$. $6H_2O$	55 mg/l
KI	28 mg/l
Na_2MoO_4 . $2H_2O$	25 mg/l

KH_2PO_4 are dissolved in 100 ml of milliQ water; the solution is adjusted to pH 7 using 4M KOH.

6.2 Methods

6.2.1 Reagents and Enzymes

All chemicals were purchased from Fluka and Sigma-Aldrich (Milan, Italy). Restriction enzymes were obtained from New England Biolabs (NEB) or Promega. Phusion(^{TM}) High-fidelity DNApolymerase was from Finnzymes. TSAP Thermosensitive Alkaline Phosphatase was purchased from Promega, whereas T4 DNA Ligase from NEB. IMAC-Select Affinity Gel resin was from Sigma-Aldrich.

6.2.2 PCR protocols

Oryza sativa

First PCR reaction (PCR1 Os)

Phusion HC Buffer 5X	10 μl		Temp (°C)	time
dNTPs 10mM	1 μl	Initial denaturation	98	30sec
Phusion DNA polymerase	0,5 μl	Denaturation	98	10sec
Sp6 10 μM	2,5 μl	Annealing	50	30sec
T7 term 10 μM	2,5 μl	Extension	72	1min
pCMVFL3	0,3 μl (60 ng)	Final extension	72	10min
H_2O	33,2 μl			
Total volume	50 μl			

Mutagenic PCR reaction (PCR2 Os)

Phusion HC Buffer 5X	10 μl		Temp (°C)	time
dNTPs 10mM	1 μl	Initial denaturation	98	30sec
Phusion DNA polymerase	0,5 μl	Denaturation	98	10sec
Os1for 10 μM	2,5 μl	Annealing	57	30sec
Os2 rev 10 μM	2,5 μl	Extension	72	1min
PCR1 Os	0,5μl	Final extension	72	10min
H_2O	33 μl			
Total volume	50 μl			

Phiscomitrella patens

First PCR reaction (PCR1 Pp)

Phusion HC Buffer 5X	10 μl		Temp (°C)	time
dNTPs 10mM	1 μl	Initial denaturation	98	30sec
Phusion DNA polymerase	0,5 μl	Denaturation	98	10sec
Pp3 for 10 μM	2,5 μl	Annealing	66	30sec
Pp4 rev 10 μM	2,5 μl	Extension	72	1min
Genomic DNA	0,5 μl (22 ng)	Final extension	72	10min
H_2O	33μl			
Total volume	50 μl			

Mutagenic PCR reaction (PCR2 Pp)

Phusion HC Buffer 5X	10 μl		Temp (°C)	time
dNTPs 10mM	1 μl	Initial denaturation	98	30sec
Phusion DNA polymerase	0,5 μl	Denaturation	98	10sec
Pp1for 10 μM	2,5 μl	Annealing	66	30sec
Pp2 rev 10 μM	2,5 μl	Extension	72	1min
PCR1 Pp	0,5 μl	Final extension	72	10min
H_2O	33 μl			
Total volume	50 μl			

Cyanidioschizon merolae

First PCR reaction (PCR1 Cm)				
Phusion HC Buffer 5X	10 μl		Temp (°C)	time
dNTPs 10mM	1 μl	Initial denaturation	98	30sec
Phusion DNA polymerase	0,5 μl	Denaturation	98	10sec
Cm3 for 10 μM	2,5 μl	Annealing	64	30sec
Cm4 rev 10 μM	2,5 μl	Extension	72	1min
Genomic DNA	0,5 μl (100 ng)	Final extension	72	10min
H_2O	33 μl			
Total volume	50 μl			

Mutagenic PCR reaction (PCR2 Cm)				
Phusion HC Buffer 5X	10 μl		Temp (°C)	time
dNTPs 10mM	1 μl	Initial denaturation	98	30sec
Phusion DNA polymerase	0,5 μl	Denaturation	98	10sec
Cm1for 10 μM	2,5 μl	Annealing	72	30sec
Cm2 rev 10 μM	2,5 μl	Extension	72	90sec
PCR1 Cm	0,5 μl	Final extension	72	10min
H_2O	33 μl			
Total volume	50 μl			

Trichodesmium erythraeum

First PCR reaction (PCR1 Te)				
Phusion HC Buffer 5X	10 μl		Temp (°C)	time
dNTPs 10mM	1 μl	Initial denaturation	98	30sec
Phusion DNA polymerase	0,5 μl	Denaturation	98	10sec
Te5 for 10 μM	2,5 μl	Annealing	57	30sec
Te6 rev 10 μM	2,5 μl	Extension	72	1min
Genomic DNA	0,5 μl (27 ng)	Final extension	72	10min
H_2O	33 μl			
Total volume	50 μl			

PCR reaction (Te1-Te9)				
Phusion HC Buffer 5X	10 μl		Temp (°C)	time
dNTPs 10mM	1 μl	Initial denaturation	98	30sec
Phusion DNA polymerase	0,5 μl	Denaturation	98	10sec
Te1 for 10 μM	2,5 μl	Annealing	61	30sec
Te9 rev 10 μM	2,5 μl	Extension	72	90sec
PCR1 Te	0,5 μl	Final extension	72	10min
H_2O	33 μl			
Total volume	50 μl			

