Enzyme-substrate matching in biocatalysis: in silico studies to predict substrate preference of ten putative ene-reductases from *Mucor circinelloides* MUT44.

Gianluca Catucci¹, Alice Romagnolo¹, Federica Spina¹, Giovanna Cristina Varese¹, Gianfranco Gilardi¹,², Giovanna Di Nardo¹,²,*

¹Department of Life Sciences and Systems Biology, University of Torino, via Accademia Albertina 13, 10123 Torino, Italy

²CrisDi, Interdepartmental Center for Crystallography, via Pietro Giuria 7, 10125, Torino, Italy

*Corresponding author

E-mail: giovanna.dinardo@unito.it
Abstract

Ene-reductases are flavoproteins able to catalyse the reduction of carbon-carbon double bonds with many potential applications in biocatalysis. The fungus *Mucor circinelloides* MUT44 has high ene-reductase activity when grown in the presence of substrates carrying different electron-withdrawing groups. Genome sequencing revealed the presence of ten putative genes coding for ene-reductases that can be potentially exploited for biocatalytic purposes. To this end, the availability of a method able to predict which isoform binds and turns over a specific substrate would help to choose the best catalyst for the desired bioconversion.

Here, homology models of the ten putative enzymes are first generated, validated and show that the proteins share the typical TIM barrel fold with a conserved β-hairpin cap on one side of the barrel and a non-conserved subdomain capping the other side, where the FMN cofactor is accommodated. The active site of the ten enzymes is different in terms of both volume and charge distribution whereas the residues responsible for substrate recognition and catalysis are generally conserved.

Docking of cyclohexenone into the active site of the ten enzymes shows a binding almost superimposable to that found in pentaerythritol tetranitrate reductase in complex with this substrate (PDB ID 1GVQ) in isoforms 1, 2, 6 and 10.

The data demonstrate that *in silico* predictions can be used for new putative fungal ene-reductases to predict the best substrate-enzyme matching for the selection of the most suitable catalyst for the desired biotransformation.

Keywords: ene-reductase, Old Yellow enzyme, substrate selectivity, carbon–carbon double bond, biocatalysis
1. Introduction

Ene-reductases (ERs) are members of the Old Yellow Enzymes (OYEs) family and are NAD(P)H-dependent flavin-containing enzymes able to catalyse the reduction of carbon-carbon double bonds (C=C) on a wide range of α,β-unsaturated substrates [1]. OYEs were first isolated from the yeast Saccharomyces pastorianus [2] and then from other sources such as bacteria, plants and filamentous fungi [3] where they can participate to the metabolism of both endogenous and xenobiotic compounds [1]. Some OYEs are involved in the biosynthesis of fatty acids [4] or, in plants, 12-oxophytodienoate reductase (OPR) is involved in the biosynthesis of jasmonic acid, a compound that regulates gene expression in plant development and defense [5]. Many other OYEs are “orphans” since the physiological substrates and their role in metabolism is still unknown. However, some OYEs have attracted a lot of attention due to their ability to perform a biotechnologically important reaction, that is the stereoselective reduction of activated C=C, on a wide range of substrates of different sizes [6-8]. The resulting chiral compounds are industrially relevant and therefore OYEs are very attractive as biocatalysts [9].

Among the organisms where these enzymes have been described, fungi have been shown to possess a different number of OYEs homologs in their genomes [10]. Most of the species analysed have from 3 to 7 genes coding for these proteins. Although some of them may be pseudogenes or expressed under control of different promoters, these data suggest a possible coexistence of different isoenzymes in fungal cells.

Since fungi are highly versatile organisms, able to grow in different environmental conditions and using different substrates, it is also possible that there is a differential expression of the isoenzymes depending on the external stimuli such as nutrients, presence of noxious substrates, physico-chemical parameters of the
surrounding environment. It can be expected that the different isoenzymes have evolved toward the recognition and reduction of specific substrates, making them even more attractive in the field of biocatalysis because of the possibility to have an enzyme portfolio catalytically optimized toward a wide range of substrates.

