Published in final edited form as:

Methods Mol Biol. 2013 ; 987: . doi:10.1007/978-1-62703-321-3_6.

Identification of endogenous substrates of orphan cytochrome P450 enzymes through the use of untargeted metabolomics approaches

Qian Cheng and **F. Peter. Guengerich**

Vanderbilt University School of Medicine, Department of Biochemistry, 638 Robinson Research Building, 2200 Pierce Avenue, Nashville, Tennessee 37232 U.S.A

Summary

Metabolomics provides an invaluable means to interrogate the function of “orphan” enzymes, i.e., those whose endogenous substrates are not known. Here we describe a high performance liquid chromatography-coupled mass spectrometry (HPLC-MS)-based metabolomics approach to identify an endogenous substrate of an orphan cytochrome P450.

Keywords

orphan enzyme; cytochrome P450; metabolomics; HPLC-MS

1. Introduction

With the progress of recent genome-sequencing efforts, more than 15,000 cytochromes P450 (P450s) have been registered in the database (<http://drnelson.uthsc.edu/CytochromeP450.html>). Most of these P450s are considered “orphans” because no substrates or functions are associated with these enzymes. A need exists to explore these orphan P450s in order to bridge the wide knowledge gap between amino-acid sequence information and biochemical function.

Metabolomics provides an efficient approach to identify the substrates of these orphan P450s (1–3). Here we describe an untargeted metabolomics approach to identify substrates of orphan P450s from *Streptomyces coelicolor*. Briefly, an organic extract of the organism (or tissue) where the orphan P450 is normally found was prepared as a chemical library that may contain endogenous substrates. The extract was then incubated with the recombinant P450 in the presence of NADPH and redox partners, with subsequent metabolite profiling by LC-MS. The metabolic profile obtained after the reaction was compared with that before the reaction using specialized metabolomics software, e.g., XCMS (4, 5). Molecules that were depleted due to the enzymatic reaction were identified as potential substrates. A successful example of this approach is the recent study of P450 154A1 (6). This method can also be applied in the study of mammalian orphan P450s.

Address correspondence to: Professor F. Peter Guengerich, Department of Biochemistry, Vanderbilt University School of Medicine, 638 Robinson Research Building, 2200 Pierce Avenue, Nashville, Tennessee 37232-0146 U.S.A. Telephone: 1 (615) 322-2261, FAX: 1 (615) 322-4349, f.guengerich@vanderbilt.edu.

2. Materials

2.1 P450 assay

1. Recombinant orphan P450 (see Note 1).
2. Redox partners of P450: 200 μ M spinach ferredoxin and 20 μ M spinach NADPH-ferredoxin reductase. These enzymes can be purchased from Sigma-Aldrich. Rat NADPH-P450 reductase (NPR) can be prepared as described previously (7).
3. 10 mM NADP⁺ stock solution: 382 mg NADP⁺ in 50 ml of Milli-Q water. Store at 4 °C.
4. 100 mM glucose 6-phosphate: 3.4 g in 100 ml of Milli-Q water. Store at 4°C.
5. yeast glucose 6-phosphate dehydrogenase (10³ IU/ml): 1.0 mg in 1 ml of 10 mM Tris-acetate buffer (pH 7.4), containing 20% (v/v) glycerol and 1 mM EDTA. Store at 4 °C.
6. Extraction solvent: ethyl acetate with 0.1% (v/v) acetic acid (see Note 2).
7. Tissue extract: extract tissue with extraction solvent by using a ratio of 200 ml/g for dry tissue or 10/1 (vol/vol) for tissue culture. For preparing mammalian tissue extracts, homogenization before extraction is needed. The organic portion is filtered and dried in a rotary evaporator. Dissolve the extract in minimal acetonitrile so that no obvious pellet is observed. Store the extract at -20 °C.
8. DLPC solution: 5 mg L- α -dilauroyl-*sn*-glycero-3-phosphocholine in 5 ml of Milli-Q water. Sonicate (with probe, 2 $\times$ 10 s, full power). Store at 4°C. Sonicate each day before use and hold at 23 °C until use.
9. 100 mM potassium phosphate buffer, pH 7.4.
10. NADPH generating system: combine 50 parts of 10 mM NADP⁺, 100 parts of 100 mM glucose 6-phosphate, and 2 parts of yeast glucose 6-phosphate dehydrogenase (1 mg/ml). Prepare fresh daily and store on ice when not in use.

2.2 Metabolic profiling by HPLC-MS

1. HPLC: Waters Acquity UPLC system with photodiode array (PDA) detector (see Note 3).
2. Column: Acquity UPLC BEH (C₁₈) column (2 mm $\times$ 100 mm, 1.7 μ m, Waters) (see Note 4).
3. Mobile phase: solvent A: 95% H₂O, 5% CH₃CN, 0.1% HCO₂H (v/v/v); solvent B: 5% H₂O, 95% CH₃CN, 0.1% HCO₂H (v/v/v).

