Identification of a Baeyer-Villiger monooxygenase sequence motif
Fraaije, Marco; Kamerbeek, NM; van Berkel, WJH; Janssen, Dick; Kamerbeek, Nanne M.; Berkel, Willem J.H. van
Published in:
FEBS Letters

DOI:
10.1016/S0014-5793(02)02623-6

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2002

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Identification of a Baeyer–Villiger monooxygenase sequence motif

Marco W. Fraaije, Nanne M. Kamerbeek, Willem J.H. van Berkel, Dick B. Janssen

1. Introduction

Baeyer–Villiger monooxygenases (BVMOs) are flavoenzymes that are able to catalyze a Baeyer–Villiger reaction; using dioxygen and NAD(P)H a ketone function is converted into the corresponding ester (Scheme 1) [1]. Cyclohexanone monooxygenase from Acinetobacter NCIMB 9871 is the best characterized BVO [2,3]. Because of its wide substrate spectrum and stereoselectivity this enzyme has been extensively used for synthetic applications [1,4]. Like other flavoprotein monooxygenases, cyclohexanone monooxygenase forms a hydroperoxylavlin intermediate which is involved in substrate oxygenation [5,6].

Until a few years ago, cyclohexanone monooxygenase was the only BVO of which the sequence was known. Recently, several other bacterial BVMOs were cloned and sequenced: 4-hydroxyacetophenone monooxygenase [7], steroid monooxygenase [8], cyclododecanone monooxygenase [9] and two other cyclohexanone monooxygenases [10]. Using these data we performed a sequence alignment study which uncovered a conserved sequence motif. The function of this BVO-specific sequence motif was explored by site-directed mutagenesis of 4-hydroxyacetophenone monooxygenase. Moreover, our search for the presence of conserved sequence motifs also resulted in the disclosure of a superfamily of sequence-related FAD-dependent monooxygenases.

2. Materials and methods

2.1. Materials

Chemicals were purchased from Acros Chimica, Merck, Aldrich, or Sigma. Molecular biology enzymes were purchased from Roche. Oligonucleotides were supplied by Eurosequence BV, Groningen, The Netherlands.

2.2. Site-directed mutagenesis

The Quickchange site-directed mutagenesis kit from Stratagene was used to introduce mutations into the 4-hydroxyacetophenone monooxygenase gene. Successful mutagenesis was confirmed by plasmid sequencing. Mutant proteins were expressed and purified as described for wild type enzyme [7].

2.3. Analytical methods

4-Hydroxyacetophenone monooxygenase activity was determined as described previously [7]. The relative molecular mass of native 4-hydroxyacetophenone monooxygenase was determined by FPLC gel filtration using a Superdex 200 HR 10/30 column (Pharmacia Biotech) running with 50 mM potassium phosphate buffer, pH 7.5. Absorption spectra were recorded at 25°C on a Perkin Elmer Lambda Bio 40 spectrophotometer. Circular dichroism (CD) spectra were recorded at 25°C from 190 to 250 nm using a 0.1-cm cuvette containing 0.1 mg/ml of enzyme (5 mM potassium phosphate, pH 7.5), and an AVIV 62A DS spectrometer.

2.4. Sequence homology analysis

The BLAST programs at the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov/BLAST) were used to search for proteins showing sequence similarity. Multiple sequence alignments were made with the ClustalW program at the Centre for Molecular and Biomolecular Informatics (www.cmbi.kun.nl/bioinf/tools). Sequence alignments were visualized using the BOXSHADE 3.21 program (http://www.ch.embnet.org/software/BOX_form.html). TreeView 1.5.2 was used for generating a tree representation using the ClustalW output.

3. Results and discussion

3.1. The BVO family

A non-redundant database search at the NCBI using the protein sequence of 4-hydroxyacetophenone monooxygenase and the BLAST program [11] yielded 35 protein sequences with significant similarity (S > 100, E < 1 x 10^-20). Besides the above-mentioned bacterial BVMOs (sequence identities of 27–33%), this set of homologs contains two fungal gene products that have been shown to be involved in biosynthetic pathways.
of mycotoxins which include Baeyer-Villiger oxidation steps [12,13]. All BVMO homologs found are of bacterial or fungal origin, indicating that BVMOs play a role in a variety of microbial oxidative metabolic pathways. Interestingly, a relatively large number of BVMO homologs was found in pathogenic bacteria (e.g. seven in *Mycobacterium tuberculosis* H37Rv). In addition, by performing a BLAST search in the NCBI database of patented nucleotide sequences we found that a DNA probe (accession nr. L04542) used for specific detection of pathogenic *Mycobacterium avium* isolates [14] encodes a major part of a putative BVMO (35% sequence identity). This indicates that these oxidative enzymes represent attractive targets for drug development. In fact, one of the seven BVMO-related genes from *M. tuberculosis*, was recently shown to be responsible for the activation and therefore the efficacy of the widely used anti-tuberculosis pro-drug ethionamide [15,16]. It has been shown that the *etaA* gene product mediates ethionamide activation by sulfoxidation; a reaction typically catalyzed by BVMOs.

