UPLC/MS/MS Analysis of Human Plasma and Serum Samples using the Biocrates Absolute IDQ p180 Kit

Duke University School of Medicine, Proteomics and Metabolomics Shared Resource

Objective: Measure the levels of selected metabolites in human plasma and serum samples using the Biocrates Absolute IDQ p180 kit.

Introduction:

The Absolute IDQ p180 assay quantifies over 180 metabolites from five analyte groups: acylcarnitines, amino acids, biogenic amines, glycerophospholipids, and sphingolipids. The p180 kit includes all requisite calibration standards, internal standards, and QC samples. The use of these standards according to the detailed analysis protocol, which was validated in Biocrates' lab in Austria, assures assay harmonization and standardization within a project, across projects, and across laboratories. Selective analyte detection is accomplished by use of a triple quadrupole tandem mass spectrometer operated in Multiple Reaction Monitoring (MRM) mode in which specific precursor to product ion transitions are measured for every analyte and stable isotope labeled internal standard. There are two separate tandem mass spectrometric analyses of each sample. For the analysis of acylcarnitines, glycerophospholipids, and sphingolipids, samples are introduced using a Flow Injection Analysis method (FIA-MS/MS, Figure 1A). Sample analyses of amino acids and biogenic amines are performed by a UPLC (ultra-high pressure liquid chromatography) tandem MS method using a reversed phase analytical column for analyte separation (LC-MS/MS, Figure 1B).

Figure 1. Schematic of data collection methodologies for Biocrates p180 kit, including Flow-Injection MS/MS (A) and LC-MS/MS (B).
The calibration standards provided in the Biocrates Absolute IDQ kit were used for quantitation. A single point calibration standard is used for the reproducible quantitation of the acylcarnitines, glycerophospholipids, and sphingolipids. Therefore, the results for these analytes are described as semi-quantitative (schematic shown in Figure 2A). Seven calibration standards are used for highly accurate and reproducible quantitation of the amino acids and biogenic amines as shown in Figure 2B. Calibration standards were fit with a linear regression using $1/x^2$ weighting. Figure 2 shows a schematic and representative examples of how the single-point calibrator or calibration curve would be used to back-calculate QC and sample concentrations for the different analyte classes.

**Figure 2.** Schematic depicting the quantitative methodologies used in the Biocrates p180 kit for flow-injection analysis (A) and LC-MS/MS analysis (B).

The samples were prepared in a 96-well plate format using the layout shown in Figure 3.

**Figure 3.** Schematic depicting the 96-well plate layout for the analysis of study samples including: blanks, calibration standards, and QC samples from Biocrates. Three additional QC samples were
analyzed: the DPMCF Global Reference QC, the study sample pool QC, and the NIST SRM-1950 reference plasma.

Sample Preparation:

Samples were prepared using the Absolute IDQ® p180 kit (Biocrates Innsbruck, Austria) in strict accordance with their detailed protocol. After the addition of 10 µL of the supplied internal standard solution to each well of the 96-well extraction plate, 10 µL of each study sample was added to the appropriate wells. The plate was then dried under a gentle stream of nitrogen. The samples were derivatized with phenyl isothiocyanate, then eluted with 5mM ammonium acetate in methanol. Samples were diluted with either water for the UPLC analysis (4:1) or running solvent (a proprietary mixture provided by Biocrates) for flow injection analysis (20:1).

A pool of equal volumes of all samples analyzed on the first plate was created (SPQC). The pooled sample was prepared and analyzed in the same way as the study samples on all plates. From each plate this sample was injected once before, once during, and once after the study samples in order to measure the performance of the assay across the sample cohort. The analyses of this pool can be used to assess potential batch effects. The order of injection of the samples is shown in Figure 4.

