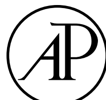

ACADEMIC PRESS

Archives of Biochemistry and Biophysics 403 (2002) 277–283

ABB

www.academicpress.com

Molecular cloning and characterization of a full-length flavin-dependent monooxygenase from yeast

Man Zhang and Jon D. Robertus*

Department of Chemistry and Biochemistry, Institute of Cellular and Molecular Biology, University of Texas, Austin, TX 78712, USA

Received 10 April 2002, and in revised form 20 May 2002

Abstract

Eucaryotes contain a class of enzymes called flavin-dependent monooxygenases (FMOs). Unlike mammals, yeast have only a single isoform—yFMO. Deletion mutants suggested that yFMO may play a role in folding proteins which contain disulfide bonds. Recently we detected two nucleotide errors in the GenBank sequences attributed to the yFMO gene. This previously led us to express and characterize a 373-residue catalytically active protein instead of the correct 432-residue enzyme. Here we report the sequencing, expression, and enzyme characterization of the full-length form of yFMO. Comparison of the two forms of yFMO showed similar pH profiles and K_m, K_{cat}, and V_{max} values using glutathione as a substrate. These results indicate that the full-length yeast FMO has biochemical and catalytic properties similar to those of the truncated protein. Therefore, it is likely that the hypotheses concerning the enzyme's function proposed earlier are still valid. © 2002 Elsevier Science (USA). All rights reserved.

Keywords: Flavin-dependent oxygenase; Kinetic parameters; Sequence error; Protein expression

In the past three decades, a group of flavin-containing monooxygenases (FMOs)[1] have been identified in many mammals, including humans [1], pigs [2], and rabbits [3,4]. At least five common isozymes have also been discovered in various organs and tissues, including lung, liver, and kidney [5]. Studies on pig liver FMO [6] have provided a model of the catalytic mechanism and substrate specificity for FMOs. The major catalytic steps are summarized in Fig. 1 [5]: NADPH reduces the FAD bound in FMO to $FADH_2$; molecular oxygen then oxidizes the $FADH_2$ to 4α-hydroperoxyflavin, FAD-HOOH. Incoming substrate (SH) is oxygenated to SOH, while the enzyme-bound peroxide is reduced to hydroxyflavin, FADHOH, and, finally, water and $NADP^+$ are released. The intermediate $FADO_2H_2$ is quite stable, giving a characteristic spectrum [7,8]. Another interesting feature of pig liver FMO is the broad substrate specificity. The enzyme recognizes many secondary and tertiary amines, hydrazines, and organic sulfur compounds [6]. Even though other mammalian FMOs differ in substrate specificity, they generally can recognize a range of substrates. Given the fact that many substrates of mammalian FMOs are xenobiotics and the high concentration of the enzymes in liver, it was suggested that the physiological function of mammalian FMOs was to oxidize xenobiotic nucleophiles [5].

In recent years, proteins sharing sequence homology with mammalian FMOs were identified from several nonmammalian eukaryotes and assigned different physiological functions [9–11]. For example, *YUCCA* from *Arabidopsis* contains putative FAD and NADPH binding regions and was found to be involved in auxin biosynthesis, catalyzing the crucial step of N-oxygenation of tryptamine [12].

Based on the known genomic sequence of *Saccharomyces cerevisiae* [13], there is only one FMO-like protein in yeast, yFMO. We used the genomic sequence to clone and express that open reading frame (ORF) in *Escherichia coli* [14]. The purified yeast enzyme differed from the mammalian FMOs in that it recognized only certain biological thiols as substrates; these included glutathione, cysteine, and cysteamine. Subsequent gene knockout and other experiments suggested that the yeast FMO-like protein was important for the correct folding of proteins which contained disulfide bonds. It was

* Corresponding author. Fax: +1-512-471-8696.
E-mail address: jrobertus@mail.utexas.edu (J.D. Robertus).

[1] *Abbreviations used:* FMO, flavin-containing monooxygenases; ORF, open reading frame; ER, endoplasmic reticulum.

0003-9861/02/$ - see front matter © 2002 Elsevier Science (USA). All rights reserved.
PII: S0003-9861(02)00237-0

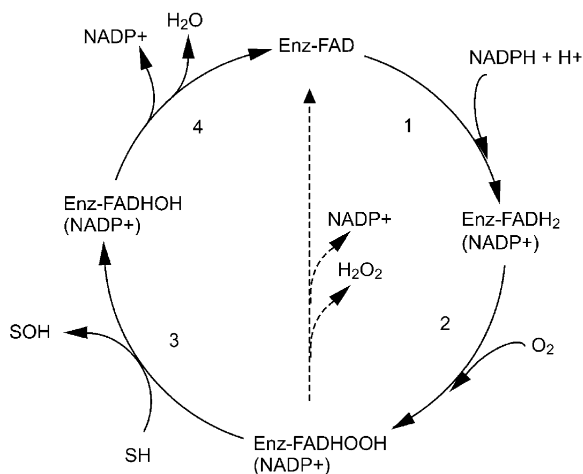


Fig. 1. The catalytic cycle of FMO. The dashed arrow shows the substrate-independent NADPH oxidase or "blank" rate.

proposed that it functioned by generating oxidizing equivalents that ultimately influenced the environment in the endoplasmic reticulum [15–17].

