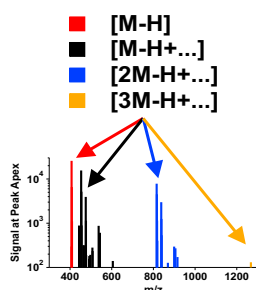


## RESEARCH ARTICLE

# Ion-neutral Clustering of Bile Acids in Electrospray Ionization Across UPLC Flow Regimes

Patrick Brophy,<sup>1</sup> Corey D. Broeckling,<sup>1</sup> James Murphy,<sup>2</sup> Jessica E. Prenni<sup>1</sup><sup>1</sup>Proteomics and Metabolomics Facility, Colorado State University, Fort Collins, CO 80523, USA<sup>2</sup>Waters Corporation, Milford, MA 01757, USA

**Abstract.** Bile acid authentic standards were used as model compounds to quantitatively evaluate complex in-source phenomenon on a UPLC-ESI-TOF-MS operated in the negative mode. Three different diameter columns and a ceramic-based microfluidic separation device were utilized, allowing for detailed descriptions of bile acid behavior across a wide range of flow regimes and instantaneous concentrations. A custom processing algorithm based on correlation analysis was developed to group together all ion signals arising from a single compound; these grouped signals produce verified compound spectra for each bile acid at each on-column mass loading. Significant adduction was observed for all bile acids investigated under all flow regimes and across a wide range of bile acid concentrations. The distribution of

bile acid containing clusters was found to depend on the specific bile acid species, solvent flow rate, and bile acid concentration. Relative abundancies of each cluster changed non-linearly with concentration. It was found that summing all MS level (low collisional energy) ions and ion-neutral adducts arising from a single compound improves linearity across the concentration range (0.125–5 ng on column) and increases the sensitivity of MS level quantification. The behavior of each cluster roughly follows simple equilibrium processes consistent with our understanding of electrospray ionization mechanisms and ion transport processes occurring in atmospheric pressure interfaces.

**Keywords:** Electrospray, Ion-neutral clustering, Adduction, Liquid chromatography, Time-of-flight, Mass spectrometry, Bile acids, Quantification

Received: 6 October 2017/Revised: 18 December 2017/Accepted: 19 December 2017/Published Online: 9 February 2018

## Introduction

Discussions of ion-neutral clusters in the electrospray ionization (ESI) mass spectrometry literature are limited despite being a ubiquitous feature of the mass spectrum obtained from these techniques [1–9]. Singly and multiply charged ion-neutral clusters consisting of at least one ion and one or more neutral species are formed during the electrospray process [7–9]. Neither Dole's charge residue model (CRM) [10] nor the Iribarne-Thomson ion evaporation model (IEM) [11–14] alone completely explain the presence of the various ions and ion-neutral clusters observed by Gamero-Castaño and de la Mora in their study of electrospray produced salt clusters using

a high-resolution differential mobility analyzer [7]. This work suggests a mixed set of ESI processes leading to gas-phase ions and ion-neutral clusters that do not neatly fit into a pure-CRM or pure-IEM. An earlier investigation of cluster ion behavior using an ESI quadrupole mass spectrometer at concentrations leading to ESI saturation also claims that these species can be rationalized under the CRM, IEM, and by additional processes occurring in the atmospheric pressure interface [8]. Other investigations by Meng and Fenn propose an entirely different theory where clusters are formed during the electrospray process due to the precipitation of solutes as solvent evaporates [9]; they reject the hypotheses that gas-phase reactions could produce ion-neutral clusters in the ESI source and also reject the notion that adiabatic expansion during transfer to the vacuum system would result in cluster condensation due to their instrument's design.

The understanding of ion-neutral clusters is further obscured when considering the processes occurring during the transit of ions from the ESI source through the atmospheric pressure

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s13361-017-1878-6>) contains supplementary material, which is available to authorized users.

Correspondence to: Corey Broeckling; e-mail: [cbroeckl@colostate.edu](mailto:cbroeckl@colostate.edu)

interface (API). APIs transport ions and bulk gas from an atmospheric pressure ion source to low pressure vacuum chambers during which adiabatic expansion and rapid cooling occur, resulting in cluster condensation [15–17]. API design for coupling ESI sources to mass analyzers has significantly improved over the years by preventing neutral species from entering the mass spectrometer [18]. This has decreased the extent to which ions are observed as solvated ion-neutral clusters, but some buffer additives (e.g.,  $\text{Na}^+$ ) and contaminants are present in high abundance and can form very stable clusters. Importantly, the inherent design and operation of APIs biases the mass spectrum towards observing ion-neutral clusters and fragments depending on the strength of applied electric fields. These processes have been addressed but are underappreciated and remain poorly understood [19–22]. Recent comparisons between clustering in ion trap, triple quadrupole, and quadrupole-Orbitrap mass spectrometers highlight the significant influence of API design on the observed cluster distribution in the mass spectrum [2]. This comparative study demonstrates the influence of buffer gas and contaminants in buffer gases on the observed ion distribution.