¹ Recombinant P450s are required for this experiment to minimize the metabolic interference conferred by other enzymes. Expression of bacterial P450s in *Escherichia coli* is generally straightforward. On the other hand, expression of mammalian P450s in *E. coli* often requires sequence optimization and may involve other expression hosts (e.g., baculovirus-based insect cells).

² Extraction with ethyl acetate will enrich moderately hydrophobic compounds. To enrich highly hydrophobic compound, e.g., lipids and steroids, solvents such as chloroform or hexane should be used. If the suspected substrate is relatively hydrophilic, a polar solvent (e.g., methanol) should be considered. Avoid using plastic containers or pipettes when preparing tissue extract because plastic polymers may leach plasticizers (especially in organic solvents) and contaminate the prepared extract.

³ UPLC is recommended for its short gradient times, high resolution and highly reproducible retention time. If not available, other analytical HPLC systems can also be used in these experiments. PDA is not required but highly recommended because the additional UV information can be very helpful in the isolation and identification of the substrate.

⁴ An octadecylsilane (C₁₈) reversed-phase column is the most common choice in metabolomics studies of this type. Depending on the extraction method, a C₈ reversed-phase or normal-phase column may be used for the separation of highly hydrophobic compounds.

4. Mass spectrometer: LTQ (Thermo). An ion-trap instrument is preferred in this protocol due to the broader scan range compared with quadrupole-based mass spectrometers. A q-TOF instrument (Waters) is used for high-resolution mass spectrometry (HRMS).

2.3 Processing HPLC-MS dataset using XCMS

Dell Dimension XPS-Gen 5 with Intel Pentium Quad Processor, 4 GB memory, 150 GB hard drive with R language, and XCMS installed (see Note 5).

3. Methods

3.1 Orphan P450 enzyme assay

Skip step 1 if working with bacterial P450s; mix bacterial P450 with ferredoxin and NADPH-ferredoxin reductase in potassium phosphate buffer (pH 7.4) so that the reaction mix (before addition of the NADPH generating system) contains 1 μ M P450, 2 μ M NADPH-ferredoxin reductase and 10 μ M ferredoxin. All reactions should be performed at least in duplicate.

1. Reconstitute the mammalian P450 system: sonicate the DLPC solution for 2×10 -second bursts or until the solution is clear. Add reagents in following sequence: recombinant P450, NPR and DLPC. Gently mix the reaction components for 20–30 s and then leave the mix at room temperature for 10 min. Add an appropriate amount of 100 mM potassium phosphate buffer. A typical reaction has a volume of 1.0 ml and contains 0.1 μ M P450, 0.5–1 μ M NPR and 20 μ g/ml DLPC. Omit the P450 from the negative controls.
2. Add an appropriate amount of tissue extract so that the total organic content is 1% (vol/vol). Initiate the reaction by the addition of 100 μ l of the NADPH-generating system (10% of reaction volume). Incubate the reaction at 37 °C for 1 h.
3. Extract the reaction mix 3 $\times$ with 5 ml of ethyl acetate, combine all organic phases, and evaporate to dryness under a nitrogen stream. Dissolve the extract in 50 μ l of methanol.

3.2 Metabolic profiling by HPLC-MS

Tune and calibrate the mass spectrometer according to the manufacturer's instructions before the experiment. Positive electrospray (+ESI) mode was used in this protocol. Other ionization methods, e.g., atmospheric chemical ionization (APCI), may be employed if the orphan P450s are suspected to metabolize compounds that do not ionize well in the ESI mode, e.g., steroids.

1. HPLC gradient: flow rate 0.35 ml/min, 0–1 min, 15% B (v/v); 1–9 min, 15%–95% (v/v) B; 9–12 min, 95% (v/v) B; 12–13 min, 95%–15% B (v/v); 13–15 min, 15% B (v/v). The column temperature is set at 40 °C.
2. Each sample (10 μ l) of is injected using the partial-loop mode. After each injection, it is recommended to run a blank sample (100% methanol) to ensure there is no carry over of the previous sample.

⁵ For R language, go to <http://www.r-project.org/> and follow the instructions for download and installation. For XCMS, go to website: <http://metlin.scripps.edu/xcms/installation.php> and follow the instruction to download and install XCMS into R.