Recently, the fungus *Mucor circinelloides* MUT44 has been shown to be the most efficient compared to other selected fungal strains, in reducing three model substrates, cyclohexenone, \(\alpha\)-methylcinnamaldehyde and \((E)\)-\(\alpha\)-methylnitrostyrene, all characterized by the presence of different electron-withdrawing groups (EWG) and different steric hindrance [11]. Ten putative sequences of OYE genes (McOYE1-10) were found in its genome by means of a BlastP analysis, in which the query was OYE1 sequence from *S. pastorianus*. A recent fungal OYE classification by Nizam and collaborators [10,12] clearly showed that this class of enzymes is divided in 3 distinct groups according to the structural peculiarity (e.g. core of the active site, accessory residues, loop regions): class I, class II and class III. By applying the same analysis parameters, nine out of ten OYEs from *M. circinelloides* MUT44 clustered together in class I, where OYE from bacteria, yeasts, filamentous fungi, animals and plants can be found, showing a species-specific clade, whereas only McOYE10 resulted located in class II [13]. Moreover, the expression profile of the ten enzymes is different when the fungus is grown in the presence of different substrates, suggesting that the different isoenzymes could be specialized for the conversion of different molecules [13]. In such a case, the conversion potential of the fungus can be exploited for biocatalytic purposes by using only selected isoenzymes specialized for the compound of interest.

In this work, sequence analysis and structure prediction through homology modeling are used to gain information about the degree of similarity among the ten putative isoenzymes and possible structural differences also compared to homologs...
from yeasts and plants. Moreover, we performed docking simulations with cyclohexenone and compare the results to the crystal structure of pentaerythritol tetranydrate reductase in complex with the same molecule to find out which isoforms are able to bind this substrate effectively.

2. Experimental

2.1 Sequence analysis and alignments.
The putative OYE homologues of *M. circinelloides* and sequence IDs according to JGI database (http://genome.jgi.doe.gov/programs/fungi/index.jsf) are reported in Table 1. Multiple sequence alignments were performed using Clustal Omega software [14]. Primary structure analysis was performed using PROTPARAM [15] and TOPPRED [16]. Secondary structure prediction was performed using PSIPRED server [17].

2.2 Homology modeling of *Mucor circinelloides* MUT44 (McOYEs)
The software Modeller v9.11 [18,19] was used for homology modeling. The search of homologs of known crystal structure was carried out using PSI-Blast. Multiple templates were chosen on the basis of sequence identity, full coverage and high resolution of their crystal structure. The templates chosen for McOYE1-9 models are the crystal structures of plant 12-oxophytodienoate reductase 1 from tomato (PDB ID 1ICS, 3HGR) [20,21], the Old Yellow Enzyme from *Saccharomyces pastorianus* (PDB ID 1OYA) [22] and from *Kluyveromyces marxianus* AKU4588 (PDB ID 4TMB) [23]. For McOYE10, the templates used were the thermostable OYE from *Thermoanaerobacter pseudethanolicus* E39 (PDB ID 3KRU) [24], *Thermus scotoductus* SA-01 (PDB ID 3HF3) [25], YqjM from *Bacillus subtilis* (PDB ID 1Z41) [26], OYE from *Geobacillus kaustophilus* (PDB ID 3GR7) [27] and the xenobiotic
reductase A from *Pseudomonas putida* 86 (PDB ID 2H8X) [28]. Thus, for McOYE10, the best structural homologs were all bacterial enzymes.

For each McOYE, ten models were generated and the best model was chosen on the basis of the normalized Discrete Optimized Protein Energy (DOPE) parameter, an atomic distance-dependent statistical score [29]. In order to optimize the side chain packing and interactions, all models were subjected to energy minimization through AMBER 03 force field. The models were analyzed from a structural point of view and their quality was checked by PROCHECK [30] ProSA [31] and the QMEAN scoring function [32].

### 2.2 Analysis of active site, surface charge and FMN binding site

The active site volume of each model was estimated through the CASTp server [33]. The surface charge of the models was also analyzed through the Coulombic surface coloring tool available on the UCSF Chimera software [34], in order to identify densely charged regions on the enzyme surface that could suggest interactions with other partners such as proteins bearing surface patches with opposite charge.

FMN binding sites were analyzed with the program LigPlot' [35].