These complexities and the pervasiveness of ion-neutral clusters raise the question as to why ion-neutral clusters have received limited attention. There exists a widely adopted canonical model of ESI that assumes ions observed in the mass spectrum should adhere to the forms of  $[\text{M}+\text{H}]^+$  and  $[\text{M}-\text{H}]^-$  in the positive and negative modes, respectively. Generally speaking, these protonated and deprotonated molecules provide sufficient sensitivity for many analytical tasks, are abundantly formed by a majority of compounds, and have been the historical focus of targeted assays on triple-quadrupole systems due to their interpretability in the mass spectrum and the quality of their fragmentation spectra. However, complexity of the mass spectrum (and particularly the high-resolution mass spectrum) is extremely complicated and multiplexed. Numerous software packages have been developed to help deal with these complexities by grouping or annotating all the signals arising from a single molecule using a variety of approaches [23–25]. In non-targeted profiling, recognizing the presence of these adducts in the mass spectrum is of great importance for correctly interpreting mass spectral features.

Here, we quantitatively examine the behavior of all the ions arising from a single molecule across various column diameters in order to investigate the influence of solvent flow rate on the observed mass spectrum. We chose to focus on bile acids in the negative mode because previous work in our laboratory identified these molecules as being prone to forming ion-neutral clusters. Additionally, bile acids are complex, small molecules representative of real world metabolomics studies [26, 27]. Many previous investigations focusing on ion-neutral clustering rely on infusion based experiments and unit-mass resolution quadrupole mass analyzers. The use of reverse phase ultra-performance liquid chromatography (UPLC), high-resolution mass analyzers, and programmatic data analysis aims to extend the early work of this field towards real-world applications where the complexity of these processes can be embraced,

explored, and the implications understood within the scope of modern analytical tasks.

## Experimental/Methods

### *Preparation of Standards*

Bile acid standards were obtained from the Human Metabolome Library through the Human Metabolome Database [28]. All solvents were LCMS grade or better. Bile acid standards were prepared from 1 mg  $\text{mL}^{-1}$  methanol stock solutions as both single component standards and calibration standard mixes by diluting the methanol stock in a 70:30 v/v solution of LCMS grade water (Fisher Scientific, Hampton, NH, USA) and LCMS grade acetonitrile (Fisher Scientific, Hampton, NH, USA) with 0.1% formic acid by volume (Thermo Fisher Scientific Pierce, Rockford, IL, USA). Single component standards were diluted to a final concentration of 1  $\mu\text{g mL}^{-1}$ . The seven calibration standard mixes span the concentration range from 125  $\text{ng mL}^{-1}$  to 10  $\mu\text{g mL}^{-1}$  with each individual component present at the same concentration within a single standard mix. The prepared standards were stored at  $-80^\circ\text{C}$ .

### *UPLC Conditions*

A Waters Acquity UPLC system was used for reverse phase liquid chromatography with either a 2.1 mm  $\times$  50 mm or 1.0 mm  $\times$  50 mm HSS T3 Acquity column (Waters Corporation) with 1.8  $\mu\text{m}$  diameter particles. A Waters Acquity UPLC M-Class was used for micro-flow reverse phase liquid chromatography with a 300  $\mu\text{m}$   $\times$  50 mm HSS T3 M-Class capillary column (Waters Corporation) with 1.8  $\mu\text{m}$  diameter particles. The M-Class UPLC was also used with a custom passivated iKey Separation Device (Waters Corporation) with a 150  $\mu\text{m}$   $\times$  50 mm inner diameter channel packed with 1.8  $\mu\text{m}$  diameter HSS T3 particles. All columns were thermostatically controlled to 45  $^\circ\text{C}$  using column heaters onboard the UPLC stack or within the iKey device. Fluidic connections and transfer lines were installed according to manufacturer recommendations.