PCR reaction (Te8-Te2)				
Phusion HC Buffer 5X	10 μl		Temp (°C)	time
dNTPs 10mM	1 μl	Initial denaturation	98	30sec
Phusion DNA polymerase	0,5 μl	Denaturation	98	10sec
Te8 for 10 μM	2,5 μl	Annealing	61	30sec
Te2 rev 10 μM	2,5 μl	Extension	72	90sec
PCR1 Te	0,5 μl	Final extension	72	10min
H_2O	33 μl			
Total volume	50 μl			

PCR reaction (Te1-Te9)				
Phusion HC Buffer 5X	10 μl		Temp (°C)	time
dNTPs 10mM	1 μl	Initial denaturation	98	30sec
Phusion DNA polymerase	0,5 μl	Denaturation	98	10sec
Te1for 10 μM	2,5 μl	Annealing	61	30sec
Te9 rev 10 μM	2,5 μl	Extension	72	90sec
PCR1 Te	0,5 μl	Final extension	72	10min
H_2O	33 μl			
Total volume	50 μl			

PCR reaction (Final)				
Phusion HC Buffer 5X	10 μl		Temp (°C)	time
dNTPs 10mM	1 μl	Initial denaturation	98	30sec
Phusion DNA polymerase	0,5 μl	Denaturation	98	10sec
Te1for 10 μM	2,5 μl	Annealing	61	30sec
Te2 rev 10 μM	2,5 μl	Extension	72	90sec
PCR Te1-Te9 + PCR Te8-Te2	0,5 μl	Final extension	72	10min
H_2O	33 μl			
Total volume	50 μl			

Colony PCR

GoTaq 5x Reaction Buffer	10 μl		Temp (°C)	time
dNTPs 10mM	1μl	Initial denaturation	94	5min
Primer "seq" 10 μM	2,5 μl	Denaturation	94	1min
Primer T7 term 10 μM	2,5 μl	Annealing	54	45sec
GoTaq Promega (5u/μl)	0,4 μl	Extension	72	90sec
H_2O	33,6 μl	Final Extension	72	10min
Total volume	50 μl			

6.2.3 Expression

The pET28a constructs containing the relative BVMO gene were transformed into *E. coli* BL21 (DE3) or *E.coli* ArcticExpress. Preculture were carried out in 5 ml LB medium at 37°C containing 50 μg/ml of kanamycin. Big culture were carried out using *E. coli* BL21(DE3) in 1L LB medium with riboflavin addition (100 μM). The cells were grown in shacking incubator at 37°C to an optical density at 600 nm (OD600) of 0.4-0.6, then induce by the addition of IPTG to a final concentration of 0.2 mM and cultivated at 18°C over night. The, cells were

harvested by centrifugation (4°C, 10 min, 4500xg) and washed with Tris/HCl buffer (Tris 50 mM, pH 8.0). Cell disruption was carried out using French Press and crude extract was centrifugated (4°C, 20 min, 15000xg) to separate soluble e insoluble fractions. FAD cofactor (100 μM) was added at the crude extract before cell disruption. To verify the expression of the BVMO, samples were normalized by OD600, denatured with a SDS buffer, centrifugated, resuspended and used for SDS-PAGE analysis. To verify the identity of the protein also immunoblotting assay was performed using anti His-tag antibody.

6.2.4 Enzyme purification and concentration

Overexpressed BVMOs were purified using His6-Tag metal ion affinity chromatography. The soluble fractions obtained from 1l culture were incubated together with the column material for 1h at 4°C and then load on a 5 ml column (Biorad). The column was washed by gravitational flow, first with three column volumes of the 50 mM Tris/HCl and 300 mM of NaCl and, then twice with the same buffer with the addition of 5 mM of imidazole. The elution of the His-tagged BVMOs were performed with five column volumes of 50 mM Tris/HCl containing 300 mM of imidazole. Excess of imidazole was removed using an Econo-PAC 10DG desalting column (BioRad). The enzyme concentration was measured by calculation the concentration of free flavin in a solution of denatured protein. In fact the concentration of the flavoenzyme is in this case equal to the concentration of free flavin. Using the law of Lambert-Beer A=l·ε·[flavin] is possible calculate this concentration.

6.2.5 Activity assay, kinetics and *Termo*FAD

BVMO activity is usually determined spectrophotometrically by monitoring the decrease of NADPH at 340 nm (ε= 6.22 mM-1 cm-1). Reaction mixtures (1 ml) contained 50 mM Tris–HCl buffer pH 8.0, 100 μM NADPH, 0,4 μM of pure enzyme, and 10 μl of 1 mM of bicyclo[3.2.0]hept-2-en-6-one in dioxane. The reaction was started by adding the enzyme to the mixture. One unit of BVMO is defined as the amount of protein that oxidizes 1 μmol NADPH per minute.

Steady-state kinetic parameters of the different substrates were determined

using purified enzymes and using substrate concentrations ranging from 0 to 5 mM. Data were fitted using the Michaelis-Menten equation using the program SigmaPlot.

The unfolding temperatures, *(Tm)*, were determined using the *Thermo*FAD method. 30 μl of 1 mg/mL protein in Tris-HCl buffer (50 mM, pH 7.5) were loaded in a Real Time PCR machine (Eppendorf) fitted with a 470 – 543 nm excitation filter and a SYBR Green emission filter (523 – 543 nm). A temperature gradient from 20 to 90 °C was applied (1°C/min), and fluorescence data was recorded. A sigmoidal curve was obtained after plotting the fluorescence against the temperature. The unfolding temperature, *Tm*, is then determined as the maximum of the derivative of this sigmoidal curve [34].