### 2.3 Ligand docking

The chemical structure of cyclohexenone was subjected to molecular geometry optimization with YASARA [36]. The substrates were docked into the models of the putative McOYEs with the program AutoDock v4 [37,38] available in the YASARA package using 100 runs of flexible docking. A simulation cell (18 Å x 18 Å x18 Å) was built around the FMN group of McOYEs. Since the crystal structure of an OYE in complex with cyclohexenone is available (PDB ID 1GVQ) [39], the substrate was removed from this PDB entry and a first docking simulation was performed on this
protein with the same substrate to validate our approach and to calculate the binding energy. The same simulation cell was then used for the McOYEs models. The binding energies and dissociation constant (K_D) values were predicted using the scoring function included in the YASARA embedded AutoDock package. A set of structures having a root mean square deviation (RMSD) of less than 1Å was included into single clusters and the clusters were ranked according to their binding energies. The best cluster in terms of binding energy for each simulation was used for predicting the final pose of the protein/ligand interaction.

3. Results and discussion

3.1 Sequence alignments and analysis of the primary structure of McOYEs

Primary sequences of the ten McOYEs considered in this work are from 364 to 396 amino acids long, and with an identity percentage ranging from 24.6% for isoforms 9 and 10 to 91.8% for isoforms 1 and 2 (Table S1).

Sequence alignments of McOYEs with representative members of different classes of OYEs of known crystal structure (Figures S1 and S2) show that the consensus sequence for substrate binding, that is HX_1X_2HGY or HX_1X_2NGY, where X_1 and X_2 are often small residues such as glycine, alanine or serine, is almost fully conserved with some significant exceptions. The two His or one His and one Asn of this consensus sequence are directly involved in H-bonds with the substrates [40], and the Tyr residue is catalytically important since it donates a proton to the substrate during catalysis [41]. McOYE1-8 presents the HX_1X_2NGY motif whereas, in McOYE9, the motif is not conserved, as a phenylalanine replaces the second His or Asn residue known to directly form a hydrogen bond with substrates. This substitution is highly significant since it can
affect the substrate binding ability of the enzyme. In McOYE10, the motif is HX$_1$X$_2$HGY, showing a second histidine acting as hydrogen bond donor.

The YGGS fingerprint motif, typical of OYEs and, more in general, of TIM-barrel proteins, is also present in all the fungal enzymes (Figures S1 and S2). This motif is important for the movement of loops involved in substrate binding [42-47].

McOYE proteins have molecular weights ranging from 40,166 to 43,058 kDa and the analysis with Wolf PSORT server [48] predicts mitochondrial localization, except for McOYE5 and 10 that are predicted to be in the cytoplasm (Table 1). A significant difference among the putative ten isoforms is found in the theoretical pI of McOYE8 that shows the highest theoretical pI (7.18) compared to the other isoforms, where a more acidic pI ranging from 5.45 to 6.13 is predicted. This difference is justified by an increase in the number of basic residues in the sequence of McOYE8.

3.2 Homology modeling, refinement and validation of McOYEs
Homology models were built using multiple templates that were chosen on the basis of high sequence identity, that ranged from 40 to 50% for all the McOYEs aligned with the corresponding templates. Moreover, the templates were chosen also on the basis of their high coverage of the McOYE sequences, meaning that possible insertions in McOYEs were considered and aligned where possible. It is interesting to notice that the highest identity for McOYE8 1-9 was found with plant enzymes compared to yeast proteins.

Out of the 10 models built using MODELLER for each sequence, the best one according to the zDOPE score was selected. The models were then energy minimized using AMBER 3 force field and their quality assessed by PROCHECK, Q-MEAN and ProSA. Furthermore, since the models are then used for docking purposes, their quality in terms of packing of the residues side chains and atom contacts was also checked by Verify3D server [49]. The main results of model validation are summarized in Table S2.
The Ramachandran plots show that at least 96% of the residues fall in the most favorable and allowed regions in all cases. The G-factor overall score is within the range of acceptable values in PROCHECK in all cases. The z-scores calculated through ProSA and Q-mean is in all cases within the range of scores typically found for native proteins of similar size [31].