Both UPLC systems were operated with a 5  $\mu\text{L}$  injection loop in the partial loop injection mode (1  $\mu\text{L}$  injection volume) using a 70:30 v/v water and acetonitrile with 0.1% formic acid weak wash solution to match both the composition of the standards and gradient starting conditions. Buffer A was water with 0.1% formic acid and buffer B was acetonitrile with 0.1% formic acid. Gradient elution was conducted using a 10 minute linear gradient starting at 70% A and ending at 3% A. The column was re-equilibrated at the starting conditions after the end of the gradient. The 2.1 mm column, 1.0 mm column, 300  $\mu\text{m}$  column, and 150  $\mu\text{m}$  iKey Separation Device were operated at flow rates of 600, 250, 15, 5  $\mu\text{L min}^{-1}$  corresponding to initial backpressures of 6300, 7525, 7750, 5000 psi, respectively.

### *Ion Sources and MS Configuration*

The 2.1 mm, 1.0 mm, and 300  $\mu\text{m}$  columns were connected to a Waters Z-Spray ion source. The iKey Separation Device was

operated with the Waters ionKey source. Both sources were installed and configured on a Waters Xevo G2 TOF and operated with LockSpray to provide accurate mass measurements and compensate for  $m/z$  calibration drift [29, 30]. It should be noted that this mass spectrometer does not have the Waters StepWave found in their more modern implementation of the API. The sources were optimized based off manufacturer recommended settings with slight modifications. TOF operation and data acquisition was achieved using Waters MassLynx (ver. 4.1) appropriately configured for each instrument configuration. Mass spectra were acquired as centroided spectra at 5 Hz in the negative ion mode from 50 to 2000  $m/z$  using the  $MS^E$  acquisition mode [31].  $MS^E$  acquisition alternates between acquiring two independent mass spectra: one mass spectrum is acquired with low collisional energy and the other with high collisional energy without any precursor ion selection. A collisional energy ramp from 50 to 110 V was performed during high energy operation [32]. All source configurations were operated with the sampling cone voltage set to 10 V, the extraction cone set to 4 V, the source temperature set to 150 °C, and the cone gas set to a flow rate of 30 L Hr<sup>-1</sup>. The Z-Spray and ionKey sources were operated at a capillary voltages of -1.8 and -2.5 kV, respectively. The desolvation gas flow rate and temperature were different dependent on the column flow rate following manufacture guidelines. The Z-Spray source operated with the 300  $\mu$ m, 1.0 mm, and 2.1 mm columns was operated with desolvation gas flows rates and temperatures of 600 L Hr<sup>-1</sup> (300 °C), 600 L Hr<sup>-1</sup> (450 °C), and 800 L Hr<sup>-1</sup> (600 °C), respectively; the ionKey source was operated without desolvation gas.

### Data Collection, Peak Picking, and Compound Spectra Determination

A standard curve was acquired on each of the four instrument configurations starting with the lowest concentration bile acid standard mix. This process was then repeated in triplicate. Uncertainties are reported as either standard deviations or propagated and added in quadrature when appropriate. Single component standards were used to verify compound retention time and identity across the column formats as isobaric bile acids were included in the analysis (Supplement).

The Waters .raw files were converted to the Net Common Data Format (netCDF) [33] using Waters Databridge conversion software. Next, these files were grouped into individual calibration curves for processing with xcms (ver. 1.46.0) [34, 35] using the centWave algorithm [36] within the R environment (ver. 3.2.5) [37]. Relevant portions of the xcms R object were exported from the R environment and imported into Igor Pro (WaveMetrics, Inc. ver. 7.04) for all subsequent data processing, analysis, and visualization. Additionally, the exact mass bins of each xcms feature were used to bin and extract high-resolution ion chromatograms to avoid arbitrarily binning the high-resolution data. These data were also exported from the R environment and imported into Igor Pro allowing for

analysis on features (peak areas, descriptive peak statistics, etc.) and scan-by-scan as extracted ion chromatograms (EICs).