6.2.6 Substrate screening

For the phosphate 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 was incubated at 25°C for 2 h. Twenty μl of the reaction mixture was than 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. By measuring the absorption at 600 nm, the concentration of produced phosphate is quantitatively determined using a calibration curve. Each conversion can be performed with a standard amount of BVMO allowing a comparison between samples. All substrates tested in 96 wells plates and were dissolved in dioxane (maximum 5%).

6.2.7 Conversions and GC/GC-MS analysis

For GC and GC-MS analysis, 500 μl incubations of 50 mM Tris–HCl (pH 7.5), containing 2 mM substrate, 5% dioxane, 100 μM NADPH, 3.0 μM PTDH, 10 mM phosphite and 1 μM BVMOs was incubated shaking at 25°C from 1h to 20 h. The reactions were then stopped by extracting with ethyl acetate (3 x 0.5 ml, including 0.1% mesitylene as an internal standard), dried with magnesium sulfate

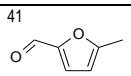
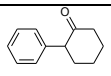
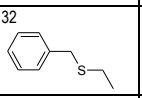
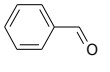
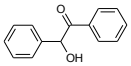
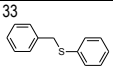
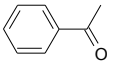
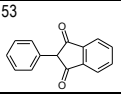
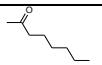
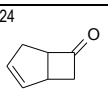
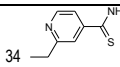
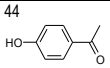
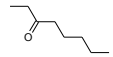
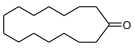
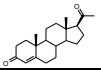
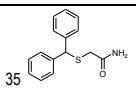
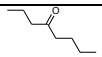
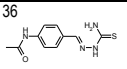
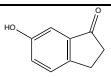
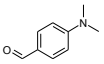
and analyzed directly by GC or GC-MS to determine the degree of conversion.

Columns used:

- GC-Helifex AT5
- Chirasil Dex CB
- GC-MS HP1

Compound	Column	program	t_R (min) substrate	t_R (min) product
Phenylacetone	AT5 (GC)	2min at 100°C 100°C→120°C 5°C/min 120°C→160°C 20°C/min	6.9	7.5
Cyclopentanone	AT5 (GC)	5min at 50°C 50°C→120°C 5°C/min 120°C→150°C 15°C/min	5.5	14.7
Cyclohexanone	AT5 (GC)	5min at 50°C 50°C→120°C 5°C/min 120°C→150°C 15°C/min	9.0	17.7
Bicyclohept-2-en-6-one	Chiraldex CB (GC)	35°C→130°C 10°C/min 12min at 130°C 130°C→35°C 10°C/min	9.0 9.2	16.5 A (1R,5S) 16.9 N (1R,5S) 17.2 A (1S,5R) 17.4 N (1S,5R)
Octanal	HP1 (GC-MS)	70°C→120°C 5°C/min 120°C→280°C 20°C/min	7.4	8.5
2-octanone	HP1 (GC-MS)	70°C→120°C 5°C/min 120°C→280°C 20°C/min	8.1	8.9
3-octanone	HP1 (GC-MS)	70°C→120°C 5°C/min 120°C→280°C 20°C/min	7.9	8.7
4-octanone	HP1 (GC-MS)	70°C→120°C 5°C/min 120°C→280°C 20°C/min	7.5	8.3

Substrate structures:

	1	2	3	4	5	6
A	 1	 11	 21	 31	 41	 51
B	 2	 12	 22	 32	 42	 52
C	 3	 13	 23	 33	 43	 53
D	 4	 14	 24	 34	 44	 54
E	5 	15 	25 	 35	 45	 55
F	 6	 16	26 	36 	 46	 56
G	7 	 17	27 	 37	47	dioxane
H	8	18	28	38	48	dioxane

Bibliography

[1] V. Alphand, A. Archelas, and R. Furstoss. Microbial transformations 16. one-step synthesis of a pivotal prostaglandin chiral synthon via a highly enantioselective microbiological baeyer-villiger type reaction. *Tetrahedron letters*, 30(28):3663–3664, 1989.

[2] V. Alphand and R. Furstoss. Microbiological transformations. 22. microbiologically mediated baeyer-villiger reactions: a unique route to several bicyclic. gamma.-lactones in high enantiomeric purity. *The Journal of organic chemistry*, 57(4):1306–1309, 1992.

[3] P. T. Anastas, L. B. Bartlett, M. M. Kirchhoff, and T. C. Williamson. The role of catalysis in the design, development, and implementation of green chemistry. *Catalysis Today*, 55(1):11–22, 2000.

[4] P. T. Anastas and M. M. Kirchhoff. Origins, current status, and future challenges of green chemistry. *Accounts of Chemical Research*, 35(9):686–694, 2002.

[5] P. T. Anastas, M. M. Kirchhoff, and T. C. Williamson. Catalysis as a foundational pillar of green chemistry. *Applied Catalysis A: General*, 221(1):3–13, 2001.

[6] P. T. Anastas and T. C. Williamson. Green chemistry: an overview. In *ACS Symposium Series*, volume 626, pages 1–19. ACS Publications, 1996.