The server Verify3D was also used to check the atomic contacts and the compatibility of the model (3D) with the amino acid sequence (1D). For the quality of the predicted model to be considered satisfactory, it is expected to have the Verify3D score more than 80%. For all McOYE models, this value ranges from 89 for McOYE9 to 98% for McOYE4.

### 3.3 Analysis of the overall structure of McOYE models

The models of the ten McOYEs are shown in Figure 1. The scaffold of the proteins consists of a well-conserved TIM barrel domain with eight twisted β-strands surrounded by eight α-helices. The N-terminus is also a well-conserved β-hairpin that closes one side the barrel. On the other side of the barrel, the FMN cofactor is accommodated.

According to the known crystal structures of different OYEs, the most structurally variable and flexible region that differentiates these enzymes is the so-called “capping subdomain” that is an extended loop forming 2 two-stranded beta-sheets in tomato 12-oxophytodienoate reductase (OPR) whereas, in the OYE from yeasts, this region forms a two-stranded beta-sheet and 2 α-helices (Figure 2A). As a result, in the plant enzymes the loop extends to the FMN binding region and delineates part of the active site whereas in the yeast enzyme this region is not part of the catalytic pocket (Figure 2A). In any case, in this region thought to be highly flexible [21] conformational changes have been observed upon ligand binding [50]. The McOYE models show that this region is highly disordered exception made for McOYE8 and 9,
where only one short α-helix is present. This is in line with the poor alignment found in this region between McOYE1-9 and the templates used for homology modeling. According to the alignment, there is a 3 amino acids deletion in McOYE1-9 compared to the tomato enzyme that shorten the loop that does not extend and does not form the 2 stranded β-sheet as in the plant enzyme. On the other hand, the alignment of McOYE5-9 with the yeast enzymes is also poor in this region and the propensity to form helical structure is lower due to the substitution of good helix-forming amino acids such as glutamate by proline and glycine residues. The low propensity to form α-helices was also confirmed by PSI-Pred server that did not predict any secondary structure in this region for McOYE1-7, whereas a short helix is predicted for McOYE8 and 9. Thus, in general, the McOYE5s seem to have an even more flexible capping subdomain when compared to the yeast and plant counterparts. Moreover, in McOYE3, 4 and 9, the two stranded β-sheet connecting the TIM barrel and the capping subdomain is also missing.

McOYE10 model shows a significant difference compared to the other McOYE5s and the template structures, due to an insertion in the region 246-255 that forms a long loop connecting strand 5 and helix 5 (Figure 2B). Also in this case, the capping subdomain is long and disordered at the top of the barrel and, according to the alignment, this region is particularly long in the case of McOYE10 compared to the other OYE5s used as templates, exception made for the FMN-containing xenobiotic reductase A (XenA) form Pseudomonas putida 86. Furthermore the crystal structure of XenA confirms the presence of a highly disordered capping subdomain, that in McOYE10 contains 6 glycines and 1 proline residues out of the 20 amino acids forming this stretch.

The crystal structure of YqiM from Bacillus subtilis revealed a dimeric organization of the protein where the so-called “arginine finger” of one monomer
protrudes toward the active site of the other and it is directly involved in substrate recognition. This arginine residue is conserved in McOYE10, suggesting the possibility to form dimers that cooperates in catalysis also in this fungal protein (Figure 2B).

3.4 Analysis of the FMN binding site

The FMN cofactor in OYEs is non-covalently bound to the protein through highly conserved residues present at the top of the barrel. In OYEs, FMN is roughly perpendicular to the barrel axis [22,51] and the re-face of FMN is completely buried, while the si-face is in contact with the solvent [21,52]. The residues contacting FMN through their side chains are highly conserved among the OYE family, while those forming main chain contacts are variable [21].

In McOYE models, the FMN cofactor is accommodated as in the other members of the OYE family. Sequence alignments and a Ligplot analysis of the homology models (Figure S3) show that the residues involved in H-bonds, charge-charge and Van der Waals interactions are highly conserved with few exceptions. In some cases, due to the involvement of backbone groups in H-bond with FMN, these substitutions do not alter the type of interaction with FMN. It should be noted that Asn325, involved in FMN binding in the plant OYE and absent in the yeast protein, is conserved in all the McOYEs, with the exception of McOYE9, where a serine is present.