Sample Injection Sequence

- double blank
- PBS ‘zero’ samples
- Calibration curve Low to High
- Low, Mid, High QC Samples
- Global Reference QC
- SPQC
- NIST SRM 1950
- Study Samples 1 to 38
- Low, Mid, High QC Samples
- Global Reference QC
- SPQC
- NIST SRM 1950
- Study Samples 39 to 76
- Global Reference QC
- SPQC
- NIST SRM 1950
- High, Mid, Low QC Samples
- Calibration Curve Low to High

Figure 4. Schematic depicting the injection order of the samples for UPLC analysis. For Flow Injection Analysis the injection order is the same except the calibration curves are not analyzed. Note that the Global Reference QC was prepared once and analyzed three times giving a measure of the analytical variability of the MS/MS analyses. The SPQC and NIST samples were prepared three times and each analyzed once giving a measure of total analytical variability: sample preparation variability and UPLC/MS/MS variability.

Sample Analysis:

UPLC separation of amino acids and biogenic amines was performed using a Waters (Milford, MA) Acquity UPLC with a Waters Acquity 2.1 mm x 50 mm 1.7 µm BEH C18 column fitted with a Waters Acquity BEH C18 1.7 µm Vanguard guard column. Analytes were separated using a gradient from 0.2% formic acid in water, to 0.2% formic acid in acetonitrile. Total UPLC analysis time was approximately 7 minutes per sample. Acylcarnitines, sphingolipids, and glycerophospholipids were analyzed by flow injection analysis (FIA) with total analysis time of approximately 3 minutes per sample. Using electrospray ionization in positive mode, samples for both UPLC and flow injection analysis were introduced directly into a Xevo TQ-S triple quadrupole mass spectrometer (Waters) operating in the Multiple Reaction Monitoring (MRM) mode. MRM transitions (compound-specific precursor to product ion transitions) for each analyte and internal standard were collected over the appropriate retention time. The UPLC-MS/MS data were imported into Waters application TargetLynx™ for peak integration, calibration, and concentration calculations. The UPLC-MS/MS data from TargetLynx™ and FIA-MS/MS data were analyzed using Biocrates MetIDQ™ software.
Data Return Document Descriptions

A number of documents have been added to the Express Data Repository including this sample analysis summary. A link to the repository is included here. A description of the data spreadsheets follows the link.  https://discovery.genome.duke.edu/express/resources/project/

UPLC p180 Data.xlsx

This is an Excel workbook containing three worksheets. The first worksheet contains the calculated concentration data (µM) acquired in the study for analyte classes Amino Acids and Biogenic Amines. The second row in this worksheet lists the analytes measured and the status for each one. Row 4 lists the lower limit of detection (LOD) for the analytes. The Biocrates-defined lowest calibration standard and highest calibration standard are listed in Rows 5 and 6. The barcode number of the plate on is listed in column A. The Sample Bar Code number (column B) is assigned to every sample by the Biocrates MetIDQ™ software. The Sample Identification Number in column D is the unique sample identifier assigned by the Proteomics and Metabolomics Sample Submission System. Column G lists the Customer Sample Identification information which were listed on every sample tube provided for analysis. Table 1 below gives the unique plate barcodes for the LC-MS/MS and Flow-Injection Analysis-MS (FIA-MS) analysis.

Table 1. Barcodes for Plate Analysis

<table>
<thead>
<tr>
<th>Plate Barcode for UPLC-MS/MS Analysis</th>
<th>Plate Barcode for FIA-MS Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1010101011-1</td>
<td>1010101012-1</td>
</tr>
<tr>
<td>1010101013-1</td>
<td>1010101014-1</td>
</tr>
<tr>
<td>1010101015-1</td>
<td>1010101016-1</td>
</tr>
</tbody>
</table>

The results for the SPQC sample are in rows 167 through 175. Row 177 has the %CV for the SPQC sample results. The average %CV for the UPLC data for SPQC is 9.4%, with a range from 2.9 to 64.2%. It is our recommendation that if the inter-day variability of the Study Pool QC sample (as measured by the %CV) for a particular analyte is greater than 25%, then the data for that analyte in the study samples should be flagged for removal from the dataset because of imprecision. Removing analyte histamine, whose SPQC variability was 64.2%, gives an average CV of 7.8% for the remaining analytes. Additionally, best practice suggests that analytes which have more than 40% missing values should be flagged for removal from the dataset because of excess missing values. This last point may not be the case in a test-control study, where either test or control subjects/samples may be below the LOD. In this case, replacement of missing values with LOD/2 is recommended.