3.3 Processing HPLC-MS dataset using XCMS

1. Convert the raw LCMS data to CDF or mzXML for processing by XCMS. The manufacturer normally provides a file-converter function in their instrument-operation software. For example, Xcalibur (Thermo) has a “File converter” option under the menu of “Tools,” which can convert .RAW files to .CDF files.
2. Build two file folders under the same directory and name these folders as “Assay” and “NegativeControl” respectively. Move the converted assay data to the Assay folder and the control data to the NegativeControl folder.
3. Start R, load an XCMS library, and run a standard XCMS analysis (see Note 6). Set the output number n to choose how many statistically significant differences should be reported. A good starting point is about 50.
4. Verify the reported metabolic difference by checking the original LCMS data under the extracted-ion mode. Make sure that the difference is significant and consistent among replicates (see Note 7). The metabolites that are consumed in the reaction are considered putative endogenous substrates and worthy of further follow-up. For example, a putative substrate was identified in our study of P450 107U1 (Figure 1a).

3.4 Verification of the identified substrates

Partially purify the substrate candidates. This can be accomplished by collecting the HPLC elute or using solid-phase extraction. Repeat the assay using partially purified candidates to confirm the substrate will be depleted in an enzyme-dependent manner. With fewer interfering compounds present in the partially purified substrate, it is not rare to see the formation of product. In our studies of P450 107U1, a product peak derived from the desaturation of the putative substrate was very obvious (-2 amu, Figure 1b), lending further support that a *bona fide* substrate was identified.

3.5 Structural elucidation of identified substrate

After a substrate is identified, considerable effort may be required to elucidate the structure of the compound. Detailed instructions for structure elucidation for all compounds are beyond the scope of this protocol. Here we describe a simplified and basic systematic approach for structural characterization.

1. Obtain the HR-MS of substrate (error < 5 ppm, using a qTOF, Orbitrap, or magnetic sector instrument) and deduce the molecular formula. Sometimes multiple formulae may be possible. The use of MS fragmentation and other information, e.g., UV absorbance, can often help with selecting the correct molecular formula.
2. Search chemical databases using the derived formula (e.g., chemspider: <http://www.chemspider.com/>, NCBI pubchem: <http://pubchem.ncbi.nlm.nih.gov/>, NIST Chemistry Webbook: <http://webbook.nist.gov/chemistry/>, and Scifinder: <https://scifinder.cas.org/scifinder>); determine whether the substrate formula matches any

⁶ A basic operation manual of XCMS can be downloaded at: <http://metlin.scripps.edu/xcms/faq.php>. Click “LC/MS Preprocessing and Analysis with XCMS” and follow the instruction in the downloaded manual.

⁷ Even with high-quality LCMS data, false metabolic differences may be reported by XCMS. Therefore, it is very important to first check the raw LCMS data to ensure that the reported difference is valid. Second, discard insignificant differences, which may be caused by retention-time drifting, ion-source contamination, or any instrumental variation. A standard decision point is somewhat subjective. In our practice, a substrate candidate will not be selected for follow up unless $>50\%$ of this molecule is consumed during the reaction, under prolonged conditions. If no candidate can meet the criterion, a lower setting is used until about 10 candidates are picked.

known compound. After the integration of all known information, if any known compound appears to be the unidentified substrate then obtain the standard compound (if available commercially) and compare its HPLC elution time and MS fragmentation pattern with those of the unidentified substrate.

3. If the substrate appears to be a novel chemical entity, collect enough material and identify the structure by crystallography or (more realistically) NMR. A minimum amount for 1-dimensional ^1H -NMR work is $\sim 1\ \mu\text{g}$, but an order of magnitude more is required for 2-dimensional ^1H studies and even more is needed if natural abundance ^{13}C -NMR is required.

References

1. Guengerich FP, et al. Approaches to deorphanization of human and microbial cytochrome P450 enzymes. *Biochim Biophys Acta*. 2011; 1814:139–145. [PubMed: 20493973]
2. Guengerich FP, Cheng Q. Orphans in the human cytochrome P450 superfamily: approaches to discovering functions and relevance in pharmacology. *Pharmacol Rev*. 2011; 63:684–699. [PubMed: 21737533]
3. Guengerich FP, et al. Characterizing proteins of unknown function: orphan cytochrome P450 enzymes as a paradigm. *Mol Interv*. 2010; 10:153–163. [PubMed: 20539034]
4. Fraga CG, et al. Signature-discovery approach for sample matching of a nerve-agent precursor using liquid chromatography-mass spectrometry, XCMS, and chemometrics. *Anal Chem*. 2010; 82:4165–4173. [PubMed: 20405949]
5. Smith CA, et al. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem*. 2006; 78:779–787. [PubMed: 16448051]
6. Cheng Q, et al. Cyclization of a cellular dipentaenone by *Streptomyces coelicolor* cytochrome P450 154A1 without oxidation/reduction. *J Am Chem Soc*. 2010; 132:15173–15175. [PubMed: 20979426]
7. Hanna IH, et al. Role of the alanine at position 363 of cytochrome P450 2B2 in influencing the NADPH- and hydroperoxide-supported activities. *Arch Biochem Biophys*. 1998; 350:324–332. [PubMed: 9473308]

Figure 1.