Title: Differential tissue abundance of membrane-bound drug metabolizing enzymes and transporter proteins by global proteomics

Dilip Kumar Singh¹, Deepak Ahire¹, Dmitri R. Davydov², Bhagwat Prasad¹*

¹Department of Pharmaceutical Sciences, Washington State University, Spokane, WA, USA

²Department of Chemistry, Washington State University, Pullman, WA 99164, USA
Running Title: Differential tissue DMET quantification by global proteomics

The total number of pages in the manuscript: 33
The total number of figures: 5
The total number of tables: 1
Total word count of the abstract: 250
Total word count of the introduction: 777
Total word count of the discussion: 1509
Total number of references: 47

Corresponding author: Bhagwat Prasad, Ph.D., Department of Pharmaceutical Sciences, Washington State University, Spokane, WA 99202, USA. Phone: +1-509-358-7739. Fax: +1 509-368-6561. E-mail: bhagwat.prasad@wsu.edu

Keywords: Drug metabolizing enzymes, Transporters, Global proteomics, LC-MS/MS, Targeted proteomics, Total protein approach
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>Ammonium bicarbonate</td>
</tr>
<tr>
<td>BCA</td>
<td>Bicinchoninic acid assay</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>CYPs</td>
<td>Cytochrome P450 enzymes</td>
</tr>
<tr>
<td>DIA</td>
<td>Data-independent acquisition</td>
</tr>
<tr>
<td>DMETs</td>
<td>Drug-metabolizing enzymes and transporters</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>HIM</td>
<td>Human intestinal microsomes</td>
</tr>
<tr>
<td>HKM</td>
<td>Human kidney microsomes</td>
</tr>
<tr>
<td>HLM</td>
<td>Human liver microsomes</td>
</tr>
<tr>
<td>IAA</td>
<td>Iodoacetamide</td>
</tr>
<tr>
<td>IVIVE</td>
<td>In vitro to in vivo extrapolation</td>
</tr>
<tr>
<td>LC-HRMS</td>
<td>Liquid chromatography-high resolution mass spectrometry</td>
</tr>
<tr>
<td>LFQ</td>
<td>Label-free quantification</td>
</tr>
<tr>
<td>MRM</td>
<td>Multiple reaction monitoring</td>
</tr>
<tr>
<td>PBPK</td>
<td>Physiologically based pharmacokinetic models</td>
</tr>
<tr>
<td>SC-TPA</td>
<td>Sequence coverage-informed total protein approach</td>
</tr>
<tr>
<td>SIL</td>
<td>Stable isotope-labeled peptide</td>
</tr>
<tr>
<td>TPA</td>
<td>Total protein approach</td>
</tr>
<tr>
<td>UGTs</td>
<td>Uridine diphosphate-glucuronosyltransferase enzymes</td>
</tr>
</tbody>
</table>
Abstract

Protein abundance data of drug-metabolizing enzymes and transporters (DMETs) are critical for scaling in vitro and animal data to humans for accurate prediction and interpretation of drug clearance and toxicity. Targeted DMET proteomics which relies on synthetic stable isotope-labeled surrogate peptides as calibrators, is routinely used for the quantification of selected proteins; however, the technique is limited to the quantification of a small number of proteins. Although the global proteomics-based total protein approach (TPA) is emerging as a better alternative for large-scale protein quantification, the conventional TPA doesn’t consider differential sequence coverage by identifying unique peptides across proteins. Here, we optimized the TPA approach by correcting protein abundance data by the sequence coverage (SC-TPA), which was applied to quantify 54 DMETs for characterization of i) differential tissue DMET abundance in the human liver, kidney, and intestine, and ii) interindividual variability of DMET proteins in individual intestinal samples (n=13). UGT2B7, MGST1, MGST2, MGST3, CES2, and MRP2 were expressed in all three tissues, whereas, as expected CYP3A4, CYP3A5, CYP2C9, CYP4F2, UGT1A1, UGT2B17, CES1, FMO5, MRP3, and P-gp were present in the liver and intestine. The top three DMET proteins in individual tissues were: CES1>CYP2E1>UGT2B7 (liver), CES2>UGT2B17>CYP3A4 (intestine), and MGST1>UGT1A6>MGST2 (kidney). CYP3A4, CYP3A5, UGT2B17, CES2, and MGST2 showed high interindividual variability in the intestine. These data are relevant for enhancing in vitro to in vivo extrapolation (IVIVE) of drug absorption and disposition and can be used to enhance the accuracy of physiologically based pharmacokinetic (PBPK) prediction of systemic and tissue concentration of drugs.
Significance Statement

We quantified the abundance and compositions of drug-metabolizing enzymes and transporters (DMETs) in pooled human liver, intestine, and kidney microsomes using an optimized sequence coverage-informed total protein approach. The quantification of DMETs revealed quantitative differences in their levels in the liver, intestine, and kidney. Further, the analysis of individual intestine samples confirmed high variability in the levels of CYP3A4, CYP3A5, UGT2B17, CES2, and MGST2. These data are applicable for the prediction of first-pass metabolism and tissue-specific drug clearance.
Introduction