In addition, some scaling to control for batch effects may be needed during statistical analysis; further investigation of batch effect is discussed below under Principal Components Analysis.

In the data document UPLC p180 Data.xlsx the concentration data (µM) are coded as shown below in Table 2 in order to allow presentation of additional information regarding data quality:
Table 2. Key to the Analyte Status Columns Used in the Data Tables

<table>
<thead>
<tr>
<th>Status</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td>Calculated concentrations are based on a standard curve of the analyte listed. (Not all analytes are contained in the calibration standards provided by Biocrates. Therefore, results for some analytes will be coded as Semi-Quantitative.)</td>
</tr>
<tr>
<td>&lt;Lowest CS: Lowest Calibration Standard&gt;Value&gt;LOD</td>
<td>The value is greater than the LOD but less than the Biocrates-defined lowest calibration standard. These values should be considered reliable except for analytes for which the %CV observed for the pooled sample was greater than 25%.</td>
</tr>
<tr>
<td>&gt;Highest CS</td>
<td>Greater than the highest calibration standard</td>
</tr>
<tr>
<td>&lt;LOD</td>
<td>The internal standard peak area for this analyte was outside of the normal range. (Not all analytes have internal standards.)</td>
</tr>
<tr>
<td>Internal Standard out of range</td>
<td>The internal standard peak area for this analyte was outside of the normal range. (Not all analytes have internal standards.)</td>
</tr>
<tr>
<td>Semi-Quantitative</td>
<td>Calculated concentration is not based on a calibration curve but on the peak area of an internal standard or that of a structurally similar analyte. This data is valid but is semi-quantitative instead of fully quantitative.</td>
</tr>
<tr>
<td>No Interception or NA</td>
<td>No peak was detected in the chromatogram at the appropriate retention time for this analyte. For statistical analyses apply the same value used for &lt;LOD.</td>
</tr>
</tbody>
</table>

The second worksheet in the workbook contains the study sample data with the Status columns removed.

The third worksheet in the workbook contains a statistical summarization of the data from the Biocrates MetIDQ™ application. The Analyte Statistics worksheet lists these summary statistics (Min µM, Max µM, Mean µM, Median µM, 25th Percentile µM, 75th Percentile µM, STD µM, MAD µM, Skewness, Kurtosis, CV [%], CVRobust [%]) by metabolite for all of the samples. The number of results (e.g. samples with usable data) included for each statistical calculation is indicated by “n” in Column C. These results are compiled for the entirety of the dataset, in essence showing the biological variance in the analytes measured. Most analytes show low biological variability, for instance 26 amino acids and biogenic amines have biological variance of less than 40%. Some have exceedingly low biological variability, including Glutamine (Gln, 12.6%) and SDMA (17.5%).

**UPLC QC and NIST Data.xlsx**

This workbook contains two worksheets. The first worksheet contains the Biocrates plasma QC concentration data for three levels of QCs for analyte classes amino acids and biogenic amines. The measurements for the QC samples are provided in order to confirm that quantification of the metabolites across the wide dynamic range performed in this analysis is generally accurate and reproducible. The second worksheet lists the results for the NIST SRM 1950 samples. Three replicates of this sample were prepared on every sample plate. One replicate was injected before the study samples, one during, and one after the study samples as an additional measure of assay performance. QC values showed excellent reproducibility (Table 3).
Table 3. Reproducibility of UPLC Biocrates QC Sample Analyses