In continuing our studies, we found that there is an error in the original yeast FMO-like ORF sequence in GenBank. This error, when incorporated into our cloning strategies, caused the expression of a modified yFMO, which we will refer to as the truncated form, or yFMOt. That protein lacks 38 amino acids at the C terminus. Also, due to an open reading frame shift, caused by an extra G at position 1080, the presumed stop codon was read through, adding 36 vector amino acids to the protein.

In this paper, we report the cloning of the true full-length yFMO cDNA, its expression in *E. coli*, and the comparison of substrate specificity and kinetic parameters to the previously reported truncated enzyme.

Material and methods

Yeast strains and growth conditions

The *S. cerevisiae* strains used were X2180 and TR2. They were grown at 28 °C in YPD medium, which contains 1% yeast extract, 2% Bactopeptone, and 2% glucose.

DNA sequencing and cloning of the yFMO gene

Genomic DNA was isolated from *S. cerevisiae* strains X2180 and TR2 using Dneasy Tissue kit (Qiagen). In all PCR amplifications, *Pfu* DNA polymerase (Stratagene) was used to minimize the error rate. To facilitate DNA sequencing, the *fmo* gene, including upstream and downstream regions, was amplified. The forward primer was (5′-GCAAGATAATCCTGTAGC-3′), and the reverse primer was (5′-TACAAGTACATTCCTTTTTAC G-3′). The 1.5-kb PCR product included 163-bp upstream and 24-bp downstream non-translating regions and the 1.3-kb coding region. The PCR product was recovered from DNA agarose gel using QIAquick Gel Extraction kit (Qiagen) and then sequenced.

To clone only the coding region of yFMO gene, another pair of primers was used: 5′-GAAGTAACT ACATATGACAGTGAATGAC-3′ and 5′-CATGAGG GAAACGCTCGAGTAAGCGGCAC-3′. That PCR product was recovered, digested with *Nde*I and *Xho*I, and inserted into pET16b vector (Novagen). The sequence of the recombinant construct, pET16b-fyFMO, was verified by restriction digestions and DNA sequencing. DNA autosequencing was performed at the ICMB facility center at the University of Texas at Austin.

Expression of the recombinant N-terminal His-tag yFMO protein in E. coli

The plasmid pET16b-fyFMO was transformed into *E. coli* strain HMS174(DE3) cells. A single transformant colony containing the plasmid was first grown overnight in 10 ml LB media (containing 10% NaCl, 10% tryptone, 5% yeast extract) with 100 μg/ml ampicillin at 37 °C. The overnight cell culture was then diluted into 1 liter fresh LB media, containing 100 μg/ml ampicillin, and grown at 37 °C. When the cell culture reached an A_{600} of 0.5, it was induced with 1.0 mM isopropyl-β-D-thiogalactopyranoside (Gold Biotechnology). After growth at 30 °C for 4 h, the cells were harvested by centrifugation at 7000*g* for 5 min at 4 °C. Cell pellets were stored at −20 °C.

Purification of the recombinant N-terminal His-tag yeast FMO protein

The purification process was undertaken at 4 °C. Fresh or frozen cell pellets (about 12 g wet weight) were resuspended in 20 ml cold lysis buffer (100 mM sodium phosphate, pH 7.8, 1 mM phenylmethylsulfonyl fluoride) and lysed twice in a precooled French Press cell (SLM Aminco, Urbana, IL) at 700 psi. The broken cells were subjected to centrifugation at 70,000*g* for 40 min at 4 °C, and the supernatant was loaded onto a 3-ml Ni–NTA agarose (Novagen) column equilibrated with cold lysis buffer. Once the sample was loaded, the column was run in the dark. It was washed with lysis buffer and wash buffer (lysis buffer with 80 mM imidazole) until the A_{280} of the flow-through was lower than 0.01. His-tagged yeast FMO protein was eluted from the column with lysis buffer of 0.1–1.0 M imidazole gradient. Elution fractions were collected and analyzed on 10% SDS–PAGE. Fractions containing His-tagged yeast FMO protein were pooled and dialyzed into storage buffer (50 mM sodium phosphate, 10% glycerol, pH 7.0).