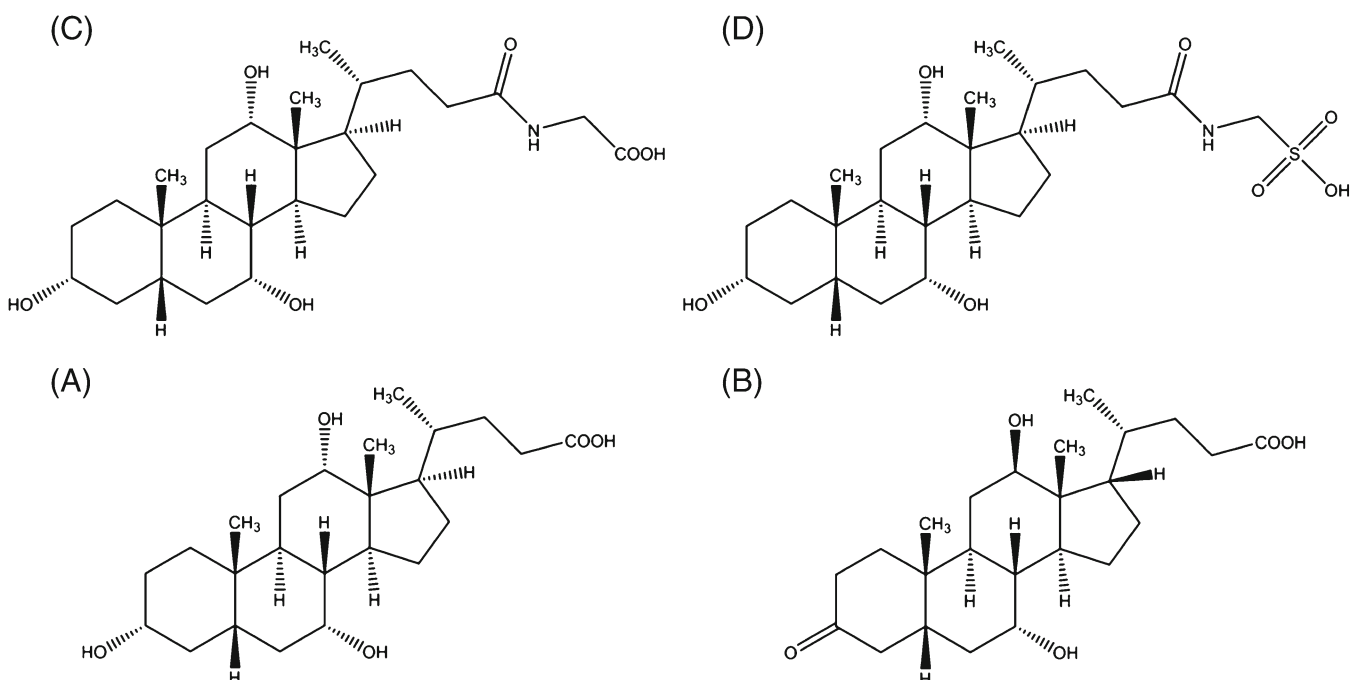
A custom processing algorithm was developed in Igor Pro with the aim of grouping together all the signals from both the MS and  $MS^E$  spectra that arise from a single compound.

These reconstructed sets of ions are termed *compound spectra* [23]. The implementation of this algorithm focuses on quantitatively characterizing in-source chemistry (ion-neutral clustering) although fragmentation data from the  $MS^E$  are also processed and presented. The algorithm was realized by first searching the xcms peak group data for any features corresponding to the mass of  $[M-H]^-$  for a species of interest. The  $[M-H]^-$  peak was known to be present in large abundance for the compounds of interest. Then, all features with retention times (peak apex) near the center of the  $[M-H]^-$  exact mass match are grouped together as preliminary compound spectra. Last, Pearson's  $r$  was calculated for every high-resolution EIC against the  $[M-H]^-$  EIC within the retention time window of the  $[M-H]^-$  EIC to verify each ion actually belongs to the compound spectrum. Subsequent analysis was conducted using these verified compound spectra.

## Results

### Species Dependent Compound Spectra

Low collisional energy compound spectra exhibited significantly different distributions of ion-neutral clusters depending on the bile acid investigated. These compounds show similar patterns by broad bile acid subclasses as defined by the functional groups found on the steroidal ring or on the conjugation site; this is consistent with previous modeling work attempting to predict in-source phenomenon for metabolite annotation where classes of compounds were found to exhibit similar, though not identical, in-source behavior [1]. Many of the same ion-neutral clusters are present for each species investigated but in differing relative abundancies. For these reasons, and to keep the data presented manageable, the presented analysis is limited to four representative bile acids: cholic acid (C<sub>24</sub>H<sub>40</sub>O<sub>5</sub>), 3-oxocholeic acid (C<sub>24</sub>H<sub>38</sub>O<sub>5</sub>), glycocholic acid (C<sub>26</sub>H<sub>43</sub>NO<sub>6</sub>), and taurocholic acid (C<sub>26</sub>H<sub>45</sub>NO<sub>7</sub>S) (Figure 1). These species were baseline resolved or nearly baseline resolved with little to no co-elution (Supplemental Information). Figure 2 shows the low-energy MS and collisional energy ramped  $MS^E$  compound spectra of the four representative bile acids. These data were acquired during the same chromatographic run on the 1.0 mm × 50 mm T3 column at 5 ng on column using the Z-Spray ESI source. Each of the four bile acids generated multiple signals arising from ion-neutral clusters. For example, cholic acid produced at least 20 features with three or more points across the calibration range for all the flow regimes with increasing numbers of features observed at higher mass loadings (supplemental information). Little to no fragmentation was observed in the low collisional energy MS mode. Many ion-neutral clusters present in the MS compound spectrum are also present in the  $MS^E$  compound spectrum; this means that some of these