Prediction of tissue-specific drug metabolism and transport is critical for assessing drug clearance and toxicity. In general, the liver is a large organ containing a higher abundance of drug-metabolizing enzymes and transporters (DMETs), thus playing a major role in drug clearance. While DMETs exhibit lower abundance within the intestine, they still significantly impact the bioavailability of high-extraction drugs as a primary organ involved in the first-pass effect. MK-7246, an investigational drug for asthma, is primarily metabolized in the intestine by UGT2B17 as evidenced by the effect of its gene deletion on the first-pass metabolism and bioavailability without any significant effect on the systemic clearance (Y.-H. Wang et al., 2012). Similarly, DMETs in the intestine and kidney may not often play a critical role in determining the clearance and systemic exposure of drugs, but they regulate tissue concentration and toxicity. For example, irinotecan, which is given as an intravenous infusion, is primarily activated to its active form SN38 by carboxylesterase 2 (CES2). Although UGT1A1 can detoxify SN-38 in the intestine and liver, the relative abundance of CES2 versus UGT1A1 is higher in the intestine as compared to the liver, which explains the accumulation and toxicity of SN38 in the intestine (Parvez et al., 2021). Accumulation and toxicity of β-lactam antibiotics and antivirals such as tenofovir and cidofovir in the kidney is mediated by uptake and efflux transporters in the proximal tubular cells (Hagos & Wolff, 2010). Therefore, it is important to quantitatively characterize differential tissue abundance of DMET proteins in the human liver, intestine, and kidney, for mechanistic understanding of drug disposition including tissue-specific toxicity prediction. In addition, while cytochrome P450 enzymes are most abundantly present in the liver, non-cytochrome P450 drug-metabolizing enzymes and transporters are ubiquitously expressed in different tissues. For example, UGT1A10 (Basit et al., 2020), OCT2, OAT1, and OAT3 (Prasad, Johnson, et al., 2016) are exclusively present in extra-hepatic organs, whereas UGT2B17 (Zhang et al., 2018), P-gp, and BCRP (Couto et al., 2020) are more abundant in the intestine as compared to the liver (Prasad, Gaedigk, et al., 2016). Although these data are
available for major DMET proteins in the liver, a comprehensive understanding of DMET abundance in the intestine and kidney is lacking. Moreover, although the intestine is an important first-pass organ that contributes variable drug bioavailability, limited data exist on the interindividual variability of DMET proteins in the intestine.

Liquid chromatography-mass spectrometry (LC-MS) based quantitative proteomics has emerged as a gold standard technique for DMET quantification because of its selectivity, non-dependence on antibodies, multiplexed quantification, high-throughputness, and low cost as compared to the conventional Western blotting (Aebersold et al., 2013; Prasad et al., 2019). Quantitative proteomics strategies are generally based on a bottom-up (shotgun) approach, where proteins are digested into peptides prior to LC separation and MS quantification. LC-MS-based quantification can be performed using either targeted or global (untargeted) approaches (D. Ahire et al., 2022). The targeted proteomics approach utilizes multiple reaction monitoring (MRM) mode to selectively quantify a list of target peptides that are specific to proteins of interest. Although targeted proteomics can be performed using a low-resolution MS instrument (triple quadrupole MS), it requires the selection and purchase of proteotypic (surrogate) peptides (Bhatt & Prasad, 2018; Prasad & Unadkat, 2014). Further, a targeted proteomics assay can only quantify a limited number of proteins (typically <20 proteins) without compromising the sensitivity. To address these challenges, the global proteomics-based total protein approach (TPA) has recently gained significant attention for the quantification of proteins (Bryk & Wiśniewski, 2017; Wiśniewski et al., 2018). TPA is a multiplexed quantitative global proteomics strategy that can quantify proteins without the need for peptide standards and stable isotope labeling (Wiśniewski et al., 2012; Wiśniewski & Rakus, 2014). TPA utilizes MS spectral intensities to simultaneously quantify a large number of proteins (3000-5000) allowing a comprehensive understanding of protein abundance data (He et al., 2019; Jorge et al., 2021; Ölander et al., 2020). However, TPA works on the assumption that the sum of total MS signals...
from all identified proteins in a dataset reflects the total protein content and that the MS signal from a single protein corresponds to its abundance in the studied sample. Therefore, TPA lacks the ability to consider differences in the protein sequence coverage by the unique peptides, thus leading to potentially biased results. Here, we applied a new approach, sequence coverage-informed TPA (SC-TPA), to quantify 54 DMET proteins in pooled microsomal samples isolated from the human liver, intestine, and kidney as well as individual intestinal microsomes. The study revealed quantitative differences in DMET abundance across tissues, which can be leveraged for the development of physiologically-based pharmacokinetic models (PBPK) for predicting systemic and tissue concentration of drugs.

**Materials and Methods**

**Chemicals**

Ammonium bicarbonate (ABC; 98% pure), dithiothreitol (DTT), iodoacetamide (IAA), Pierce trypsin protease (MS-grade), stable isotope-labeled (SIL) peptides, and bovine serum albumin (BSA) were procured from Thermo Fisher Scientific (Rockford, IL). The bicinchoninic acid (BCA) protein assay kit was procured from Pierce Biotechnology (Rockford, IL). Acetone was obtained from Sigma Aldrich (St. Louis, MO). Optima MS-grade acetonitrile, methanol, and formic acid were purchased from Fisher Scientific (Fair Lawn, NJ).

**Procurement of liver, intestinal, and kidney microsomes**

Pooled human liver microsomes (HLM; pool of n=150) and the intestinal microsomal preparations (HIM) from 13 donors (age, 33 to 67 years; mixed sex; race: Caucasian, African American, and Hispanic backgrounds) were procured from BioIVT Inc. (Baltimore, MD). Pooled HIM (pool of n=15) and pooled human kidney microsomes (HKM; pool of n=8) were procured from Xenotech (Kansas City, KS). Demographic information (age and sex) as well as number of donors and individual lot numbers for the pooled HLM, HIM, and HKM are provided in Table S1.
Eleven HLM preparations of different lots from multiple sources (Table S2, Supplementary I) were procured for CYPs quantification by targeted proteomics and their comparison with global proteomics data. Demographic details of all individual HIM donors are provided in Table S3 (Supplementary I). All samples were de-identified by the vendors.