Protein concentration was determined by the Bradford method using bovine serum albumin as the standard.

Preparation of apo-enzyme

About 5 mg of purified His-tagged yFMO protein in 10 ml volume was dialyzed into 2 liters of semidenaturing buffer (50 mM sodium phosphate, pH 7.0, 2 M urea, 2 M KBr) at 4 °C, for 12–24 h; the dialysis buffer was changed twice during this interval. Then the protein was dialyzed into 2 liters storage buffer (50 mM sodium phosphate, 10% glycerol, pH 7.0) at 4 °C for 12–24 h, again using two changes of buffer.

Measurement of FAD concentration in holoenzyme

Five hundred microliters of yFMO protein sample was incubated with 50 μl 3 M trichloroacetic acid, on ice, in the dark, for 5 min. The mixture was centrifuged at 12,000 rpm for 10 min at room temperature, and 100 μl 1 M K_2HPO_4 and 20 μl 6 M KOH were added to the supernatant. The OD_{450} was measured before and after reduction with sodium dithionite to reduce FAD. To obtain the concentration of FAD bound in the holoenzyme, the difference of OD_{450} was divided by the extinction coefficient of FAD, $11.3\,mM^{-1}$.

Oxygen uptake assay

The recombinant yFMO activity was measured by substrate-dependent oxygen uptake assay [18] at 37 °C. The assay was performed in a 2-ml thermostated oxygraph chamber (Gibson Medical Electronics) with a Clark-type oxygen electrode (Yellow Springs Inc.). The signals were recorded by an EU-200-02 DC offset module. The reaction mixture contained 0.2 M sodium phosphate buffer with 0.5 mM EDTA, 50 μM FAD, and the NADPH regenerating system (0.2 mM $NADP^+$, 1.5 mM glucose 6-phosphate, and 1 U glucose-6-phosphate dehydrogenase) in a 2 ml volume at pH 6.4. After a 2 min incubation, the enzyme was added, and the substrate-independent oxygen uptake was recorded for about 2 min. (This measures the "blank rate" oxidase activity.) Then the organic substrate was added and the substrate-dependent oxygen uptake was recorded for about 4 min. To measure the pH profile of NADPH oxidase activity, assays were performed at pH 5.5, 6.0, 6.5, 7.0, 7.5, and 8.0. To measure the K_m values for NADPH and FAD, the NADPH oxidase activity was monitored at pH 6.4 in the absence of any substrate. Apo-enzyme was used in the assay, with a range of FAD concentrations, to measure the K_m for FAD. In the assay to determine the K_m for NADPH, 50 μM FAD was introduced and the usual NADPH regenerating system was substituted by a range of NADPH concentrations. To calculate substrate-dependent oxygen uptake, in addition to the substrate-independent NADPH oxidase activity, the background from the substrate itself was subtracted from the total rate.

FPLC Superdex 200 gel filtration analysis

After the Superdex 200 (Amersham-Pharmacia Biotech.) column was equilibrated with 10 column volumes of the column buffer (0.1 M sodium phosphate, 0.25 M sucrose, pH 6.5) at room temperature, purified yFMO protein was loaded onto the column. The column was run at a flow rate of 0.5 ml/min and 0.5-ml fractions were collected. Fractions were analyzed on 10% SDS–PAGE. The standard curve for the Superdex 200 was calculated from the elution profile of the protein standards thyroglobulin, IgG 158, ovalbumin, myoglobin, and cyanocobalamin (Bio-Rad) with molecular weights of 670, 158, 44, 17, and 1.35 kDa, respectively. Protein elution was monitored by absorbance at 280 nm.

Results

Nucleotide sequence analysis

The DNA sequence of *S. cerevisiae* FMO-like cDNA is located on chromosome VIII, nucleotides 454227 to 455348 (GenBank NC_001140). When manual sequencing was performed during the previous work, no difference between the GenBank sequence and the locally cloned cDNA was detected [14]. In the process of cloning the yeast FMO cDNA into other *E. coli* expression vectors, one difference was observed at position 1080 from the automatic sequencing of the cDNA clone, as shown in Fig. 2. Our locally cloned yFMOt cDNA had an extra G base compared to the GenBank sequence (NC_001140, position 455306 of *S. cerevisiae* chromosome VIII). To verify that this difference was not introduced by the cloning process, genomic DNA from two *S. cerevisiae* strains, X2180 and TR2, was used as the template to amplify the FMO-like gene by PCR. Both PCR products had the same sequence as the newly