[7] C. Baalen and R. M. Brown. The ultrastructure of the marine blue green alga, *Trichodesmium erythraeum*, with special reference to the cell wall,

gas vacuoles, and cylindrical bodies. *Archives of Microbiology*, 69(1):79–91, 1969.

[8] A. Baeyer and V. Villiger. *Ber. Dtsch. Chem. Ges.*, 32:3625–3633, 1899.

[9] M. P. Beam, M. A. Bosserman, N. Noinaj, M. Wehenkel, and J. Rohr. Crystal structure of baeyer-villiger monooxygenase mtmoiv, the key enzyme of the mithramycin biosynthetic pathway. *Biochemistry*, 48(21):4476, 2009.

[10] E. Beneventi, G. Ottolina, G. Carrea, W. Panzeri, G. Fronza, and P. C. K. Lau. Enzymatic baeyer-villiger oxidation of steroids with cyclopentadecanone monooxygenase. *Journal of Molecular Catalysis B: Enzymatic*, 58(1):164–168, 2009.

[11] N. Berezina, V. Alphand, and R. Furstoss. Microbiological transformations. part 51: The first example of a dynamic kinetic resolution process applied to a microbiological baeyer-villiger oxidation. *Tetrahedron: Asymmetry*, 13(18):1953–1955, 2002.

[12] WJH Van Berkel, NM Kamerbeek, and MW Fraaije. Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *Journal of Biotechnology*, 124(4):670–689, 2006.

[13] UT Bornscheuer, GW Huisman, RJ Kazlauskas, S. Lutz, JC Moore, and K. Robins. Engineering the third wave of biocatalysis. *Nature*, 485(7397):185–194, 2012.

[14] Uwe T. Bornscheuer and Martina Pohl. Improved biocatalysts by directed evolution and rational protein design. *Current opinion in chemical biology*, 5(2):137–143, 4/1 2001.

[15] B. P. Branchaud and C. T. Walsh. Functional group diversity in enzymic oxygenation reactions catalyzed by bacterial flavin-containing cyclohexanone oxygenase. *Journal of the American Chemical Society*, 107(7):2153–2161, 1985.

[16] AF Buckmann, MR Kula, R. Wichmann, and C. Wandrey. An efficient synthesis of high-molecular-weight nad (h) derivatives suitable for continuous operation with coenzyme-dependent enzyme systems. *J.Appl.Biochem*, 3:301–315, 1981.

[17] S. G. Burton. Oxidizing enzymes as biocatalysts. *Trends in biotechnology*, 21(12):543–549, 2003.

[18] N. A. Campbell and J. B. Reece. Biology, international edition. 2002.

[19] D. G. Capone, J. P. Zehr, H. W. Paerl, B. Bergman, and E. J. Carpenter. *Trichodesmium*, a globally significant marine cyanobacterium. *Science*, 276(5316):1221–1229, 1997.

[20] G. Chen, M. M. Kayser, M. D. Mihovilovic, M. E. Mrstik, C. A. Martinez, and J. D. Stewart. Asymmetric oxidations at sulfur catalyzed by engineered strains that overexpress cyclohexanone monooxygenase. *New J.Chem.*, 23(8):827–832, 1999.

[21] H. K. Chenault, ES Simon, and GM Whitesides. Cofactor regeneration for enzyme-catalysed synthesis. *Biotechnol Genet Eng Rev*, 6:221–270, 1988.

[22] H. K. Chenault and G. M. Whitesides. Regeneration of nicotinamide cofactors for use in organic synthesis. *Applied Biochemistry and Biotechnology*, 14(2):147–197, 1987.

[23] R. A. Chica, N. Doucet, and J. N. Pelletier. Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design. *Current opinion in biotechnology*, 16(4):378–384, 2005.

[24] S. Colonna, V. Pironti, P. Pasta, and F. Zambianchi. Oxidation of amines catalyzed by cyclohexanone monooxygenase. *Tetrahedron letters*, 44(4):869–871, 2003.

[25] HE Conrad, R. DuBus, MJ Namtvedt, and IC Gunsalus. Mixed function oxidation. *Journal of Biological Chemistry*, 240(1):495–503, 1965.

[26] HE Conrad, R. DuBus, MJ Namtvedt, and IC Gunsalus. Mixed function oxidation. *Journal of Biological Chemistry*, 240(1):495–503, 1965.

[27] D. J. C. Constable, A. D. Curzons, and V. L. Cunningham. Metrics to 'green' chemistry - which are the best? *Green Chem.*, 4(6):521–527, 2002.

[28] D. J. C. Constable, A. D. Curzons, L. M. F. dos Santos, G. R. Geen, R. E. Hannah, J. D. Hayler, J. Kitteringham, M. A. McGuire, J. E. Richardson, and P. Smith. Green chemistry measures for process research and development. *Green Chemistry*, 3(1):7–9, 2001.

[29] R. E. Cripps. The microbial metabolism of acetophenone. metabolism of acetophenone and some chloroacetophenones by an *Arthrobacter species*. *Biochemical Journal*, 152(2):233, 1975.

[30] C. M. Crudden, A. C. Chen, and L. A. Calhoun. A demonstration of the primary stereoelectronic effect in the baeyer-villiger oxidation of α-fluorocyclohexanones. *Angewandte Chemie International Edition*, 39(16):2851–2855, 2000.

[31] M. J. Van der Werf. Purification and characterization of a baeyer-villiger mono-oxygenase from *Rhodococcus erythropolis* dcl14 involved in three different monocyclic monoterpene degradation pathways. *Biochemical Journal*, 347(Pt 3):693, 2000.