3.5 Surface potential and catalytic pocket dimensions

Since a difference in the theoretical pI of McOYE8 was found, the Coulombic surface of the ten models was generated using the UCSF Chimera tool and the results shown in Figure 3, where the yeast and tomato OYEs are also included.
The surface charge at the access of the catalytic pocket shows some differences in the plant enzyme (Figure 3A), in the yeast enzymes (Figure 3B) and McOYE s (Figure 3C-3L). In general, McOYE s show a more neutral charge surface exception made for McOYE6 and McOYE8, where the active site face is mainly negatively and positively charged, respectively. This is in line with the more basic theoretical pI previously observed in McOYE8. However, according to sequence alignments, the extra basic residues found in McOYE8 compared to the other fungal proteins are homogeneously distributed on the protein surface suggesting that they may also confer a generally different stability at basic pH rather than being involved in a positively charged cluster that can be important for substrate recognition as well as for protein-protein interaction.

Furthermore, it can be noticed from Figure 3 that McOYE s show different active site mouths dimensions as well as different volumes (Table 1). In fact, the volume of the catalytic pocket ranges from 112.5 for McOYE9 to 1,587.4 Å³ for McOYE10. These differences are significant even if it should be taken into account that the highly flexible region found in McOYE s can affect the active site dimensions. However, what is interesting to note is the large and accessible active site of McOYE10, where the FMN cofactor is highly exposed compared to the other fungal proteins (Figure 3). The FMN cofactor is poorly accessible in McOYE5, as in the case of the yeast enzyme (PDB ID 1OYA). These data suggest that McOYE putative proteins can attract and accommodate substrates of different sizes and carrying different charges/dipoles possessing a different substrates preference.

3.6 Active site

The crystal structures of OYEs in complex with different substrates revealed a well-conserved binding mode. The binding to the protein occurs through hydrogen bonds involving an oxygen atom of the substrate and 2 His or 1His/1Asn residues. According
to sequence alignments and the models generated, the residues involved are highly conserved in McOYE\(s\) and they are 1 His and 1 Asn/His. As mentioned above, there is an exception represented by McOYE\(9\), where a phenylalanine residue is present. For McOYE\(9\), the presence of a hydrophobic residue not able to form hydrogen bonds with substrates, can affect its binding ability or affinity.

Other relevant residues, known to be important for substrate binding in yeast OYE are Phe250, Phe296 and Tyr375. The first phenylalanine residue, corresponding to a tyrosine in the OPR plant enzymes, is highly conserved in McOYE\(s\), exception made for McER\(8\) and 9, where a valine and an alanine are present. Phe296 is not conserved in McOYE\(s\) whereas Tyr375 is highly conserved in all the fungal isoenzymes.

Another important residue (Thr37) known to increase the redox potential of FMN in OYE [53], forms a hydrogen bond with N(5) and O(4) of the pyrimidine moiety, is also highly conserved.

For bacterial OYE\(s\), other residues important for substrate binding (Cys26, Tyr28, Arg336 in YqjM) are well conserved in McOYE\(10\).

### 3.7 Substrate docking

Since ene-reductase activity has been detected in *Mucor circinelloides* MUT44 on cyclohexenone [11], and the crystal structure of a related enzyme (pentaerythritol tetranitrate reductase) was obtained in complex with this substrate (PDB ID 1GVQ) [38], we used this molecule to perform molecular docking on McOYE\(s\) to further validate our models as well as the docking results. According to the crystal structure of the tetranitrate reductase, the binding mode of this substrate to the enzyme is the “classical” one through two hydrogen bonds formed with two well-conserved histidines. The substrate is located between the FMN ring and a conserved tyrosine residue, important for catalysis. In order to validate our docking method, the substrate was first
removed from the PDB entry of the crystal structure of pentaerythritol tetranitrate reductase and docked back into the structure. The RMSD obtained between the substrate present in the crystal structure and the one docked is 0.126 Å demonstrating an almost complete superimposition.