Alignment of the BVMO homologs revealed several conserved regions containing sequence motifs that are known to be involved in dinucleotide cofactor binding [17]. The conserved sequence motifs can also be clearly recognized from the alignment of all above-mentioned biochemically identified BVMOs, as shown in Fig. 1. The presence of two Rossmann folds, as evidenced by two GXXGXXGXX(Y/F) motifs, clearly discriminates these enzymes from the mechanistically related flavoprotein hydroxylases [18]. Furthermore, we noticed a stretch of conserved residues just before the second Rossmann fold motif. With two exceptions, this FXGXXXXXXGXX(P/D) motif is strictly conserved in all BVMO homologs and represents, apart from the Rossmann fold motif regions, the sequence region with the highest number of conserved residues. In the two aberrant protein sequences from *Caulobacter crescentus* the central histidine is not conserved.

To probe the specificity of the BVMO sequence motif we also used the motif as a seed for a Pattern Hit Initiated BLAST search [19]. By this approach, the above-mentioned sequences could be specifically retrieved. Furthermore, a trawl of the PEDANT genomic sequence database (http://pedant.gsfc.de/) using only the BVMO sequence motif as seed for a pattern search resulted in the specific identification of 32 microbial BVMO homologs. These pattern-based searches exemplify the potential of the described sequence motif for efficient annotation of putative BVMO genes.

### 3.2. BVMOs are sequence-related to flavin-containing monooxygenases (FMOs) and N-hydroxylating monooxygenases (NMOs)

The sequential organization of the dinucleotide binding sequence motifs observed in known BVMO sequences is reminiscent of two other flavoprotein monooxygenase families: multi-functional FMOs (eukaryotic enzymes oxygenating a broad range of amines and other heteroatom containing compounds) [20] and NMOs (microbial enzymes catalyzing N-hydroxylations of long-chain primary amines) [21]. A search in the non-redundant database of NCBI using the PSI-BLAST program [22] confirmed sequence homology between BVMOs, FMOs and NMOs, revealing the presence of more than 150 sequences belonging to one of these three flavoprotein families.

Two representative members of the FMO and NMO families have been included in the multiple sequence alignment depicted in Fig. 1. The observed sequence similarity between members of the three different flavoprotein families is reflected in their structural and catalytic properties: all characterized members are single-component FAD-containing monooxygenases of about 50 kDa which specifically use NADPH as electron donor. A dendrogram of all homologous sequences gives insight in the sequence distance relationship within the discovered flavoprotein superfamily nicely illustrating the clustering of sequences in the three distinct monooxygenase families (Fig. 2).

From the three monooxygenase families, members of the FMO family have been most extensively studied as they play an essential role in the detoxification of xenobiotics (e.g. nicotine, drugs) in vertebrates [20]. Vertebrates typically contain five FMO isoforms. Several naturally occurring mutations in human FMO3 have been shown to underlie inheritable trimethylaminuria (also called ‘fish-odor syndrome’) illustrating the important role of FMOs in metabolic detoxification [23]. Except for the known mammalian FMOs, our sequence similarity search revealed a surprising number of FMO homologs in plants (26 in the *Arabidopsis thaliana* genome). Recently, the function of three of these homologs (YUCCA 1–3) has been identified, confirming a FMO-type reactivity as these enzymes catalyze an N-oxygenation of tryptamine yielding a precursor of the plant hormone auxin [24]. Interestingly, in all FMO sequences a similar sequence motif as found for BVMOs can be identified, FXGXXXXXXGXX(Y/F), differing by only one residue (Fig. 1). Only in the FMO homologs from yeast this sequence signature is not fully conserved. In this respect it is interesting to note that yeast FMO has lost its ability to catalyze typical FMO reactions [25].
Members of the NMO family have been less well studied which is partly due to their low affinity for the FAD cofactor [26]. In NMO sequences, only the histidine appears to be strictly conserved in the sequence region corresponding to the location of the above-mentioned BVMO and FMO sequence motifs (Fig. 1) (see [27]). The conservation of this histidine residue throughout the monooxygenase superfamily indicates that it fulfills a role in a shared function within the monooxygenase superfamily.

3.3. Probing the functional role of the newly identified sequence motif

To examine the functional role of the BVMO-specific sequence motif we performed a site-directed mutagenesis study on 4-hydroxyacetophenone monooxygenase from *Pseudomonas fluorescens*. This BVMO has recently been cloned and characterized in our laboratory providing a good opportunity to investigate the observed sequence conservation [7]. As targets for mutagenesis, the conserved His296 and Trp300 residues were chosen. Expression of both W300A and W300Y mutant enzymes resulted in formation of insoluble and inactive protein. This indicates that the conserved tryptophan serves a critical role in attaining and/or maintaining a correctly folded state of the enzyme. The H296A mutant enzyme was expressed in soluble form and could be purified using the procedure used for wild type 4-hydroxyacetophenone monoxygenase. However, during purification the mutant enzyme lost its FAD cofactor to some extent. Holo H296A enzyme could be obtained by saturating the apoprotein fraction with FAD or by purifying the protein in the presence of 50 μM FAD. Both procedures resulted in fully reconstituted H296A as indicated by an A280/A440 ratio of 16.8 which is similar to that observed for wild type enzyme. Gel filtration experiments revealed that, similar to wild type enzyme, the holo form of the H296A mutant is present as a dimer. Furthermore, the far-UV CD spectra of wild type 4-hydroxyacetophenone monoxygenase and the H296A mutant were identical, indicating that the mutation did not affect overall protein folding. Kinetic analysis revealed that the H296A mutant showed nearly no activity (kcat < 0.01 s⁻¹). Moreover, the flavin spectral properties of the H296A mutant were slightly different from those of wild type 4-hydroxyacetophenone monoxygenase.
References