Sample preparation

Microsomal samples were digested by trypsin using a previously described protocol (D. S. Ahire et al., 2021). In brief, an 80 µL sample (1 mg/mL microsomal protein) was mixed with 30 µL ABC buffer (100 mM), 10 µL DTT (250 mM), and 20 µL BSA (0.02 mg/mL). The mixture was incubated at 95 °C for 10 min with a gentle 300 rpm shaking for denaturation and reduction. After cooling to room temperature, the sample was treated with 10 µL IAA (500 mM) in the dark for 30 min. Ice-cold acetone (1 mL) was added to the sample, vortex mixed, and kept at -80 °C for 1 h before a 15 min centrifugation at 16,000 x g and 4 °C. The supernatant was discarded, and the pellet was rinsed with cold methanol (0.5 mL), and centrifuged again under the same conditions for 10 min. The supernatant was removed, and the pellet was dried for 30 min at room temperature. The dried protein was reconstituted in 60 µL of 50 mM ABC buffer and digested with a freshly prepared trypsin solution (20 µL in 50 mM ABC buffer, with a protein:trypsin ratio of 25:1) at 37 °C for 16 h with gentle shaking at 300 rpm. The digestion was quenched by adding 5 µL of 5% formic acid. The sample was vortex-mixed and centrifuged at 10,000 x g for 10 min at 4 °C. The samples were stored in a -80 °C freezer prior to LC-MS analysis.

To assess the reliability of global proteomics-based SC-TPA, eleven digested HLM preparations of different lots (Table S2, Supplementary I) were divided into two aliquots. The first aliquot was diluted with 10 µL of 0.1% formic acid solution for global proteomics, while the second aliquot was diluted with 10 µL of SIL peptides cocktail (Table S4, Supplementary I) for targeted
proteomics. A previously characterized pooled liver tissue membrane preparation with a known abundance of major CYP enzymes was used as a calibrator for targeted proteomics.

**Liquid chromatography and high-resolution mass spectrometry (LC-HRMS) for global proteomics data acquisition**

The global proteomics analysis was conducted on the Easy Spray 1200 series nanoLC coupled with Q-Exactive HF mass spectrometer (Thermo Fisher Scientific, Waltham, MA) using Acclaim PepMap™ trap column (75 µm x 2 cm, 3µ) (Thermo Scientific, part number 164535) connected with analytical PepMap™ RSLC C18 column (75 µm x 250 mm, 2µ) (Thermo Scientific, part number ES902). The mobile phase consisting of 0.1% formic acid in water (A) and 0.1% formic acid in 80% acetonitrile (B), was run in a gradient mode (%B) of 2-6% (0-10 min), 6-30% (10-110 min), 30-100% (110-115 min), and 100% (115-131 min) with a flow rate of 300 nL/min for analysis of all samples except for 11 pooled HLM samples (Table S2, Supplementary I). The injection volume and column temperature were 1 µL and 40 °C, respectively. Data were acquired in the positive ion mode using the data-independent acquisition (DIA) feature of Xcalibur software (Thermo Scientific, Version 4.3.73.11; Table S5, Supplementary I). The pooled 11 HLM samples were analyzed separately using a different gradient program: %B of 2-6% (0-5 min), 6-30% (5-60 min), 30-100% (60-65 min), and 100% (65-80 min). Other LC and MS parameters used for the analysis of individual HIM samples were similar to the above-described conditions for the pooled microsomes.

**Liquid chromatography and tandem mass spectrometry (LC-MS/MS) for targeted proteomics data acquisition**

The 11 pooled HLM samples were also analyzed using targeted proteomics utilizing MRM scan. The analysis was carried out using an M-class Waters UPLC system coupled with Waters Xevo TQ-XS µLC-MS/MS instrument. The peptides were separated on ACQUITY UPLC® HSS T3 column (1.8 µm, 1.0 x100 mm, part number 186003536) connected with VanGuard™ ACQUITY
UPLC® HSS T3 pre-column (1.8 µm, 2.1 x 5 mm, part number 186003976) (Waters, Milford, MA). The mobile phase consisting of 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B) was used with a flow rate of 50 µL/min and a gradient program: %B of 5% (0–3 min), 5–13% (3-8 min), 13-45% (8-20 min), 45–80% (20-21.5 min), 80% (21.5-22.5 min), 5% (22.5-23 min), and 5% (23-28 min). The column temperature and injection volume were 40 °C and 1 µL, respectively. The optimized MRM transitions with cone voltage and collision energy for selected CYPs are provided in Table S4 (Supplementary I).

**Global proteomics data analysis**

The LC-HRMS data acquired in DIA mode were analyzed using MaxQuant (version 2.1.4.0; Max Plank Institute of Biochemistry, Germany) with an inbuilt MaxDIA feature. Proteins were identified by searching against a reference human proteome database ([https://datashare.biochem.mpg.de/s/gle1qKbz2iRuf?path=%2FDiscov](https://datashare.biochem.mpg.de/s/gle1qKbz2iRuf?path=%2FDiscov)eryLibraries). Using MaxQuant, both peptide-spectrum match (PSM) and protein false discovery rate (FDR) were set to 0.01 (i.e., 1%) for decoy database generation. Specific “Trypsin/P” was selected as the proteolytic enzyme for peptide identification with a maximum of “2” missed cleavages. N-terminal acetylation and methionine oxidation were chosen as variable modifications, while the carbamidomethylation of cysteine was selected as a fixed modification, and unique peptides were selected for the label-free quantification (LFQ). The maximum number of modifications per peptide was fixed to “5” for global proteomics data processing. All detected proteins (at least 1 out of 3 replicates) features were assessed for the total protein count in HLM, HIM, and HKM. Venn diagrams were generated for comparison of total proteins as well as DMET proteins in different tissue microsomes using an open-access tool ([http://bioinformatic](http://bioinformatics.psb.ugent.be/webtools/Venn/)s.psb.ugent.be/webtools/Venn/). Proteins that were detected in all 3 replicates with the coefficient of variation (%CV) less than 40 were considered for further analysis.
Since the conventional TPA approach (Wiśniewski, 2017) assumes that all proteins are digested similarly, and all peptides are consistently ionized and detected in LC-HRMS, we utilized SC-TPA (Eq. 1) in this study. We hypothesized that differences in sequence coverage can lead to biased results, i.e., proteins with high and low sequence coverages will be over- or underestimated, respectively. Further, equation 1 assumes that only those proteins that were detected with sequence coverage >10% can be quantified with good confidence. 10% cut-off was selected because proteins with sequence coverage less than 10% showed greater than 30% average CV, whereas proteins with > 10% sequence coverage showed better precision (CV < 30%).