<table>
<thead>
<tr>
<th>QC Level</th>
<th>Average % CV</th>
<th>Min %CV</th>
<th>Max %CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>13.8</td>
<td>8.6</td>
<td>28.4</td>
</tr>
<tr>
<td>Medium</td>
<td>9.6</td>
<td>5.0</td>
<td>15.9</td>
</tr>
<tr>
<td>High</td>
<td>6.8</td>
<td>2.5</td>
<td>10.9</td>
</tr>
</tbody>
</table>

FIA p180 Data.xlsx

This is an Excel workbook containing two worksheets. The first worksheet contains the calculated concentration data (µM) acquired in the study for analyte classes Glycerophospholipids, Sphingolipids, and Acylcarnitines. The second row in this worksheet lists the analytes measured and the status for each one. Rows 3 through 6 list the lower limit of detection (LOD) for the analyte in each analytical run. The Biocrates-defined lowest calibration standard and highest calibration standard are listed in Rows 7 and 8 if applicable. The barcode number of the plate on which each sample was analyzed is listed in column A. The Sample Bar Code number (column B) is assigned to every sample by the Biocrates MetIDQ software. The Sample Identification Number in column D is the unique sample identifier assigned by the Proteomics and Metabolomics Sample Submission System. Column G lists The Customer Sample Identification information which were listed on every sample tube provided for analysis. Table 1 above gives the unique plate barcode for the LC-MS/MS and Flow-Injection Analysis-MS (FIA-MS) plates for each sample box number. The results for the SPQC sample are in rows 169 through 177. Row 179 has the %CV for the SPQC sample results. The average %CV for the FIA data for SPQC is 14.3%, with a range from 3.5 to 52.8%. It is our recommendation that if the inter-day variability of the Study Pool QC sample (as measured by the %CV) for a particular analyte is greater than 25%, then the data for that analyte in the study samples should be flagged for removal from the dataset because of imprecision. If the two outliers are removed (PCaa C36:0 (49.9%CV) and PCaa C30:1 (52.0%CV), the remaining analytes give an average CV of 13.7%. Additionally, best practice suggests that analytes which have more than 40% missing values should be flagged for removal from the dataset because of excess missing values. The NIST Samples also showed excellent reproducibility, with an average % CV of 8.65%.

In the data document FIA p180 Data.xlsx the concentration data (µM) are coded as shown above in Table 2 in order to allow presentation of additional information regarding data quality.

The second worksheet in the workbook contains the study sample data with the Status columns removed.

The third worksheet in the workbook contains a statistical summarization of the data from the Biocrates MetIDQ™ application. The Analyte Statistics worksheet lists these summary statistics (Min µM, Max µM, Mean µM, Median µM, 25th Percentile µM, 75th Percentile µM, STD µM, MAD µM, Skewness, Kurtosis, CV [%], CVRobust [%]) by metabolite for all of the samples. The number of results (e.g. samples with usable data) included for each statistical calculation is indicated by “n” in Column C.

FIA QC Data.xlsx

This workbook contains two worksheets. The first worksheet contains the Biocrates plasma QC concentration data for three levels of QC for analyte classes acylcarnitines, glycerophospholipids, and sphingolipids. QC concentrations are not provided for all the analytes detected, only for the subset of analytes within each class which have reference values available from Biocrates. The measurements for the QC samples are provided in order to confirm that quantification of the metabolites across the wide dynamic range performed in this analysis is accurate and reproducible. The second worksheet lists the results for the NIST SRM 1950 samples. Two replicates of this sample were prepared on every sample plate. One replicate was injected before the study samples and one after the study samples as an additional measure of assay performance. These Biocrates QC values showed excellent reproducibility (Table 4). The NIST Samples also showed excellent reproducibility with an average % CV of 8.94%.