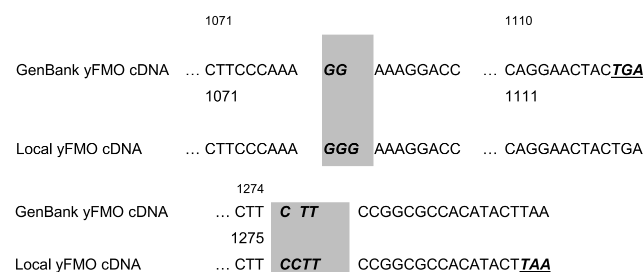


Fig. 2. Partial sequence alignment of the locally sequenced yeast FMO cDNA and the GenBank sequence. The differences between them are shaded; stop codons are underlined.

```
      1              15|             30              45              60              75              90
yFMO  MTVNDKKRLAIIGGG PGGLAAARVFSQSLP NFEIEIFVKDYDIGG VWHYPEQKSDGR-VM YDHLETNISKKLMQF SGFPFEENVPLYPSR 89
yFMOt MTVNDKKRLAIIGGG PGGLAAARVFSQSLP NFEIEIFVKDYDIGG VWHYPEQKSDGR-VM YDHLETNISKKLMQF SGFPFEENVPLYPSR 89
pFMO1 ----MAKRVAIVGAG VSGLASIKCCLEEG- -LEPTCFERSDDLGG LWRFTEHVEEGRASL YKSVVSNSCKEMSCY PDFPFPEDYPNYVPN 84
hFMO1 ----MAKRVAIVGAG VSGLASIKCCLEEG- -LEPTCFERSDDLGG LWRFTEHVEEGRASL YKSVVSNSCKEMSCY SDFPFPEDYPNYVPN 84

      91             105             120             135             150             165             180
yFMO  RNIWEYLKAYYKTFI ANKDAISIHFSTEVT YLKKKNS--QWEITS KDELRTTKSDFDFVI VASGHYSVPKLPTNI AGLDLWFDNKGAFHS 177
yFMOt RNIWEYLKAYYKTFI ANKDAISIHFSTEVT YLKKKNS--QWEITS KDELRTTKSDFDFVI VASGHYSVPKLPTNI AGLDLWFDNKGAFHS 177
pFMO1 SHFLEYLRMYANQFN LLKCIQFKTKVCSVT KHEDFNTTGQWDVVT LCEGKQESAVFDAVM VCTGFLTNPYLPLDS FPGINTFKG-QYFHS 173
hFMO1 SQFLEYLKMYANHFD LLKHIQFKTKVCSVT KCSDSAVSGQWEVVT MHEEKQESAIFDAVM VCTGFLTNPYLPLDS FPGINAFKG-QYFHS 173

      181            195             210             225             240             255             270
yFMO  KDFKNCEFAREKVVI VVGNGSSGQDIANQL TTVAKKVYNSIK--- -----------EP-- ASN---------QLK AKLIETVQTIDSADW 242
yFMOt KDFKNCEFAREKVVI VVGNGSSGQDIANQL TTVAKKVYNSIK--- -----------EP-- ASN---------QLK AKLIETVQTIDSADW 242
pFMO1 RQYKHPDIFKDKSVL VVGMGNSGTDIAVEA SHLAKKVFLSTTGGA WVISRVFDSGYPWDM VFMTRFQNMFRNSLP TPIVNWLIAKKMNSW 263
hFMO1 RQYKHPDIFKDKRVL VIGMGNSGTDIAVEA SHLAEKVFLSTTGGG WVISRIFDSGYPWDM VFMTRFQNMLRNSLP TPIVTWLMERKINNW 263

      271            285             300             315             330             345             360
yFMO  KNRS----------- --------------- --------------- --VTLSDGRVLQNID YIIFATGYYYSFPFI EPSVRLEVLGEGVTG 289
yFMOt KNRS----------- --------------- --------------- --VTLSDGRVLQNID YIIFATGYYYSFPFI EPSVRLEVLGEGVTG 289
pFMO1 FNHANYGLIPEDRIQ LREPVLNDELPGRII TGKVLIKPSIKEVKE NSVVFNSSPEEEPID IIVFATGYTFAFPFL DESVVKVEDGQASLY 353
hFMO1 LNHANYGLIPEDRTQ LKEFVLNDELPGRII TGKVFIRPSIKEVKE NSVIFNNTSKEEPID IIVFATGYTFAFPFL DESVVKVEDGQASLY 353

      361            375             390             405             420             435             450
yFMO  DKHSSVNLHNLWEHM IYVKDPTLSFILTPQ LVIPFPLSELQAAIM VEVFCKSLPITTTFD SNACG--THNFPKGK DLEYYAELQELLNSI 377
yFMOt DKHSSVNLHNLWEHM IYVKDPTLSFILTPQ LVIPFPLSELQAAIM VEVFCKSLPITTTFD SNACG--THNFPKGK DLEYYAELQELLNSI 377
pFMO1 KYIFPAHLQKPTLAV IGLIKPLGSLLPTGD TQARWAVRVLKGVNK LPPSSVMIQEVNTRK ENKPSGFGLCYCKAL QSDYIAYIDELLTYI 443
hFMO1 KYIFPAHLQKPTLAI IGLIKPLGSMIPTGE TQARWAVRVLKGVNK LPPPSVMIEEINARK ENKPSWFGLCYCKAL QSDYITYIDELLTYI 443

      451            465             480             495             510             525            540
yFMO  PRRVGHFEPVVWDDR LIDLRNS-------- ----SYTDKEERNVL LAEHAQALKKKKAPY FLPAPHT-------- --------------  432
yFMOt PRRVGHFEPVVWDDR LIKPNSA-------- ----DIHHTGGRSRI RLLTKPERKLSWLLP PLSNN---------- --------------  430
pFMO1 DAKPNMFSLLLTDPH LALTIFFGPCTPYQF RLTGPGKWEGARNAI MTQWDRTFKVTKTRI VKESPSPFASLLKLF SFLALLVAIFQIFL  532
hFMO1 NAKPNLFSMLLTDPH LALTVFFGPCSPYQF RLTGPGKWEGARNAI MTQWDRTFKVIKARV VQESPSPFESFLKVF SFLALLVAIFLIFL  532
```