\[
\text{Protein}_{i,SC-TPA} = \frac{\text{Protein intensity}_i}{\text{Total protein intensity (\%SC > 10\%) } \times MW_i} \times \frac{\text{Average \%SC (> 10\%)}}{\%SC \text{ of protein}_i} \quad (Eq. 1)
\]

Both raw intensity (sum of all identified peptide mass intensities for the group) and LFQ intensity (normalized intensity generated by MaxLFQ algorithm) (Cox et al., 2014) have been used for SC-TPA-based quantification.

**Targeted proteomics data analysis**

Targeted proteomics data acquired in the MRM scan mode were processed using Skyline 22.2 (University of Washington, Seattle, WA). The target peptide for each protein was confirmed by aligning the retention time and fragmentation profiles with the corresponding SIL peptide. To account for any technical variability in MS analysis, the light peptide peak area of the target protein was normalized with the corresponding SIL peptide area. BSA response was compared across samples to assess trypsin digestion variability. The absolute protein abundance (pmol/
mg microsomal protein) was calculated by using the previously characterized pooled sample with known CYP abundance as the calibrator. All statistical analysis of the data was performed using GraphPad Prism (v8.4.3; San Diego, CA) and Microsoft Excel (v2209; Redmond, WA).

Results

Conventional TPA versus SC-TPA

While TPA has been extensively used in recent literature (Barber et al., 2022; He et al., 2019; X. Wang et al., 2020), we observed large differences in the protein sequence coverage for DMET proteins (Fig. S1), which suggested that the sum of intensities of individual proteins that were detected with high sequence coverage could be over-estimated using the conventional TPA approach and vice-versa. Indeed, we observed proteins such as Cytochrome b5 (CYB5A) and CYP2E1 with high sequence coverage showing overestimated protein levels as compared to the historical values, whereas proteins such as MGST2 and transporters (e.g., OATP1B1) with low sequence coverage showed underestimation of protein levels (Fig. S2). These results support that SC-TPA is a better method for protein quantification using global proteomics data.

Differential tissue abundance of DMET proteins abundance in human liver, kidney, and intestine by SC-TPA

Using the raw MS intensity data, 4401, 3978, and 4364 proteins were detected in HLM, HIM, and HKM, respectively (Fig. S3), whereas the LFQ intensity data showed 2422, 2384, and 2781 proteins in these three tissues, respectively (Fig. S3). A total of 3004 proteins were common across three tissues based on the raw MS intensity, while 1579 proteins were common with LFQ intensity. Therefore, we utilized DMET abundance data derived from the raw MS intensity (Tables S1-S3, Supplementary II) for further analysis. Differential tissue abundance of 16 CYPs, 11 UGTs, 11 other non-CYPs, and 16 transporters in the liver, intestine, and kidney (Fig. 1) are discussed below.
CYP enzymes: CYPs were primarily expressed in the HLM samples in the following order of abundance (pmol/mg protein) (Fig. 1A): CYP2E1 (86.4) > CYP2C8 (53.4) > CYP2C9 (39.4) ~ CYP3A4 (37.8) > CYP1A2 (35.7) > CYP2A6 (25.6) > CYP3A5 (19.0) ~ CYP2D6 (18.3) > CYP2B6 (16.4) > CYP4F3 (7.8) ~ CYP4A11 (7.7) > CYP4F2 (5.9) > CYP4F11 (5.8) > CYP2C19 (3.6) > CYP2J2 (2.7) > CYP2C18 (0.7). Only six CYPs were expressed in the intestine, however, their levels were significantly lower as compared to the liver, i.e., HLM: HIM abundance ratio of ~14.4 (CYP2C9), 10.4 (CYP4F11), 6.3 (CYP2J2), 3.3 (CYP3A5), 2.4 (CYP3A4), and 1.5 (CYP4F2). Amongst CYPs, only CYP4A11 was detected in the kidney, where its abundance was ~10-fold lower than that observed in HLM.

UGT enzymes: Unlike CYPs that were the most abundant in the liver, UGTs were more ubiquitously expressed in all three tissues (Fig. 1B). The abundance order of UGT1As in HLM (pmol/mg protein) was: UGT1A1 (34.8) > UGT1A4 (32.4) > UGT1A6 (14.3) > UGT1A9 (5.3) > UGT1A3 (1.6). The rank order for UGT2Bs in HLM (pmol/mg protein) was: UGT2B7 (54.2) > UGT2B4 (43.2) > UGT2B15 (6.5) > UGT2B17 (3.8) ~ UGT2B10 (3.6). UGT1A10 was only detected in the intestine and UGT2B17 levels were ~8-fold higher in the intestine than in the liver. The intestinal abundance of UGT1A1 and UGT2B7 was 4 and 6-fold lower than in the liver, respectively. UGT1A6 (33.9), UGT2B7 (25.8), and UGT1A9 (12.9) were abundant in the kidney with comparable levels (within 2.5-fold) as compared to the liver.