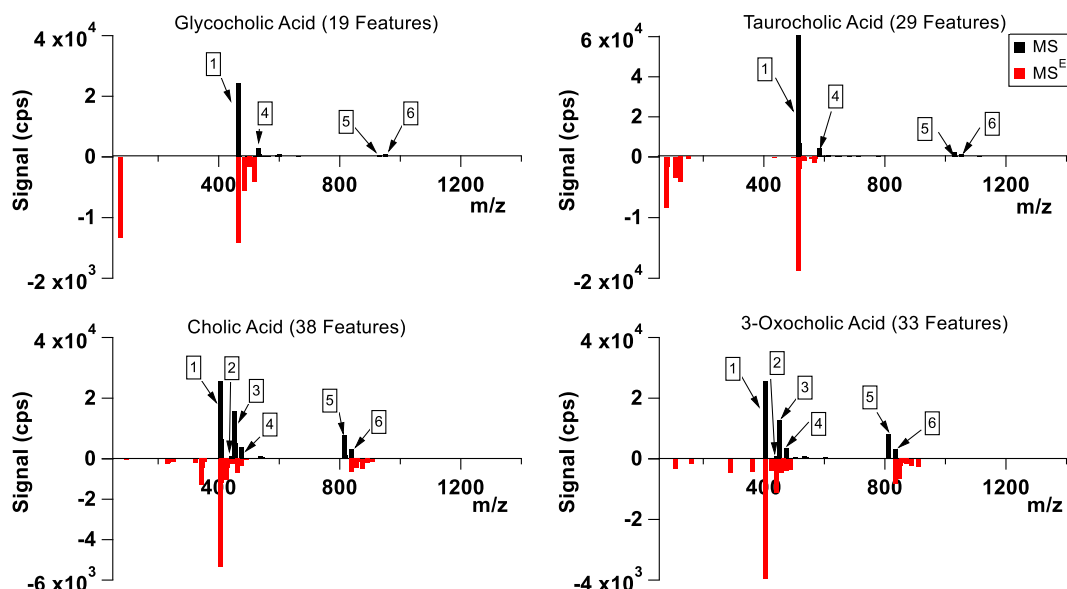


**Figure 1.** Structures of the representative bile acid classes. (A) Cholic acid, (B) 3-oxocholeic acid, (C) glycocholic acid, (D) taurocholic acid

clusters are stable enough to survive at least a portion of the  $MS^E$  collisional energy ramp which starts at 50 V.

The spectral similarity of the compound spectra was investigated by calculating the cosine similarity of the four representative bile acids across all concentrations and other species investigated [38, 39] (see Supplement for details). The maximum concentration (10 ng on column) was used as the

reference mass spectrum upon which all other spectra are compared. The procedure uses mass-gridded and normalized compound spectra where each EIC extracted across all files are ordered in a fixed list regardless of whether or not a peak was detected in that individual file. Changes in concentration for a bile acid of interest lead to substantial changes in the compound spectrum reflected in the calculated spectral similarity scores.



**Figure 2.** Verified compound spectra of cholic acid, 3-oxocholeic acid, glycocholic acid, and taurocholic acid. Ion signal intensities are shown at the chromatographic peak maximum. All data were acquired on the Z-Spray source with a 1.0 mm  $\times$  50 mm T3 column at a mass loading of 5 ng on column. MS (low-energy) ion signals are plotted as positive (black) bars and  $MS^E$  (high-energy) ion signals are plotted as negative (red) bars. The number of features (isotopes included) in the MS (low-energy) spectra are included in the title of each graph subset. Various high-abundance and commonly observed ions are annotated based on their exact mass values and isotopic ratios: (1)  $[M-H]^-$ , (2)  $[M-H+HCl]^-$ , (3)  $[M-H+FA]^-$ , (4)  $[M-2H+FA+Na]^-$ , (5)  $[2M-H]^-$ , (6)  $[2M-2H+Na]^-$

This highlights the fact that all the ions or ion-neutral clusters arising from a single species do not increase linearly with concentration and each ion or cluster has its own sensitivity, limit of detection, and linear range (discussed in subsequent sections). The ions in each compound spectrum are not fixed in relative abundance but change as a function of concentration.