Table 4. Reproducibility of FIA Biocrates QC Sample Analyses
### QC Level

<table>
<thead>
<tr>
<th></th>
<th>Average % CV</th>
<th>Min %CV</th>
<th>Max %CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>16.7</td>
<td>9.0</td>
<td>20.7</td>
</tr>
<tr>
<td>Medium</td>
<td>9.7</td>
<td>6.7</td>
<td>17.2</td>
</tr>
<tr>
<td>High</td>
<td>8.7</td>
<td>2.4</td>
<td>20.2</td>
</tr>
</tbody>
</table>

### Supplemental Data Archive

An additional Zip archive has been added to the Express Data Repository for this project. The archive is named ADNI_p180_Format.zip. The following data files are contained in this folder:

**UPLC180DATA.csv**
This csv table was created from UPLC p180 Data.xlsx. The Customer Sample Identification number was moved from Column G to Column A. The analyte status columns were deleted. The row between the study sample data and the study pool (SPQC) data was deleted. The %CV calculations for the study pool samples was deleted. All “<LOD” results were replaced with “NA”.

**P180UPLCLODvalues.csv**
This csv table contains the Lower Limit of Detection values for the amino acids and biogenic amines for each plate analyzed in this study by Plate Barcode.

**FIA180DATA.csv**
This csv table was created from FIA p180 Data.xlsx. The Customer Sample Identification number was moved from Column G to Column A. The analyte status columns were deleted. The row between the sample data and the study sample data and the study pool (SPQC) data was deleted. The %CV calculations for the study pool samples was deleted. All “<LOD” results were replaced with “NA”.

**FIALODvalues.csv**
This csv table contains the Lower Limit of Detection values for the amino acids and biogenic amines for each plate analyzed in this study by Plate Barcode.

### Principal Components Analysis (PCA) Plots

In order to assess general variability for the samples within each analyte class, to assess batch effect, and to look for sample outliers; a Principal Components Analysis (PCA) was performed for amino acids, biogenic amines, acylcarnitines, glycerophospholipids, and sphingolipids using JMP® Pro v12.0 software (SAS, Cary, NC) and the data from tables UPLC p180 Data.xlsx and FIA p180 Data.xlsx. Analytes with measurable concentrations for more than half of the samples were included in the analysis. For analytes for which there were missing values, missing values were replaced with the Biocrates LOD value for that analyte and the restricted maximum likelihood (REML) method was used for correlation.

Figures below show PCA plots for each analyte class. Separate colors were assigned to samples (red) and SPQC (blue) in order to evaluate variability and to check for outliers.

The results for the pools are generally clustered together centrally within the samples which implies that there is greater biological than technical variability.
Principal Components Analysis of Amino Acids
Principal Components Analysis of Biogenic Amines

<table>
<thead>
<tr>
<th>Eigenvalue</th>
<th>20</th>
<th>40</th>
<th>60</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.3902</td>
<td>2.0178</td>
<td>1.8084</td>
<td>1.4125</td>
</tr>
<tr>
<td></td>
<td>1.1677</td>
<td>1.0810</td>
<td>0.9118</td>
<td>0.7113</td>
</tr>
<tr>
<td></td>
<td>0.6387</td>
<td>0.6209</td>
<td>0.5124</td>
<td>0.3947</td>
</tr>
</tbody>
</table>
Principal Components Analysis of Acylcarnitines

Eigenvalue

<table>
<thead>
<tr>
<th>Eigenvalue</th>
<th>20</th>
<th>40</th>
<th>60</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.2950</td>
<td>7.5188</td>
<td>2.1475</td>
<td>2.0480</td>
<td></td>
</tr>
<tr>
<td>1.3991</td>
<td>0.9918</td>
<td>0.8939</td>
<td>0.7449</td>
<td></td>
</tr>
<tr>
<td>0.6455</td>
<td>0.5963</td>
<td>0.5514</td>
<td>0.4878</td>
<td></td>
</tr>
</tbody>
</table>

Sample Type
- Sample
- SPQC

Component 1 (38 %)

Component 2 (21.5 %)
Principal Components Analysis of Glycerophospholipids
Principal Components Analysis of Sphingolipids

![PCA plot showing the distribution of sphingolipid samples across principal components 1 and 2. The figure includes a table of eigenvalues and a wheel of sample types, including SM C24:0, SM (OH) C24:1, SM C26:1, SM C18:0, and SM C18:1.](image_url)