Other non-CYP enzymes: The protein abundance of these proteins in HLM (pmol/mg protein) was as follows: CES1 (556.7) > MGST1 (355.1) > MGST2 (53.8) ~ FMO3 (52.6) > CES2 (41.6) > AADAC (28.2) ~ FMO5 (26.8) > MGST3 (5.3) > FMO4 (0.6) (Fig. 1C). Similar to the UGT abundance data, other non-CYP enzymes were also abundant in the extra-hepatic tissues (Fig. 1C). The intestinal abundance of CES2 and MGST3 was ~3.5- and 5.4-fold higher in HIM than in HLM, whereas CES1, MGST1, FMO5, AADAC, and MGST2 were ~235-, 8-, 4-, 1.3, and 1.1-fold lower in HIM than in HLM, respectively. Similarly, MGST3 was 3.3-fold higher in HKM than
in HLM, whereas the abundance of CES2, MGST1, and MGST2 was ~5.1, 2.6, and 2.2-fold lower in HKM than in HLM. FMO1 was uniquely expressed in HKM with abundance values of 13.8 pmol/mg protein, whereas FMO3 and FMO4 were only detected in HLM.

**Transporters:** OATP1B1, OATP1B3, and OAT2 were uniquely detected in HLM, whereas OCT2, OCTN2, OAT1, OAT3, OAT4, MRP4, MATE1, OATP2A1, and OATP4C1 were detected only in HKM (Fig. 1D). Similarly, BCRP expression was only detected in HIM. MRP2 was ubiquitously expressed in all three tissues. P-gp was detected in both HKM and HIM samples. A few low abundant enzymes and transporters such as FMO4, MRP2, OAT2, and OATP1B3 in HLM and MRP3, MRP4, OCT2, OCTN2, OAT3, OATP2A1, and OATP4C1 in HKM and CYP2J2 and MRP2 in HIM showed less than 10% protein sequence coverage, thus resulting in low confidence in their measurement using global proteomics-based approach. Further, although only microsomal samples were analyzed in this study, the global proteomics analysis showed signals for many cytosolic proteins as possible contamination. The differential abundance of all detected cytosolic enzymes is presented in Fig. S4.

**Relative Composition of CYPs and UGTs in Pooled Human Liver, Kidney, and Intestinal Microsomes by SC-TPA**

The relative composition (%) of CYP enzymes in the pooled HLM sample was as follows (Fig. 2A): CYP2E1 (23.6%) > CYP2C8 (14.6%) > CYP2C9 (10.8%) ~ CYP3A4 (10.3%) > CYP1A2 (9.7%) > CYP2A6 (7.0%) > CYP3A5 (5.2%) ~ CYP2D6 (5.0%) ~ CYP2B6 (4.5%) > CYP4F3 (2.1%) ~ CYP4A11 (2.1%) > CYP4F11 (1.6%) ~ CYP4F2 (1.6%). Other CYPs, such as CYP2C18, CYP2C19, and CYP2J2 were below 1%. In HIM, the relative CYP abundance was as follows: CYP3A4 (54.2%) > CYP3A5 (19.6%) > CYP4F2 (13.3%) > CYP2C9 (9.5%) > CYP4F11 (1.9%) ~ CYP2J2 (1.5%). The relative distribution of UGTs in HLM was UGT2B7 (27.1%) > UGT2B4 (21.6%) > UGT1A1 (17.4%) ~ UGT1A4 (16.2%) > UGT1A6 (7.2%) > UGT2B15 (3.2%) > UGT1A9 (2.6%) > UGT2B17 (1.9%) ~ UGT2B10 (1.8%) > UGT1A3 (0.8%).
In HIM, UGT2B17 (53%) showed the highest abundance, followed by UGT1A1 (17.8%), UGT2B7 (16.5%), and UGT1A10 (12.8%). Similarly, the SC-TPA data suggest that UGT composition in the HKM was: UGT1A6 (46.6%), UGT2B7 (35.6%), and UGT1A9 (17.8%).

Microsomal cytochrome b5 (CYB5A) and NADPH cytochrome P450 reductase (POR) are co-enzymes essential for CYP-mediated metabolic reactions. The levels of CYB5A were approximately 1.5-, 2.8-, and 24-fold higher in HLM, HIM, and HKM, respectively, compared to the total CYP content (Fig. 3). Conversely, the levels of POR were approximately 5.3- and 2.3-fold lower than the total CYP content in HLM and HIM, respectively. POR abundance was 2.3-fold higher in HKM over total CYPs (Fig. 3).

**Inter-individual Variability in DMET Abundance in Human Intestine by SC-TPA**

MGST2, UGT2B17, and CYP3A5 showed >100-fold variability (maximum to minimum), and CYP3A4, MGST3, CES2, and UGT2B7 showed 38-, 27-, 20-, and 12-fold variability, respectively (Figure 4 and Table 1), which was non-monotonic in nature. For example, in the case of HIM 2 and HIM 12 (Fig. 4A), higher content of CYP3A4 led to a significant shift from the average CYP composition in these samples. High variability in these proteins is relevant for differential fractional contributions of these enzymes in different individuals.

**Comparison of SC-TPA and targeted proteomics data for representative proteins**

The protein levels of CYP1A2, CYP2A6, CYP2B6, CYP2C8, CYP2C9, CYP2C19, CYP2D6, CYP2E1, CYP3A4, CYP3A5, and CYP4F2 were quantified using both SC-TPA and targeted proteomics methods in 11 different pooled HLM, which revealed that both techniques were comparable with <2.5-fold difference in the measured values except CYP2B6, CYP2C8, and CYP3A5 (Fig. 5A). Further, the CYP abundance data by SC-TPA were compared with the reported mean abundance data (curated for Simcyp software and literature (Achour et al., 2014) (Figs. 5B and 5C; Table S6, Supplementary I). Except for CYP2C19, CYP3A4, and CYP3A5,
the SC-TPA data of all other CYP enzymes fell within 2.5-fold of the reported data. Similarly, the mean abundance of hepatic UGTs measured in this study was comparable to the data curated from Simcyp (Fig. 5D, Table S6, Supplementary I) and a meta-analysis report (Fig. 5E, Table S6, Supplementary I) (Ladumor et al., 2019). UGT1A3, UGT1A9, UGT2B10, and UGT2B15 showed 15.8-, 5.2-, 4.1-, and 6.4-fold difference in the abundance values measured by the SC-TPA method versus the meta-analysis data (Ladumor et al., 2019), whereas UGT1A9 and UGT2B15 showed 5.9- and 6.1-fold difference in the abundance values measured by the SC-TPA versus Simcyp values. The conventional TPA, SC-TPA and MRM methods were compared using a single batch of 11 HLM samples for the quantification of 11CYPs (CYP1A2, CYP2A6, CYP2B6, CYP2C8, CYP2C9, CYP2C19, CYP2D6, CYP2E1, CYP3A4, CYP3A5, and CYP4F2; Fig. S5). Additionally, the SC-TPA results of DMETs quantified in HIM and HKM were compared with previously reported mean abundance data (available in Simcyp software; Fig. S6), which revealed that 75% of DMETs fell within a 5-fold range, except for CYP2J2 in HIM and UGT1A6, UGT1A9, MRP2, and OAT4 in HKM.