### Solvent Flow Rate

The source configuration significantly influenced the compound spectrum observed for each bile acid driven primarily by the solvent flow rate determined by the column inner diameter (Figure 3). The solvent flow rate has two major effects on the electrospray process: (1) lower flow rates generate smaller droplets leading to increased ionization efficiency and higher sensitivity [40, 41], and (2) the lower flow rates bias the concentration of eluting compound towards higher instantaneous analyte concentration despite increased chromatographic peak width on the smaller scale separation devices. It should be noted that the physical source design (e.g., spray direction, distance from ESI emitter to MS inlet, emitter type, etc.) does not appear to significantly change the observed ion-neutral clustering in a way that deviates from our analysis of solvent flow rate. The use of the ionKey source and iKey Separation Device compared with the use of standard analytical columns with a standard ESI source highlights this finding. It is the ESI process itself occurring at each flow rate that primarily changes the behavior of ion-neutral clusters rather than the source hardware. The results obtained with the completely different ionKey source system follow the results obtained using standard columns.

We identified two general classes of ion-neutral clusters in these systems that arise from the bile acids. First, *contaminant-*

*containing clusters*: these clusters consist of ions or neutrals arising from species not intentionally added to the system (e.g., no sodium was added to our buffer system, but  $\text{Na}^+$  is present in trace amounts in the solvents). Second, *buffer-additive-containing clusters* (e.g., formic acid was intentionally added to our buffer system) and *retained species clusters* such as homodimers/homo-multimers arising from the compound of interest, which has been chromatographically retained (e.g.,  $[\text{nM-H}]^+$ , where  $n = 1, 2, 3, \dots$ ). Low-flow LC configurations (M-Class UPLC:  $5, 15 \mu\text{L min}^{-1}$ ) bias the compound spectrum towards homodimers and buffer containing clusters. High-flow LC configurations (Acquity UPLC:  $250, 600 \mu\text{L min}^{-1}$ ) also contain homodimers and buffer containing clusters in substantial abundance, but other contaminant-containing clusters congest the mass spectrum due to the increased solvent flux.

### Concentration Dependence of Observed Ion-Neutral Adducts

Calibration curves of select ions in the four representative compound spectra are presented in Figure 4 along with the sum of all the ions in the MS level compound spectra. These data were acquired on the  $1.0 \text{ mm} \times 50 \text{ mm}$  T3 column and reported in triplicate with the error bars representing the standard deviation of the averaged, integrated peak areas. Calibration curves obtained on the other columns are presented in the [Supplemental Information](#). The average relative standard deviation (RSD) of each point in Figure 4 is less than  $\sim 10\%$ ; this is consistent with nearly all the columns investigated. More importantly, the RSD of the ion-neutral adducts for all column/source configurations is consistent with the corresponding  $[\text{M-H}]^+$  RSD showing that these clusters are highly reproducible under constant instrument conditions.