Fig. 3. Alignment of FMOs. The full-length yFMO is marked as yFMO, the truncated form as yFMOt, porcine FMO1 [25] as pFMO1, and human FMO1 [26] as hFMO1. The alignment was made using ClustalW. Absolute position numbers are marked along the top of each row, and marker sequence numbers of specific FMOs are shown at the right end of rows. Conserved signature sequences, described in the text, are shaded, as is the bar at absolute position 467 which marks the end of sequence identities between yFMO and yFMOt.

determined yeast FMO cDNA clone. Further sequencing found that a C base was missing from the GenBank sequence at position 1277 (position 455503 of *S. cerevisiae* chromosome VIII). This C base is after the originally proposed stop codon in the GenBank sequence corresponding to the FMO open reading frame. As a consequence of these changes, the correct *S. cerevisiae* FMO cDNA was found to contain 1299 bp nucleotides, encoding 432 amino acids rather than 373 amino acids as originally indicated by the GenBank sequence.

The newly determined protein sequence is aligned with the original yFMOt sequence and with human and pig FMO protein sequences in Fig. 3. The full-length yFMO and yFMOt are identical for the first 394 amino terminal residues. The last identical residue is marked by a shaded bar in Fig. 3 at absolute postion 467. All conserved domains and motifs found previously [14], including the hypothetical FAD signature sequence (absolute position 13–18), the $NADP^+$ signature sequence (absolute position 198–203), the "FATGY" sequence (absolute position 334–338), and the "D/QELL" sequence (absolute position 444–447) are retained. After yFMO residue 394, however, the new and old yFMOt sequences diverge.

Expression and purification of His-tagged full-length yeast FMO (His_{10}-fyFMO) protein

Since kinetic measurements of yFMOt involved a His-tagged protein, a similar construction was made for the new full-length yFMO. pET16b-fyFMO was chosen to express the His_{10}-fyFMO protein in *E. coli*. Unlike yFMOt, the full-length clone does not require plasmid containing GroEL and GroES to be expressed as soluble fractions in *E. coli* [14]. The purification process was performed in darkness to minimize the photoreduction of free FAD. About 20 mg of purified His_{10}-yFMO protein could be obtained from 1 liter of cell culture, whereas about 10 mg purified truncated (His_{10}-yFMOt) protein was obtained under the same condition. As shown in Fig. 4, the molecular weight of the His_{10}-yFMO is estimated as 50 kDa, close to the molecular weight predicted from the true ORF. While the His_{10}-yFMOt proteins tended to precipitate after being purified, the full-length His_{10}-yFMO remained soluble.

During the purification process, it was observed that the cell lysate with the full-length FMO protein had a bright yellow color in contrast to the lysate containing yFMOt. The full-length protein retained its yellow color even after elution from the Ni–NTA column, whereas the yFMOt protein required 50 μM exogenous FAD to regain activity. This indicates that the His_{10}-fyFMO protein was able to bind FAD more tightly than the yFMOt form, in *E. coli* and throughout the purification protocol.