[32] W. E. Doering and E. Dorfman. Mechanism of the peracid ketone-ester conversion. analysis of organic compounds for oxygen-181. *Journal of the American Chemical Society*, 75(22):5595–5598, 1953.

[33] H. M. Dudek, G. de Gonzalo, D. E. Torres Pazmino, P. Stepniak, L. S. Wyrwicz, L. Rychlewski, and M. W. Fraaije. Mapping the substrate binding site of phenylacetone monooxygenase from *Thermobifida fusca* by mutational analysis. *Applied and Environmental Microbiology*, 77(16):5730–5738, 2011.

[34] F. Forneris, R. Orru, D. Bonivento, L. R. Chiarelli, and A. Mattevi. Thermofad, a *Thermofluor*®-adapted flavin ad hoc detection system for protein folding and ligand binding. *FEBS Journal*, 276(10):2833–2840, 2009.

[35] M. W. Fraaije and D. B. Janssen. *in Modern Biooxidation (Ed.: R.D. Schmid, V.B. Urlacher)*. Wiley-VCH, 2007.

[36] M. W. Fraaije, N. M. Kamerbeek, A. J. Heidekamp, R. Fortin, and D. B. Janssen. The prodrug activator etaa from *Mycobacterium tuberculosis* is a baeyer-villiger monooxygenase. *Journal of Biological Chemistry*, 279(5):3354–3360, 2004.

[37] M. W. Fraaije, J. Wu, D. P. H. M. Heuts, E. W. van Hellemond, J. H. L. Spelberg, and D. B. Janssen. Discovery of a thermostable baeyer-villiger monooxygenase by genome mining. *Applied Microbiology and Biotechnology*, 66(4):393–400, 2005.

[38] S. Franceschini, H. L. van Beek, A. Pennetta, C. Martinoli, M. W. Fraaije, and A. Mattevi. Exploring the structural basis of substrate preferences in baeyer-villiger monooxygenases. insight from steroid monooxygenase. *Journal of Biological Chemistry*, 287(27):22626–22634, 2012.

[39] S. A. Goff, D. Ricke, T. H. Lan, G. Presting, R. Wang, M. Dunn, J. Glazebrook, A. Sessions, P. Oeller, and H. Varma. A draft sequence of the rice genome (*Oryza sativa* l. ssp. japonica). *Science*, 296(5565):92–100, 2002.

[40] B. T. Greenhagen, P. E. O'Maille, J. P. Noel, and J. Chappell. Identifying and manipulating structural determinates linking catalytic specificities in terpene synthases. *Proceedings of the National Academy of Sciences*, 103(26):9826–9831, 2006.

[41] M. Griffin and P. W. Trudgill. The metabolism of cyclopentanol by *Pseudomonas* ncib 9872. *Biochemical Journal*, 129(3):595, 1972.

[42] J. Havel and W. Reineke. Microbial degradation of chlorinated acetophenones. *Applied and Environmental Microbiology*, 59(8):2706–2712, 1993.

[43] F. K. Higson and D. D. Focht. Bacterial degradation of ring-chlorinated acetophenones. *Applied and Environmental Microbiology*, 56(12):3678–3685, 1990.

[44] I. Hilker, M. C. Gutiérrez, V. Alphand, R. Wohlgemuth, and R. Furstoss. Microbiological transformations 57. facile and efficient resin-based in situ sfpr preparative-scale synthesis of an enantiopure "unexpected" lactone regioisomer via a baeyer-villiger oxidation process. *Organic letters*, 6(12):1955–1958, 2004.

[45] I. Hilker, R. Wohlgemuth, V. Alphand, and R. Furstoss. Microbial transformations 59: First kilogram scale asymmetric microbial baeyer-villiger oxidation with optimized productivity using a resin-based in situ sfpr strategy. *Biotechnology and bioengineering*, 92(6):702–710, 2005.

[46] F. Hollmann and A. Schmid. Electrochemical regeneration of oxidoreductases for cell-free biocatalytic redox reactions. *Biocatalysis and Biotransformation*, 22(2):63–88, 2004.

[47] F. Hollmann, A. Taglieber, F. Schulz, and M. T. Reetz. A light-driven stereoselective biocatalytic oxidation. *Angewandte Chemie International Edition*, 46(16):2903–2906, 2007.

[48] F. Hollmann, B. Witholt, and A. Schmid. $[cp^*rh(bpy)(h_2o)]^{2+}$: a versatile tool for efficient and non-enzymatic regeneration of nicotinamide and flavin coenzymes. *Journal of Molecular Catalysis B: Enzymatic*, 19:167–176, 2002.

[49] H. Iwaki, Y. Hasegawa, M. Teraoka, T. Tokuyama, H. Bergeron, and P. C. K. Lau. Identification of a transcriptional activator (chnr) and a 6-oxohexanoate dehydrogenase (chne) in the cyclohexanol catabolic pathway in *Acinetobacter* sp. strain ncimb 9871 and localization of the genes that encode them. *Applied and Environmental Microbiology*, 65(11):5158–5162, 1999.

[50] H. Iwaki, Y. Hasegawa, S. Wang, M. M. Kayser, and P. C. K. Lau. Cloning and characterization of a gene cluster involved in cyclopentanol metabolism in *Comamonas* sp. strain ncimb 9872 and biotransformations effected by *Escherichia coli*-expressed cyclopentanone 1, 2-monooxygenase. *Applied and Environmental Microbiology*, 68(11):5671–5684, 2002.