Then, the docking was run on McOYEs homology models (Figure S4). As summarized in Table 2, the binding energies calculated for McOYEs are comparable to the one predicted for the docking of cyclohexenone in the known crystal structure, exception made for McOYE10, where the energy is significantly lower. However, the substrate cyclohexenone is almost superimposable to the substrate present in the known crystal structure in isoforms 1, 2, 6 and 10 where, according to the distances measured, it can form two hydrogen bonds with the HXXH/N cluster (Figure S4 and Table 2). In McOYE3, the substrate cannot reach the active site of the protein and form H-bonds with Arg324. In McOYE9, the asparagine residue forming the H-bond with the keto moiety of the substrate is missing and substituted by a phenylalanine residue that has no H-bond acceptors.

In McOYE4 and McOYE7, the residues Ile271 and Met273, respectively, pull the substrate far away from the asparagine residue usually involved in a H-bond. These data show that McOYE1, 2, 6 and 10 are potentially the most suitable biocatalysts to convert cyclohexenone into cyclohexanone.

4. Conclusions
The fungus *Mucor circinelloides* MUT44, previously shown to have ene-reductase activity [11], possesses ten genes coding for putative ene-reductases belonging to the Old Yellow Enzymes family. Since the reduction of C=C double bonds is one of the most important strategies for the production of compounds with up to two chiral centers [1], these enzymes are very important for biocatalytic purposes. To this end, *in silico* studies can
be used to gain information about the structural and functional properties of these new putative enzymes. For example, analysis of the primary sequence of the McOYEs, highlight that one isoform (McOYE10) is homolog to thermophilic-like bacterial OYEs whereas McOYE8 has a theoretical pI of 7.18 due to the presence of extra Arg/Lys residues distributed on the protein surface. Thus, these two isoforms are likely to have a higher stability at high temperatures (McOYE10) and at more basic pH (McOYE8) compared to the other McOYEs.

From a structural point of view, the McOYEs share the TIM barrel scaffold typical of the OYEs family where structural variability is associated with the capping subdomain that is a flexible region involved in active site shaping and also in NAD(P)H and substrate binding. Interestingly, the crystal structure of an OYE from *Shewanella oneidensis* in complex with p-hydroxybenzaldehyde shows that a second molecule of the substrate is present in a hydrophobic cleft next to the entry of the active site tunnel in the capping subdomain [50], indicating an important role of this region in substrate binding. According to our models, the volume can vary from 126.8 to 1587.4 Å³ indicating that the fungal enzymes can accommodate substrates of different size and therefore possess different substrate selectivity. This hypothesis found a confirmation in the high versatility shown by *M. circinelloides* MUT44 in the reduction of compounds with different EWG and different steric hindrance such as ketone, aldehyde, nitro and carboxylic group [11].

In conclusion, our data show that it is possible to develop and validate *in silico* methods to predict structural differences among isoenzymes from the same organism in order to select the most suitable catalyst for the desired application.

**Acknowledgments**

This work was partially funded by Compagnia San Paolo and Fondazione CRT (Italy).
Figure 2. Comparison of the structurally different elements of OYEs. A) Superimposition of the crystal structures of yeast OYE (PDB ID 1OYA) and tomato 12-oxophytodienoate reductase 1 (OPR1) from tomato (PDB ID 1ICS). The most conserved regions are shown in light grey and the variable subcapping domain region is colored. The α-helices of the yeast enzyme are shown in violet, whereas the 2-stranded β-sheet of the plant enzyme is shown in magenta. The FMN cofactor is shown in yellow. B) Superimposition of the crystal structure of YqjM from Bacillus subtilis (dark grey) and the homology model of McOYE10 (light grey). The long loop connecting strand 5 and helix 5 in McOYE10 model is shown in red. In the bacterial protein, the so-called “arginine finger” (orange) of one monomer protrudes toward the active site of the other and it is directly involved in substrate recognition. This residue is conserved in McOYE10 and shown in green. The FMN cofactor is shown in yellow.
Figure 3. Analysis of the surface charge distribution in OYEs. Coulombic surface of A) yeast OYE (PDB ID 1OYA), B) tomato 12-oxophytodienoate reductase 1 (OPR1) from tomato (PDB ID 1ICS), C) McOYE1, D) McOYE2, E) McOYE3, F) McOYE4, G) McOYE5, H) McOYE6, I) McOYE7, J) McOYE8, K) McOYE9 and L) McOYE10.