Comparisons of enzymatic features in the full-length and truncated yeast FMOs

In the absence of organic substrates, FMOs are known to have an NADPH oxidase activity, as indi-

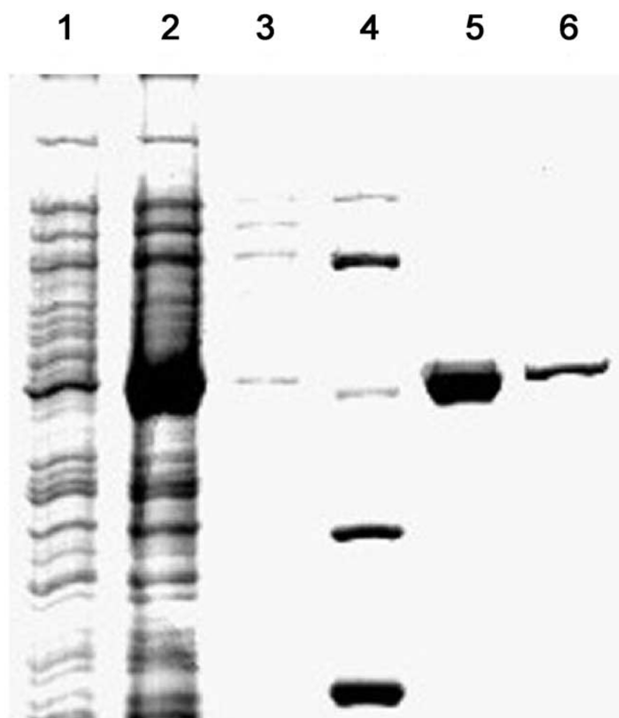
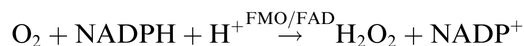


Fig. 4. Expression and isolation of His_{10}-yfMO analyzed by SDS–PAGE. Lane 1, total protein from uninduced *E. coli* with pET16b-yfFMO; lane 2, total protein from *E. coli* with pET16b-yfFMO after induction with IPTG; lane 3, the soluble fraction from *E. coli* with pET16b-yfFMO induced by IPTG; lane 4, protein molecular weight standards (94, 67, 43, 30, 21 kDa); lane 5, purified full-length His_{10}-yfMO protein from the Ni–NTA column (20 μg loaded); lane 6, purified truncate protein His_{10}-yfMOt from the Ni–NTA column.

cated in Fig. 1. Rates vary with pH but are generally around 20 mol/mol/min [14]. With fixed concentrations of NADPH (0.2 mM) and FAD (50 μM), the pH profiles of NADPH oxidase activity of both proteins were obtained. Both protein forms had higher NADPH oxidase activity at lower pH, suggesting that protons are reactants, as shown below.



$$O_2 + NADPH + H^+ \xrightarrow{FMO/FAD} H_2O_2 + NADP^+$$

To measure the K_m for FAD, apo-protein of both forms was prepared at pH 6.4, the cytoplasmic pH in yeast cells [19]. Exogenous FAD was titrated in at 0–100 μM, and the NADPH oxidase activity was measured (data not shown). The K_m for FAD of the full-length His_{10}-fyFMO was 12.9 μM and the K_m for FAD of the truncated His_{10}-yFMOt was 17.8 μM.

As reported previously [14], His_{10}-yFMOt recognizes only certain biological thiols as substrates and, unlike many mammalian FMO isozymes, was not active against nitrogen-containing organic molecules. To find out whether the full-length His_{10}-fyFMO has the same substrate specificity, a few key compounds were tested. These included *N,N*-dimethylaniline and trimethylamine, typical nitrogen-containing substrates for mammalian FMOs [5], and tryptamine, a substrate for *Arabidopsis* FMO-like protein [12]. None of the N-containing compounds was recognized as a substrate by the full-length His_{10}-fyFMO protein. However, the thiol-containing molecules, cysteamine and glutathione, were good substrates, as they were for yFMOt. Therefore, we conclude that the full-length yeast FMO shares the same type of substrates as the truncated enzyme.

Table 1
Steady state kinetic parameters for two forms of yFMO using glutathione as a substrate

Enzyme	pH optimum	K_m (mM)	V_{max} (μM/min)	k_{cat} (min^{-1})
His_{10}-yFMO	6.4	16.7	15.6	42
His_{10}-yFMOt	6.4	22.7	18.2	48

Glutathione is very concentrated in cellular cytoplasm and is likely to be oxidized by yFMO. We therefore focused our kinetic studies on glutathione. The pH profile of activity against glutathione was measured. The profile for the full-length yFMO is similar to that of the previously measured yFMOt (data not shown). The optimal pH was about pH 6.4, which is very close to the cytoplasmic pH in yeast cells. Initial rate data for both yFMO and yFMOt were measured. The kinetic constants are very similar, as summarized in Table 1.

Yeast FMO protein exists as monomer

The active form of pig liver FMO was known to exist as dimers in solution (personal communication with Dr. D.M. Ziegler). To determine in what forms the yeast FMO exists, size exclusion chromatography was performed. At 2 mg/ml, in 0.1 M sodium phosphate and 0.25 M sucrose, at pH 6.5, both the full-length and the truncated yeast FMO proteins existed as monomers.