[51] E. S. Johnson and G. Blobel. Cell cycle-regulated attachment of the ubiquitin-related protein sumo to the yeast septins. *The Journal of cell biology*, 147(5):981–994, 1999.

[52] K. H. Jones, R. T. Smith, and P. W. Trudgill. Diketocamphane enantiomer-specific "baeyer-villiger" monooxygenases from camphor-grown *Pseudomonas putida* atcc 17453. *Journal of general microbiology*, 139(4):797–805, 1993.

[53] M. Kadow, K. Loschinski, H. Mallin, S. Saβ, and U. Bornscheuer. Discovery, application and protein engineering of baeyer-villiger monooxygenases for organic synthesis. *Organic & Biomolecular Chemistry*, 2012.

[54] N. M. Kamerbeek, D. B. Janssen, W. J. H. van Berkel, and M. W. Fraaije. Baeyer-villiger monooxygenases, an emerging family of flavin-dependent biocatalysts. *Advanced Synthesis & Catalysis*, 345:667–678, 2003.

[55] N. M. Kamerbeek, A. J. J. Olsthoorn, M. W. Fraaije, and D. B. Janssen. Substrate specificity and enantioselectivity of 4-hydroxyacetophenone monooxygenase. *Applied and Environmental Microbiology*, 69(1):419, 2003.

[56] D. M. Karl. Nutrient dynamics in the deep blue sea. *Trends in microbiology*, 10(9):410–418, 2002.

[57] K. M. Koeller and Chi-Huey Wong. Enzymes for chemical synthesis. *Nature*, 409:232–240, 2001.

[58] K. Kostichka, S. M. Thomas, K. J. Gibson, V. Nagarajan, and Q. Cheng. Cloning and characterization of a gene cluster for cyclododecanone oxidation in *Rhodococcus ruber* sc1. *Journal of Bacteriology*, 183(21):6478–6486, 2001.

[59] T. Kuroiwa. The primitive red algae *Cyanidium caldarium* and *Cyanidioschyzon merolae* as model system for investigating the dividing apparatus of mitochondria and plastids. *Bioessays*, 20(4):344–354, 1998.

[60] D. Lang, A. D. Zimmer, S. A. Rensing, and R. Reski. Exploring plant biodiversity: the *Physcomitrella* genome and beyond. *Trends in plant science*, 13(10):542–549, 2008.

[61] J. Lebreton, V. Alphand, and R. Furstoss. A short chemoenzymatic synthesis of (+)-multifidene and (+)-viridiene. *Tetrahedron letters*, 37(7):1011–1014, 1996.

[62] J. Lebreton, V. Alphand, and R. Furstoss. Chemoenzymatic synthesis of marine brown algae pheromones. *Tetrahedron*, 53(1):145–160, 1997.

[63] F. Leipold, R. Wardenga, and U. T. Bornscheuer. Cloning, expression and characterization of a eukaryotic cycloalkanone monooxygenase from *Cylindrocarpon radicicola* atcc 11011. *Applied Microbiology and Biotechnology*, pages 1–13, 2012.

[64] H. Leisch, R. Shi, S. Grosse, K. Morley, H. Bergeron, M. Cygler, H. Iwaki, Y. Hasegawa, and P. C. K. Lau. Cloning, baeyer-villiger biooxidations, and tructures of the camphor pathway 2-oxo-δ3-4, 5, 5-trimethylcyclopentenylacetyl-coenzyme a monooxygenase of *Pseudomonas putida* atcc 17453. *Applied and Environmental Microbiology*, 78(7):2200–2212, 2012.

[65] XJ Luo, HL Yu, and JH Xu. Genomic data mining: an efficient way to find new and better enzymes. *Enzyme Engg*, 1(104):2, 2012.

[66] S. Lutz. Beyond directed evolution-semi-rational protein engineering and design. *Current opinion in biotechnology*, 21(6):734, 2010.

[67] P. Macheroux. Uv-visible spectroscopy as a tool to study flavoproteins. *Methods in Molecular Biology-Clifton Then Totowa-*, 131:1–8, 1999.

[68] M. P. Malakhov, M. R. Mattern, O. A. Malakhova, M. Drinker, S. D. Weeks, and T. R. Butt. Sumo fusions and sumo-specific protease for efficient expression and purification of proteins. *Journal of structural and functional genomics*, 5(1):75–86, 2004.

[69] E. Malito, A. Alfieri, M. W. Fraaije, and A. Mattevi. Crystal structure of a baeyer-villiger monooxygenase. *Proceedings of the National Academy of Sciences of the United States of America*, 101(36):13157–13162, 2004.

[70] M. Matsuzaki, O. Misumi, T. Shin-i, S. Maruyama, M. Takahara, S. Miyagishima, T. Mori, K. Nishida, F. Yagisawa, and K. Nishida. Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10d. *Nature*, 428(6983):653–657, 2004.

[71] C. Mazzini, J. Lebreton, V. Alphand, and R. Furstoss. Enantiodivergent chemoenzymatic synthesis of ®-and(s)-β-proline in high optical purity. *Journal of organic chemistry*, 62(15):5215–5218, 1997.