**Discussion**

A significant body of literature exists on DMET abundance in the liver using either targeted or global proteomics (Couto et al., 2019; Vildhede et al., 2018; X. Wang et al., 2020). Although these data are useful, the lack of such data on extra-hepatic organs creates a critical knowledge gap, especially for developing mechanistic PBPK models. In this study, we utilized a new optimized global proteomics-based approach, SC-TPA for quantifying a comprehensive list of DMETs in three key human tissues involved in drug disposition, intestine, liver, and kidney. We also studied inter-individual variability in the intestinal abundance of DMETs using individual intestinal donor samples.

We quantified 54 proteins, some of which were characterized for the first time in the kidneys and intestines (e.g., AADAC, MGSTs, and FMOs). In the liver, our results on DMET abundance are
largely consistent with the literature. For example, the top 5 CYP and UGT enzymes in the liver are CYP2E1, CYP2C8, CYP2C9, CYP3A4, CYP1A2, and UGT2B7, UGT2B4, UGT1A1, UGT1A4, UGT1A6, respectively. Interestingly, CYP2C8 levels in this study were unexpectedly higher than CYP2C9 and CYP3A4 as compared to the meta-analysis data (Achour et al., 2014). However, this pattern was consistent with a more recent study using the global proteomics-based TPA method (X. Wang et al., 2020), where CYP2C8 abundance was higher than CYP2C9 and CYP3A4. Collectively, these recent data along with our findings align with greater importance of CYP2C8 in hepatic drug metabolism as supported by clinical drug-drug interaction and pharmacogenomic studies (Backman et al., 2016). CES1, CES2, FMO3, FMO5, AADAC, and MGSTs were the major non-CYP proteins in the liver. The abundance order of CES1 > CES2, FMO3 > FMO5, and MGST1 > MGST2 > MGST3 are consistent with the published data (Boberg et al., 2017; Chen et al., 2016; El-Khateeb et al., 2021; Sato et al., 2012). When compared with the liver, the intestine showed disproportionate expressions of UGT2B17, UGT1A10, MGST3, and CES2, whereas FMO1 was uniquely detected in the kidney. Quantification of POR as compared to the total CYP abundance revealed that the POR to CYPs ratio is higher in the extrahepatic tissues as compared to the liver. This implies that POR could potentially become a rate-limiting factor in drug metabolism in the liver as compared to the intestine and kidney. For example, polymorphisms in the POR (Gong et al., 2017) can affect its catalytic activity for CYP-mediated drug metabolism in the liver as compared to the intestine and kidney.

Based on our data, the significance of individual DMET proteins on drug bioavailability, clearance, and toxicity can be predicted using their unique protein abundance patterns. For example, drugs that are substrates of UGT2B17, UGT1A10, and CES2 will likely experience poor and variable bioavailability. For example, UGT1A10-mediated high first-pass metabolism of its substrates such as raloxifene and troglitazone, is associated with poor bioavailability.
(Mizuma, 2009; Watanabe et al., 2002). Similarly, the high intestinal expression of UGT2B17 is linked to the extensive and variable first-pass metabolism MK-7246, which showed 25-fold higher plasma exposure (area under the curve (AUC)) in the individuals with UGT2B17 gene deletion (*2/*2) as compared with those with the reference allele (*1/*1) (Y.-H. Wang et al., 2012). These data suggest that the quantitative role of UGT1A10 and UGT2B17 variability on the pharmacokinetics of their substrates should be investigated if they are orally administered. The predominant intestinal expression of CES2 is critical for the activation of dabigatran etexilate (double prodrug) to its intermediate metabolite, BIBR0951 (mono-prodrug) in the intestine, which is further converted into dabigatran (BIBR953) by CES1 in the liver (Ishiguro et al., 2014). CES2, P-gp, MRP2, and BCRP in the intestine collectively contribute to the low bioavailability of allisartan, a prodrug of Exp3174 for hypertension treatment in humans (Li et al., 2019). Similarly, if these pathways are important in the activation or inactivation of a toxic drug, the tissue-specific abundance of DMET proteins can lead to local toxicity without affecting plasma levels. For example, organic anion transporters, OAT1 and OAT3, expressed in the kidney lead to accumulation and tissue toxicity of their substrates, e.g., tenofovir (Moss et al., 2014). Differential tissue abundance of non-CYP enzyme abundance impacts the overall metabolic reactions and drug clearance. Phospho-sulindac amide (PSA), a novel potential anti-cancer and anti-inflammatory agent, was sulfoxidized predominantly in the kidney(Xie et al., 2014) because of exclusive expression of FMO1 (Yeung et al., 2000). Differential expression of FMO5 in the liver versus intestine can lead to inter-tissue variability in the metabolism of its substrates such as nabumetone and pentoxifylline. In particular, nabumetone has recently been proposed as a typical substrate of FMO5 and was shown to be efficiently oxidized into 2-(6-methoxynaphthalen-2-yl)ethyl acetate via Baeyer-Villiger oxidation, which is an intermediate precursor of active metabolite 6-methoxy-2-naphthylacetic acid (Matsumoto et al., 2020). FMO5 expression in the intestine (Fig 1C) suggests a strong possibility of the first-pass metabolism of nabumetone, which is consistent with its reported bioavailability (35%) (Davies, 1997).
CYP3A4 and CYP3A5 were the major enzymes contributing to the variability in the total CYP abundance in the intestine. This is likely due to the high inducibility of CYP3A4 by environmental factors affecting the pregnane X receptor (PXR) and the high frequency of genetic polymorphism in CYP3A5. Similarly, POR also shows high variability (>100-fold). As expected UGT2B17 showed the highest variability among the detected UGTs across individual intestinal samples as compared to a moderate variability for UGT2B7 (<20-fold) and UGT1A1 (<10 fold), and UGT1A10 (<5-fold). Similarly, MGST2 and MGST3 showed more than 100-fold and 27-fold variability across individual intestines, respectively. CES2 and FMO5 are more variable than AADAC in the intestine.