Discussion

Our studies show that the yeast genomic sequence reported in GenBank contains two nucleotide errors in the gene coding for the protein called yFMO. Using the published GenBank sequence to design PCR primers caused us to amplify a truncated, and C-terminally modified, form of this enzyme, which we now call yFMOt. The initial clone of yFMOt, based on the GenBank sequence, was thought to contained 1119 bases encoding 373 amino acids, a stop codon, and 60 non-coding bases downstream of the stop codon. Due to the extra G base at position 1080, however, the stop codon was altered and the open reading frame continued about 110 bp nucleotides further downstream until it reached a stop codon in the expression vector pET16. As a consequence, rather than expressing the anticipated

373-amino acid protein, the yFMOt protein in fact contained 430 amino acids including 36 non-FMO amino acids at the C terminus. The two forms of yFMO have identical N-terminal 394 residues, but differ in the length and composition of the C terminus.

After identifying the error in the gene sequence, we amplified the full-length yFMO gene, cloned it, and expressed it as a His-tagged protein. The most noticeable physical difference between the full-length and the truncate FMO proteins is the expression level and solubility of the proteins. Full-length yFMO is expressed at higher levels and remains soluble longer than the yFMOt. This is indicative of a more stable protein fold for the intact protein. A second and probably related difference is their affinity for FAD. The full-length yFMO is apparently more stable and has a superior FAD binding domain, even though the amino acid sequence suggests that FAD is bound in the N-terminal part of the structure conserved in both forms. It may be that the altered C-terminal regions affect the global folding in a subtle way that weakens the affinity for flavin. This may be due to a lack of native residues that normally participate in cofactor binding, or it may reflect some deleterious interactions with the incorrect C-terminal residues of yFMOt.

Having previously done an extensive enzymological characterization of yFMOt [14–17], we were interested to know how the full-length yFMO compared with it. It appears that the two forms are very similar in enzymatic activity. Both enzymes carry out the substrate-independent oxidase activity, at about the same rate. More importantly, neither form appears to oxidize nitrogen-containing molecules, which are the major substrates of several mammalian and nonmammalian FMO isoforms. Both forms of yFMO, instead, oxidize biological thiols, including cysteamine and glutathione. A steady state kinetic analysis shows that the two forms have similar K_m and k_{cat} values for glutathione (Fig. 5 and Table 1). This suggests that the active site is largely unperturbed in the truncated form, at least as far as the oxidation of small thiols is concerned.

Deleting the yFMO gene from yeast allowed us to make a number of phenotypic observations, mainly that the deletion strains could not fold endogenous [16] or exogenous [15] proteins which contained disulfide bonds. This observation depended only on deleting the FMO gene and so is independent of the form of protein subsequently expressed for kinetic measurements. The role of yFMO in folding disulfide-bonded proteins therefore remains unchanged. As part of that study, we showed that the deletion mutant could be complemented by a plasmid expressing yFMOt. We have not shown that full-length yFMO would also complement the deleted gene, but it is very reasonable to assume that it would be even more effective in this regard.

Although we have presumptive evidence that yFMO affects the folding of disulfide-bonded proteins, its exact cellular function remains unclear. Previously, we suggested that by oxidizing glutathione to glutathione disulfide, yFMO could regulate the redox potential in the ER. Consistent with this idea we observed that wild-type yFMO and the plasmid-expressed yFMOt are feedback-inhibited by increasing redox potential [17]. Recently, it was reported [20] that yeast Ero1p binds FAD and suggested that Ero1p and PDI-mediated oxidative protein folding depends on the FAD level in the cell. The ultimate source of oxidizing potential, however, is still unclear. Therefore, it may be that yFMO, which uses molecular O_2 as its ultimate oxidizing source, may communicate with those proteins in the ER lumen by as yet unknown mechanisms. It may be that the true substrate for yFMO is an ER protein that can shuffle oxidizing equivalents into the ER. It could also be that the full-length protein, with an intact C-terminal domain, may allow us to identify a wider range of substrates, including protein targets.

FMOs are well known to associate with microsomal membranes, and we showed that wild-type yFMO is also membrane associated [15]. It has been suggested that the C terminus of FMOs may be involved in membrane binding, but this was disproven by deletion of the 26 C-terminal residues of the rabbit FMO 1B1, which showed no effect on cellular localization [21]. Our earlier work is consistent with this finding, because we showed that yFMOt, with its shortened and modified C terminus, also associated with the cytoplasmic surface of the ER [15].