[72] F. Melchior. Sumo-nonclassical ubiquitin. *Annual Review of Cell and Developmental Biology*, 16(1):591–626, 2000.

[73] M. D. Mihovilovic. Enzyme mediated baeyer-villiger oxidations. *Current Organic Chemistry*, 10(11):1265–1287, 2006.

[74] M. D. Mihovilovic, B. Müller, and P. Stanetty. Monooxygenase-mediated baeyer-villiger oxidations. *European Journal of Organic Chemistry*, 2002(22):3711–3730, 2002.

[75] M. D. Mihovilovic, B. Müller, and P. Stanetty. Monooxygenase-mediated baeyer-villiger oxidations. *European Journal of Organic Chemistry*, 2002(22):3711–3730, 2002.

[76] M. D. Mihovilovic, F. Rudroff, B. Grötzl, P. Kapitan, R. Snajdrova, J. Rydz, and R. Mach. Family clustering of baeyer-villiger monooxygenases based on protein sequence and stereopreference. *Angewandte Chemie*, 117(23):3675–3679, 2005.

[77] I. A. Mirza, B. J. Yachnin, S. Wang, S. Grosse, H. Bergeron, A. Imura, H. Iwaki, Y. Hasegawa, P. C. K. Lau, and A. M. Berghuis. Crystal structures of cyclohexanone monooxygenase reveal complex domain movements and a sliding cofactor. *Journal of the American Chemical Society*, 131(25):8848–8854, 2009.

[78] F. Müller, M. Brüstlein, P. Hemmerich, V. Massey, and W. H. Walker. Light-?absorption studies on neutral flavin radicals. *European Journal of Biochemistry*, 25(3):573–580, 1972.

[79] D. E. Torres Pazmi no, B. J. Baas, D. B. Janssen, and M. W. Fraaije. Kinetic mechanism of phenylacetone monooxygenase from *Thermobifida fusca*. *Biochemistry*, 47(13):4082–4093, 2008.

[80] H. Nozaki, M. Matsuzaki, M. Takahara, O. Misumi, H. Kuroiwa, M. Hasegawa, T. Shin-i, Y. Kohara, N. Ogasawara, and T. Kuroiwa. The phylogenetic position of red algae revealed by multiple nuclear genes from mitochondria-containing eukaryotes and an alternative hypothesis on the origin of plastids. *Journal of Molecular Evolution*, 56(4):485–497, 2003.

[81] N. Ohta, N. Sato, and T. Kuroiwa. The organellar genomes of *Cyanidioschyzon merolae*. *Enigmatic Microorganisms and Extreme Environments, J.Seckbach, ed (Dordrecht, The Netherlands: Kluwer Academic Publishers)*, pages 139–149, 1999.

[82] H. I. Oka. *Origin of cultivated rice.* Japan Scientific Societies Press, 1988.

[83] R. Orru, H. M. Dudek, C. Martinoli, D. E. T. Pazmi no, A. Royant, M. Weik, M. W. Fraaije, and A. Mattevi. Snapshots of enzymatic baeyer-villiger catalysis oxygen activation and intermediate stabilization. *Journal of Biological Chemistry*, 286(33):29284–29291, 2011.

[84] G. Ottolina, S. Bianchi, B. Belloni, G. Carrea, and B. Danieli. First asymmetric oxidation of tertiary amines by cyclohexanone monooxygenase. *Tetrahedron letters*, 40(48):8483–8486, 1999.

[85] S. Pan and B. A. Malcolm. Reduced background expression and improved plasmid stability with pet vectors in bl21 (de3). *BioTechniques*, 29(6):1234, 2000.

[86] S. Picaud, L. Olofsson, M. Brodelius, and P. E. Brodelius. Expression, purification, and characterization of recombinant amorpha-4, 11-diene syn-

thase from *Artemisia annua* l. *Archives of Biochemistry and Biophysics*, 436(2):215–226, 2005.

[87] J. Rehdorf, M. D. Mihovilovic, and U. T. Bornscheuer. Exploiting the regioselectivity of baeyer-villiger monooxygenases for the formation of β-amino acids and β-amino alcohols. *Angewandte Chemie International Edition*, 49(26):4506–4508, 2010.

[88] J. Rehdorf, M. D. Mihovilovic, M. W. Fraaije, and U. T. Bornscheuer. Enzymatic synthesis of enantiomerically pure β-amino ketones, β-amino esters, and β-amino alcohols with baeyer-villiger monooxygenases. *Chemistry-A European Journal*, 16(31):9525–9535, 2010.

[89] S. A. Rensing, D. Lang, A. D. Zimmer, A. Terry, A. Salamov, H. Shapiro, T. Nishiyama, P. F. Perroud, E. A. Lindquist, and Y. Kamisugi. The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science*, 319(5859):64–69, 2008.

[90] M. Renz and B. Meunier. 100 years of baeyer-villiger oxidations. *European journal of organic chemistry*, 1999(4):737–750, 1999.

[91] S. M. Roberts and P. W. H. Wan. Enzyme-catalysed baeyer-villiger oxidations. *Journal of molecular catalysis.B, Enzymatic*, 4(3):111–136, 1998.

[92] C. C. Ryerson, D. P. Ballou, and C. Walsh. Mechanistic studies on cyclohexanone oxygenase. *Biochemistry*, 21(11):2644–2655, 1982.