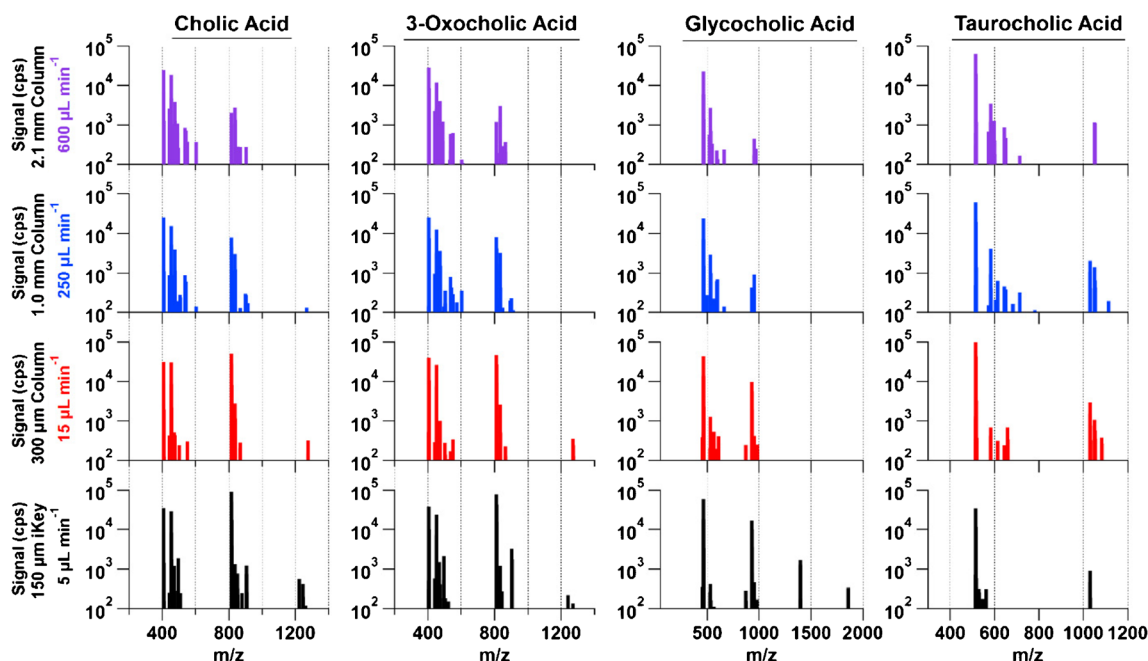
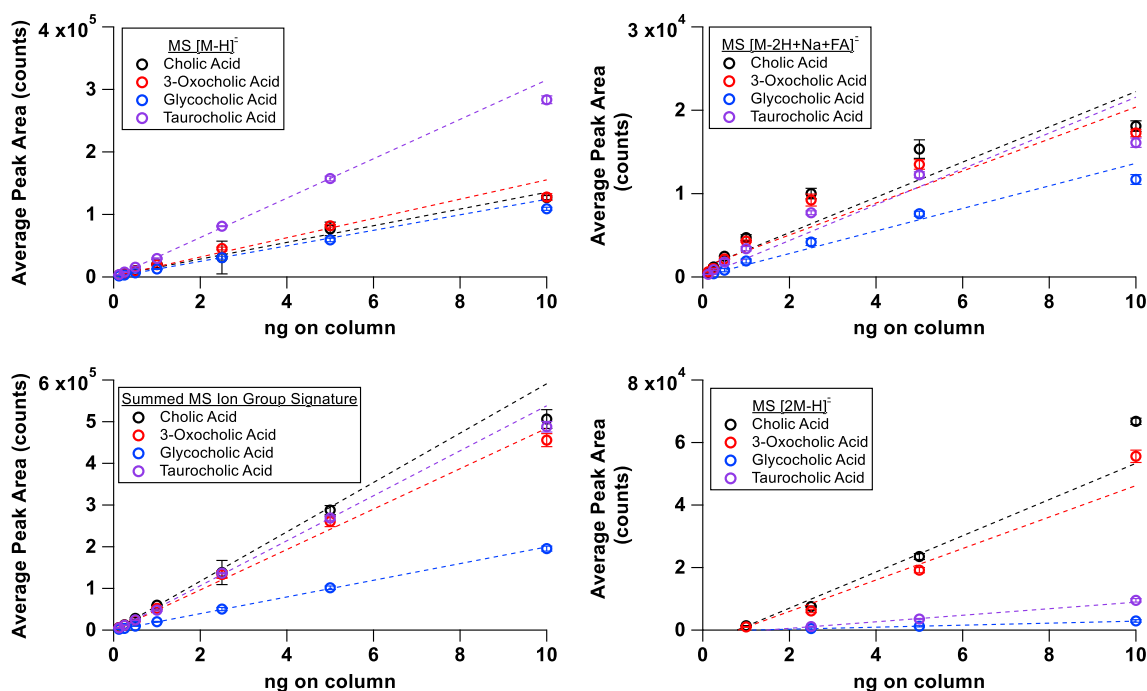


Figure 3. Compound spectra of the four representative bile acid species are shown for the four different instrument configurations:  $150 \mu\text{m}$  iKey (black),  $300 \mu\text{m}$  capillary column (red),  $1.0 \text{ mm}$  column (blue), and  $2.1 \text{ mm}$  column (purple)