Because yeast contain only a single FMO, we proposed that yFMO represents the original ancestral activity of this family of enzymes [14]. In contrast to yeast, mammals contain as many as five FMO isozymes [22]. The best-characterized forms, like the FMO1s shown in Fig. 3, play important roles in the metabolism of both nitrogen-containing and sulfur-containing xenobiotic compounds [5,23,24]. These mammalian enzymes are substantially larger than yFMO, having about 100 additional amino acids. These are largely found as major insertions, such as the 43-residue insertion at absolute position 275 in Fig. 3. It may be that this or other differences form part of the substrate recognition site which governs the substrate specificity of the mammalian enzymes.

Acknowledgments

We thank Dr. D.M. Ziegler and Dr. L.L. Poulsen for helpful discussions and technical support in the oxygen uptake assays. This work was supported by Grant GM 30048 from the National Institutes of Health, by grants from the Foundation for Research and the Welch Foundation, and by the College of Natural Sciences support to the Center for Structural Biology.

References

[1] G.A. Dannan, F.P. Guengerich, Mol. Pharmacol. 22 (1982) 787–794.
[2] D.M. Ziegler, C.H. Mitchell, Arch. Biochem. Biophys. 150 (1972) 116–125.
[3] D.E. Williams, D.M. Ziegler, D.J. Nordin, S.E. Hale, B.S. Masters, Biochem. Biophys. Res. Commun. 125 (1984) 116–122.
[4] R.E. Tynes, P.J. Sabourin, E. Hodgson, Biochem. Biophys. Res. Commun. 126 (1985) 1069–1075.
[5] D.M. Ziegler, Drug Metab. Rev. 19 (1988) 1–32.
[6] L.L. Poulsen, D.M. Ziegler, Chem–Biol Interact. 96 (1995) 57–73.
[7] N.B. Beaty, D.P. Ballou, J. Biol. Chem. 256 (1981) 4611–4618.
[8] N.B. Beaty, D.P. Ballou, J. Biol. Chem. 256 (1981) 4619–4625.
[9] M. Agosin, G.T. Ankley, Drug Metab. Dispos. 15 (1987) 200–203.
[10] D. Schlenk, D.R. Buhler, Comp. Biochem. Physiol. C Comp. Pharmacol. Toxicol. 94 (1989) 469–475.
[11] B. Kurelec, Biochem. Biophys. Res. Commuu. 127 (1985) 773–778.
[12] Y. Zhao, S.K. Christensen, C. Fankhauser, J.R. Cashman, J.D. Cohen, D. Weigel, J. Chory, Science 291 (2001) 306–309.
[13] M. Johnston, S. Andrews, R. Brinkman, J. Cooper, H. Ding, J. Dover, Z. Du, A. Favello, L. Fulton, S. Gattung, et al., Science 265 (1994) 2077–2082.
[14] J.K. Suh, L.L. Poulsen, D.M. Ziegler, J.D. Robertus, Arch. Biochem. Biophys. 336 (1996) 268–274.
[15] J.K. Suh, L.L. Poulsen, D.M. Ziegler, J.D. Robertus, Proc. Natl. Acad. Sci. USA 96 (1999) 2687–2691.
[16] J.K. Suh, J.D. Robertus, Proc. Natl. Acad. Sci. USA 97 (2000) 121–126.
[17] J.K. Suh, L.L. Poulsen, D.M. Ziegler, J.D. Robertus, Arch. Biochem. Biophys. 381 (2000) 317–322.
[18] L.L. Poulsen, D.M. Ziegler, J. Biol. Chem. 254 (1979) 6449–6455.
[19] J.M. Salhany, T. Yamane, R.G. Shulman, S. Ogawa, Proc. Natl. Acad. Sci. USA 72 (1975) 4966–4970.
[20] B.P. Tu, S.C. Ho-Schleyer, K.J. Travers, J.S. Weissman, Science 290 (2000) 1571–1574.
[21] M.P. Lawton, R.M. Philpot, J. Biol. Chem 268 (1993) 5728–5734.
[22] D.M. Ziegler, Annu. Rev. Pharmacol. Toxicol. 33 (1993) 179–199.
[23] D.M. Ziegler, Trends Pharmacol. Sci. 11 (1990) 321–324.
[24] E. Hodgson, P.E. Levi, Drug Metabol. Drug Interact. 6 (1988) 219–233.
[25] R. Gasser, R.E. Tynes, M.P. Lawton, K.K. Korsmeyer, D.M. Ziegler, R.M. Philpot, Biochemistry 29 (1990) 119–124.
[26] C. Dolphin, E.A. Shephard, S. Povey, C.N. Palmer, D.M. Ziegler, R. Ayesh, R.L. Smith, I.R. Phillips, Biol. Chem. 266 (1991) 12379–12385.