[93] E. Saunders, B. J. Tindall, R. Fähnrich, A. Lapidus, A. Copeland, T. G. Del Rio, S. Lucas, F. Chen, H. Tice, and J. F. Cheng. Complete genome sequence of *Haloterrigena turkmenica* type strain (4kt). *Standards in Genomic Sciences*, 2(1):107, 2010.

[94] D. G. Schaefer. A new moss genetics: targeted mutagenesis in *Physcomitrella patens*. *Annual review of plant biology*, 53(1):477–501, 2002.

[95] D. G. Schaefer and J. P. Zr'yd. Efficient gene targeting in the moss *Physcomitrella patens*. *The Plant Journal*, 11(6):1195–1206, 1997.

[96] A. Schmid, JS Dordick, B. Hauer, A. Kiener, M. Wubbolts, and B. Witholt. Industrial biocatalysis today and tomorrow. *Nature*, 409(6817):258, 2001.

[97] R. A. Sheldon. Fundamentals of green chemistry: efficiency in reaction design. *Chemical Society Reviews*, 41(4):1437–1451, 2012.

[98] R. A. Sheldon, I. Arends, and U. Hanefeld. *Green Chemistry and Catalysis*. Wiley-VCH, 2007.

[99] D. Sheng, D. P. Ballou, and V. Massey. Mechanistic studies of cyclohexanone monooxygenase: chemical properties of intermediates involved in catalysis. *Biochemistry*, 40(37):11156–11167, 2001.

[100] D. Sheng, D. P. Ballou, and V. Massey. Mechanistic studies of cyclohexanone monooxygenase: chemical properties of intermediates involved in catalysis. *Biochemistry*, 40(37):11156–11167, 2001.

[101] J. D. Stewart. Cyclohexanone monpoxygenase: a useful asymmetric baeyer-villiger reactions. *Current Organic Chemistry*, 2(3):195–216, 1998.

[102] F. W. Studier. Use of bacteriophage t7 lysozyme to improve an inducible t7 expression system. *Journal of Molecular Biology*, 219(1):37, 1991.

[103] A. Tanner and D. J. Hopper. Conversion of 4-hydroxyacetophenone into 4-phenyl acetate by a flavin adenine dinucleotide-containing baeyer-villiger-type monooxygenase. *Journal of Bacteriology*, 182(23):6565–6569, 2000.

[104] KD Tartoff and CA Hobbs. Improved media for growing plasmid and cosmid clones. *Bethesda Research Laboratories Focus*, 1987.

[105] M. H. Tatham, E. Jaffray, O. A. Vaughan, J. M. P. Desterro, C. H. Botting, J. H. Naismith, and R. T. Hay. Polymeric chains of sumo-2 and sumo-3 are conjugated to protein substrates by sae1/sae2 and ubc9. *Journal of Biological Chemistry*, 276(38):35368–35374, 2001.

[106] D. G. Taylor and P. W. Trudgill. Camphor revisited: studies of 2, 5-diketocamphane 1, 2-monooxygenase from *Pseudomonas putida* atcc 17453. *Journal of Bacteriology*, 165(2):489–497, 1986.

[107] S. Terui, K. Suzuki, H. Takahashi, R. Itoh, and T. Kuroiwa. Synchronization of chloroplast division in the ultramicroalga *Cyanidioschyzon merolae* (rhodophyta) by treatment with light and aphidicolin. *Journal of Phycology*, 31(6):958–961, 1995.

[108] PW Trudgill. Microbial degradation of the alicyclic ring: structural relationships and metabolic pathways. *Microbiology Series[MICROBIOL.SER.].1984.*, 1984.

[109] GE Turfitt. The microbiological degradation of steroids: 4. fission of the steroid molecule. *Biochemical Journal*, 42(3):376, 1948.

[110] W. A. van der Donk and H. Zhao. Recent developments in pyridine nucleotide regeneration. *Current opinion in biotechnology*, 14(4):421–426, 2003.

[111] A. Ventosa, M. C. Gutiérrez, M. Kamekura, and M. L. Dyall-Smith. Proposal to transfer *Halococcus turkmenicus, Halobacterium trapanicum* jcm 9743 and strain gsl-11 to *Haloterrigena turkmenica* gen. nov., comb. nov. *International Journal of Systematic Bacteriology*, 49(1):131–136, 1999.

[112] J. M. Vrtis, A. K. White, W. W. Metcalf, and W. A. van der Donk. Phosphite dehydrogenase: an unusual phosphoryl transfer reaction. *Journal-American Chemical Society*, 123(11):2672–2673, 2001.

[113] S. Wang, M. M. Kayser, H. Iwaki, and P. C. K. Lau. Monooxygenase-catalyzed baeyer-villiger oxidations: Chmo versus cpmo. *Journal of Molecular Catalysis B: Enzymatic*, 22(3):211–218, 2003.

[114] A. Willetts. Structural studies and synthetic applications of baeyer-villiger monooxygenases. *Trends in biotechnology*, 15(2):55–62, 1997.

[115] C. Wong, D. G. Drueckhammer, and H. M. Sweers. Enzymatic vs. fermentative synthesis: thermostable glucose dehydrogenase catalyzed regeneration of nad (p) h for use in enzymatic synthesis. *Journal of the American Chemical Society*, 107(13):4028–4031, 1985.

[116] J. M. Woodley. New opportunities for biocatalysis: making pharmaceutical processes greener. *Trends in biotechnology*, 26(6):321, 2008.