Although plasma membrane samples were not available in this study, we were able to detect drug transporters in HLM, HIM, and HKM samples as potential contamination. The abundance patterns of P-gp, MRP2, MRP3, MRP4, OATP1B1, and OATP1B3 were similar to previously reported data (Prasad, Gaedigk, et al., 2016). OAT1, OAT2, OAT4, and OCT2 were expressed only in the kidney, while BCRP was detected only in the intestine (Drozdzik et al., 2014; Prasad, Johnson, et al., 2016). Two important hepatic transporters, OCT1 and BSEP could not be detected in HLM samples, likely due to the selective loss of both transporters during microsomal preparation.

Although quantitative proteomics is reliable and reproducible, the quality of tissue samples could affect the interpretation of inter-individual variability. For instance, individual intestinal samples obtained in this study could be isolated from different regions of the jejunum and ileum resulting in a potential confounding effect of regional variability on inter-individual variability. Nevertheless, the interindividual variability in the marker proteins such as calnexin and Na(+)/K(+) ATPase were within 3.0 fold, suggesting that high biological variability observed in the case of DMET proteins is likely a true phenomenon (Fig. S7).
Further, although we observed a good correlation between SC-TPA versus targeted MRM data for most of the studied proteins (Fig. 5A), when the same samples were analyzed using both the techniques, CYP2B6, CYP2C8, and CYP3A5 showed 3.2-, 5.8-, and 6.7-fold difference, respectively (Fig. S8). Similarly, HIM and HKM data show differences between our results and the previously reported data for 25% DMETs (Fig. S6).

Such discrepancy could have arisen due to the limitations of any of the two methods. Regarding targeted proteomics, factors such as the sub-optimal quality of custom synthesized surrogate peptides used in targeted analysis or variability introduced by a single or small number of peptides with different ionization behavior can lead to biased results (Bhatt & Prasad, 2018). Similarly, the use of a single peptide in quantification can also lead to an underestimation of protein abundance. A similar phenomenon can be observed if the standard peptide faces stability or solubility-related challenges (Wegler et al., 2017). Further, the inclusion of historical semi-quantitative Western blotting data in the meta-analysis could be another reason for the discrepancy in CYP abundance. For HIM and HKM data, the variability in results could have originated from the population variability that was not captured by limited number of donors in the pools. On the other hand, although sequence coverage provided improved results by the SC-TPA method as compared to conventional TPA, the magnitude of ionization of the unique peptides can vary from protein to protein, resulting in over- or under-estimation of protein levels. Similarly, contrary to the literature (D. Ahire et al., 2023), UGT1A6 abundance in the kidney was higher than the abundance of UGT1A9 and UGT2B7. Nevertheless, the relative quantification of cross-tissue abundance or inter-subject variability of individual proteins presented here is accurate and not affected by the methodological limitation. For example, the kidney-to-liver ratio of UGT1A6, UGT1A9, and UGT2B7 was similar (1.8, 2.2, and 0.5, respectively) between the SC-TPA presented here and reported targeted data (D. Ahire et al., 2023).
In summary, we presented a new optimized approach, SC-TPA that leverages protein sequence coverage to correct some limitations of the global proteomics-based protein quantification. The generated DMET abundance data in the human liver, intestine, and kidney can be utilized with historical data for the development of PBPK models. In particular, the data can be used to explain the role of enzymes and transporters in bioavailability and tissue-specific toxicity.
Acknowledgment

We would like to acknowledge Dr. Guihua Yue, Research Associate, Department of Pharmaceutical Sciences, Washington State University (Spokane, WA), for her contribution to LC-MS data acquisition.

Funding

The research reported in this publication was supported in part by the National Institute on Alcohol Abuse and Alcoholism of the National Institutes of Health under Award Number R01AA030155.

Data Availability Statement

The authors declare that all the data supporting the findings of this study are available within the paper and its supplemental data.

Declaration of competing interest

Bhagwat Prasad (Corresponding Author) is cofounder of Precision Quantomics Inc. and recipient of research funding from Bristol-Myers Squibb, Genentech, Gilead, Merck, Novartis, Takeda, AbbVie, Boehringer Ingelheim, and Generation Bio. All other authors declared no competing interests for this work.

Authorship Contributions

Participated in research design: D.K.S; B.P.

Contributed new reagents or analytic tools: D.D.

Conducted experiments: D.K.S.; D.A.

Performed data analysis: D.K.S; B.P.

Wrote or contributed to the writing of the manuscript: D.K.S; D.A.; D.D.; B.P.