Supporting information captions
References


into substrate binding and specificity within the family of OYE, Structure 9 (2001) 419-429.


Table S1. Sequence identities derived from the alignment of McOYEs.

<table>
<thead>
<tr>
<th></th>
<th>McOYE1</th>
<th>McOYE2</th>
<th>McOYE3</th>
<th>McOYE4</th>
<th>McOYE5</th>
<th>McOYE6</th>
<th>McOYE7</th>
<th>McOYE8</th>
<th>McOYE9</th>
<th>McOYE10</th>
</tr>
</thead>
<tbody>
<tr>
<td>McOYE1</td>
<td>100</td>
<td>91.8</td>
<td>73.9</td>
<td>74.6</td>
<td>64.5</td>
<td>65.4</td>
<td>62.8</td>
<td>66.5</td>
<td>56.2</td>
<td>27.8</td>
</tr>
<tr>
<td>McOYE2</td>
<td>91.8</td>
<td>100</td>
<td>73.0</td>
<td>75.1</td>
<td>65.3</td>
<td>65.3</td>
<td>63.0</td>
<td>66.5</td>
<td>56.6</td>
<td>26.7</td>
</tr>
<tr>
<td>McOYE3</td>
<td>73.9</td>
<td>73.0</td>
<td>100</td>
<td>75.5</td>
<td>60.7</td>
<td>63.2</td>
<td>61.4</td>
<td>61.8</td>
<td>556.0</td>
<td>31.5</td>
</tr>
<tr>
<td>McOYE4</td>
<td>74.6</td>
<td>74.6</td>
<td>74.6</td>
<td>100</td>
<td>63.4</td>
<td>64.8</td>
<td>61.1</td>
<td>63.2</td>
<td>56.0</td>
<td>28.8</td>
</tr>
<tr>
<td>McOYE5</td>
<td>64.5</td>
<td>64.5</td>
<td>64.5</td>
<td>63.4</td>
<td>100</td>
<td>61.7</td>
<td>61.9</td>
<td>62.9</td>
<td>55.9</td>
<td>29.2</td>
</tr>
<tr>
<td>McOYE6</td>
<td>65.4</td>
<td>65.4</td>
<td>65.4</td>
<td>64.8</td>
<td>61.7</td>
<td>100</td>
<td>66.4</td>
<td>65.2</td>
<td>55.9</td>
<td>29.2</td>
</tr>
<tr>
<td>McOYE7</td>
<td>62.8</td>
<td>62.8</td>
<td>62.8</td>
<td>61.1</td>
<td>61.9</td>
<td>66.4</td>
<td>100</td>
<td>60.1</td>
<td>55.9</td>
<td>29.2</td>
</tr>
<tr>
<td>McOYE8</td>
<td>66.5</td>
<td>66.5</td>
<td>66.5</td>
<td>63.2</td>
<td>62.9</td>
<td>65.2</td>
<td>60.1</td>
<td>100</td>
<td>55.9</td>
<td>29.2</td>
</tr>
<tr>
<td>McOYE9</td>
<td>56.2</td>
<td>56.6</td>
<td>56.0</td>
<td>56.0</td>
<td>55.9</td>
<td>55.3</td>
<td>52.8</td>
<td>56.1</td>
<td>100</td>
<td>24.6</td>
</tr>
<tr>
<td>McOYE10</td>
<td>27.8</td>
<td>26.7</td>
<td>31.5</td>
<td>28.8</td>
<td>29.2</td>
<td>26.7</td>
<td>29.1</td>
<td>27.6</td>
<td>24.6</td>
<td>100</td>
</tr>
</tbody>
</table>
Figure S1. Multiple sequence alignment of McOYE1-9 with the templates used for homology modeling. The red box highlights the HX₁X₂H/N motif, important for substrate binding. The blue box shows the conserved YGGS motif, present in a loop region important for substrate binding.
Figure S2. Multiple sequence alignment of McOYE10 with the templates used for homology modeling. The red box highlights the HX₁X₂H/N motif, important for substrate binding. The blue box shows the conserved YGGS motif, present in a loop region important for substrate binding.
Table S2. Summary of the main parameters obtained from validation of McOYE homology models through PROCHECK, QMEAN, ProSA and Verify3D servers.

<table>
<thead>
<tr>
<th>McOYE isoform</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROCHECK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Main chain parameters: Ramachandran plot</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residues in most favoured and allowed regions (%)</td>
<td>96.4</td>
<td>96.7</td>
<td>97.4</td>
<td>97.4</td>
<td>96.4</td>
<td>96.5</td>
<td>97.8</td>
<td>97.8</td>
<td>97.9</td>
<td>98.5</td>
</tr>
<tr>
<td>Residues in generously allowed regions (%)</td>
<td>3</td>
<td>2.3</td>
<td>2.3</td>
<td>2</td>
<td>1.9</td>
<td>2.5</td>
<td>1.6</td>
<td>1</td>
<td>1.9</td>
<td>0.9</td>
</tr>
<tr>
<td>Residues in disallowed regions (%)</td>
<td>0.7</td>
<td>1</td>
<td>0.3</td>
<td>0.7</td>
<td>1.6</td>
<td>0.9</td>
<td>0.6</td>
<td>1.3</td>
<td>0</td>
<td>0.6</td>
</tr>
<tr>
<td>G-factor (overall)</td>
<td>-0.37</td>
<td>-0.36</td>
<td>-0.45</td>
<td>-0.45</td>
<td>-0.44</td>
<td>-0.41</td>
<td>-0.37</td>
<td>-0.44</td>
<td>-0.38</td>
<td>-0.33</td>
</tr>
<tr>
<td>Side chain parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M/c bond lengths</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>within limits</td>
<td>98.6</td>
<td>98.5</td>
<td>97.2</td>
<td>98.1</td>
<td>97.6</td>
<td>97.5</td>
<td>98.6</td>
<td>98.3</td>
<td>96.6</td>
<td>98.9</td>
</tr>
<tr>
<td>(%)</td>
<td>M/c bond angles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>highlighted</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(%)</td>
<td>1.4 1.5 2.8 1.9 2.4 2.5 1.4 1.7 3.4 1.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**within limits**

<table>
<thead>
<tr>
<th>(%)</th>
<th>Planar groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>highlighted</td>
<td></td>
</tr>
<tr>
<td>(%)</td>
<td>8.6 8.9 11.7 10.2 9.2 10.7 8.5 8.2 9.8 8.0</td>
</tr>
</tbody>
</table>

**within limits**

<table>
<thead>
<tr>
<th>(%)</th>
<th>QMEAN, ProSA and VERIFY 3D</th>
</tr>
</thead>
<tbody>
<tr>
<td>highlighted</td>
<td></td>
</tr>
<tr>
<td>(%)</td>
<td>QMEAN norm score</td>
</tr>
<tr>
<td></td>
<td>0.73 0.81 0.69 0.73 0.70 0.69 0.74 0.72 0.70 0.74</td>
</tr>
<tr>
<td></td>
<td>-0.47 -0.39 -0.90 -0.54 -0.86 -0.92 -0.38 -0.58 -0.87 -0.31</td>
</tr>
<tr>
<td></td>
<td>89.69 89.14 89.66 98.04 89.53 94.02 93.00 92.84 88.71 94.35</td>
</tr>
</tbody>
</table>

\(^a\) % amino acids have scored >= 0.2 in the 3D/1D profile.
Figure S3 LigPlot graphs for the analysis of the FMN binding site in McOYE1-10 (from panel A to J, respectively) homology model
Figure S4. Docking of 2-cyclohexenone in McOYE models. Results form the docking of the substrate 2-cyclohexenone in the active site of the ten putative McOYEs (McOYE1 to 10 from panel A to J, respectively) with the three residues important for substrate binding and catalysis, the FMN cofactor and the substrate superimposed with the crystal structure of pentaerythritol tetranitrate reductase complexed with 2-cyclohexenone (PDB ID 1GVQ, brown color)