References


Mizuma, T. (2009). Intestinal glucuronidation metabolism may have a greater impact on oral bioavailability than hepatic glucuronidation metabolism in humans: A study with raloxifene, substrate for UGT1A1, 1A8, 1A9, and 1A10. *International Journal of Pharmaceutics, 378*(1–2), 140–141. https://doi.org/10.1016/j.ijpharm.2009.05.044


Figure Legends

Fig. 1. Differential protein abundance (pmol/mg protein) of CYPs (A), UGTs (B), other non-CYPs (C), and transporters (D) in pooled HLM, HIM, and HKM by SC-TPA. CYP, cytochrome P450; UGT, UDP-glucuronosyltransferase; CES, carboxylesterase; MGST, microsomal glutathione S-transferase; FMO, flavin-containing monooxygenase; and AADAC, arylacetamide deacetylase; P-gp, P-glycoprotein; MRP, multidrug resistance-associated protein; BCRP, breast cancer resistance protein; OCT2, organic cation transporter 2; OCTN2, organic cation/carnitine transporter 2; OAT, organic anion transporter; MATE1, multidrug and toxin extrusion protein 1; OATP, organic anion transporting polypeptide.

Fig. 2. Relative composition of CYPs (A) and UGTs (B) in pooled HLM (outer circle), HIM (middle circle), and HKM (inner circle) by SC-TPA.

Fig. 3. Abundance of total CYPs versus CYB5A and POR in pooled fractions of HLM, HIM, and HKM by SC-TPA.

Fig. 4. Relative abundance of major CYPs (A) UGTs (B), and other membrane-bound non-CYPs (C) in individual HIMs (N=13) by SC-TPA.

Fig. 5. Correlations of the abundance (pmol/mg liver microsomal protein) of major CYP hepatic enzymes were analyzed using both global (SC-TPA) and targeted (MRM) proteomics approaches across 11 individual HLM lots. The dotted lines depict a 2.5-fold range in association with abundance of CYPs across the 11 HLM lots. Figures B-E show correlations of SC-TPA versus reported values for 12 CYPs (CYP1A2, CYP2A6, CYP2B6, CYP2C8, CYP2C9, CYP2C18, CYP2C19, CYP2D6, CYP2E1, CYP2J2, CYP3A4, and CYP3A5) and 10 UGTs (UGT1A1, UGT1A3, UGT1A4, UGT1A6, UGT1A9, UGT2B4, UGT2B7, UGT2B10, UGT2B15, and UGT2B17).
**Table 1.** Protein abundance data (mean and variability) of major drug-metabolizing enzymes in individual HIM samples (N=13) by SC-TPA.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Mean abundance (pmol/mg of microsomal protein)</th>
<th>SD</th>
<th>Median</th>
<th>Q1</th>
<th>Q3</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYB5A</td>
<td>109.1</td>
<td>45.0</td>
<td>97.4</td>
<td>80.3</td>
<td>143.6</td>
</tr>
<tr>
<td>POR</td>
<td>11.0</td>
<td>8.1</td>
<td>10.0</td>
<td>4.6</td>
<td>20.6</td>
</tr>
<tr>
<td>CYP2C9</td>
<td>1.9</td>
<td>1.5</td>
<td>2.4</td>
<td>0.7</td>
<td>3.0</td>
</tr>
<tr>
<td>CYP3A4</td>
<td>20.3</td>
<td>24.0</td>
<td>12.4</td>
<td>5.7</td>
<td>20.6</td>
</tr>
<tr>
<td>CYP3A5</td>
<td>7.0</td>
<td>7.9</td>
<td>2.7</td>
<td>0.0</td>
<td>14.5</td>
</tr>
<tr>
<td>CYP3A4</td>
<td>1.6</td>
<td>1.6</td>
<td>1.2</td>
<td>0.4</td>
<td>2.6</td>
</tr>
<tr>
<td>CYP4F11</td>
<td>0.2</td>
<td>0.6</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td>CYP2J2</td>
<td>0.6</td>
<td>1.1</td>
<td>0.0</td>
<td>0.0</td>
<td>1.1</td>
</tr>
<tr>
<td>UGT1A1</td>
<td>7.1</td>
<td>3.5</td>
<td>6.1</td>
<td>4.2</td>
<td>10.6</td>
</tr>
<tr>
<td>UGT1A4</td>
<td>0.6</td>
<td>1.3</td>
<td>0.1</td>
<td>0.0</td>
<td>0.6</td>
</tr>
<tr>
<td>UGT1A6</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td>UGT1A10</td>
<td>5.7</td>
<td>2.0</td>
<td>5.8</td>
<td>4.2</td>
<td>6.8</td>
</tr>
<tr>
<td>UGT2B7</td>
<td>7.1</td>
<td>5.5</td>
<td>5.4</td>
<td>3.3</td>
<td>9.4</td>
</tr>
<tr>
<td>UGT2B17</td>
<td>18.0</td>
<td>8.0</td>
<td>17.9</td>
<td>13.6</td>
<td>21.9</td>
</tr>
<tr>
<td>AADAC</td>
<td>14.6</td>
<td>6.4</td>
<td>14.5</td>
<td>9.7</td>
<td>17.1</td>
</tr>
<tr>
<td>CES1</td>
<td>1.4</td>
<td>2.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2.5</td>
</tr>
<tr>
<td>CES2</td>
<td>66.8</td>
<td>57.9</td>
<td>46.1</td>
<td>15.2</td>
<td>103.7</td>
</tr>
<tr>
<td>CES3</td>
<td>0.4</td>
<td>1.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>FMO5</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
<td>1.7</td>
<td>6.7</td>
</tr>
<tr>
<td>MGST1</td>
<td>20.8</td>
<td>11.7</td>
<td>19.0</td>
<td>11.4</td>
<td>25.3</td>
</tr>
<tr>
<td></td>
<td>MGST2</td>
<td></td>
<td>23.2</td>
<td>58.2</td>
<td>34.7</td>
</tr>
<tr>
<td>----</td>
<td>--------</td>
<td>---</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td>51.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MGST3</td>
<td></td>
<td>12.1</td>
<td>20.5</td>
<td>8.9</td>
</tr>
<tr>
<td></td>
<td>19.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>