**Figure 4.** Calibration curves of select ions from cholic acid (black), 3-oxocholeic acid (red), glycocholeic acid (blue), and taurocholeic acid (purple) are plotted in averaged triplicate with error bars representing the standard deviation of the average. These data were all collected on the Z-Spray source with a 1.0 mm  $\times$  50 mm T3 column. Weighted least squares linear regression results are also plotted. The sum of the MS ions in the compound spectra (bottom left),  $[M-H]^-$  (top left),  $[2M-H]^-$  (bottom right), and  $[M-2H+FA+Na]^-$  (top right) are provided to illustrate representative behaviors

Figure 4 and supplemental calibration curves show all calibration points above the limit of detection. Weighted least squares linear regression was performed on the calibration curves because non-linearity was observed with some of the ions at high concentrations. These fits were included in Figure 4 and the supplemental data to draw the eye to deviations from linear behavior and should not be thought of as traditional calibration curves fitted over the linear range. The calibrations were conducted using the same on-column mass loadings for each column, which lead to different regions of the linear range being investigated on each column. Some points are below the limit of quantification (excluded from analysis), and other points exhibit behavior indicative of saturating the electrospray. Saturation is obvious at the 10 ng loading on the iKey, but some deviation from linear behavior is observed for the  $[M-H]^-$  ion at the 10 ng loading for all columns. Electrospray saturation occurs where the instantaneous concentration of bile acid eluting from the column exceeds  $\sim 10^{-5}$  M [42]. The  $10^{-5}$  M saturation concentration is generally agreed upon, but compound dependent differences are known to exist and limit the linear range [43].

Least squares linear regression (non-weighted) was also performed from 0.125 ng to 5 ng to better understand the linear range and to understand how to best deal with these phenomenon in the case of non-targeted metabolomics where internal standards are absent (Supplement). The sum of the MS compound spectrum from 0.125 ng to 5 ng consistently exhibits the highest  $r^2$  value while also containing the most calibration points across the calibration range. The  $[M-H]^-$  calibration

curves also have high  $r^2$  values, but the species that exhibited extensive clustering behavior (cholic acid and 3-oxocholeic acid) show lower  $r^2$  values compared with the sum of the compound spectra. The lower  $r^2$  values for the  $[M-H]^-$  calibration curve are most obvious for the low-flow systems probably due to an increased abundance of the  $[2M-H]^-$  cluster. The summed compound spectra approach is a promising route to robust quantification when complex in-source phenomenon are observed; this is consistent with previous work quantifying ginkgolides and bilobalide, where various adducts were summed to improve linearity [5]. It should be highlighted that in the case of obvious saturation (e.g., iKey at 10 ng on column) the summed calibration method does not improve linearity since the deviation from linear behavior is so severe.

The sensitivities of the various ion source configurations follow the anticipated ESI behavior with the sensitivity increasing with lower flow rates for the mass loadings between 0.125 and 5 ng on column. For example, the  $[M-H]^-$  ion of taurocholeic acid on the 2.1 mm column exhibits an absolute sensitivity of  $1.99 \times 10^4$  counts  $\text{ng}^{-1}$  with an  $r^2$  value of 0.9994. The 1.0 mm column and 300  $\mu\text{m}$  capillary column show increasing sensitivities relative to the 2.1 mm column of 1.6x and 6.3x with  $r^2$  values of 0.9995 and 0.9971, respectively. The iKey shows a relative increase of 4.3x over the 2.1 mm column with an  $r^2$  value of 0.9908 due to peak broadening on the iKey device.

The highest mass loadings (5–10 ng on column) and lowest flow rates produce higher-order clusters observed in the mass range corresponding  $[3M-H]^-$  and  $[4M-H]^-$

(Figure 3). Homo-multimer formation occurs to the greatest extent on the iKey device which was operated at the lowest flow rate ( $3 \mu\text{L min}^{-1}$ ) leading to the highest instantaneous concentrations and saturation at 10 ng. The mass differences observed on the iKey are consistent with pure homo-multimer formation following the pattern of  $[\text{nM-H}]^-$  for cholic acid, glycolic acid, and taurocholic acid. Interestingly, 3-oxocholic acid does not form a  $[\text{3M-H}]^-$  cluster but rather a  $[\text{3M-2H+Na}]^-$ . Various other unidentified clusters exist in these mass ranges when using the iKey device, but the most important finding is that at higher flow rates the higher order clusters ( $n > 2$ ) contain other species making annotation of these signals more challenging and more complex compared to the iKey. With the exception of taurocholic acid, the pure homo-multimer,  $[\text{nM-H}]^-$ ,  $n > 2$  is absent for all species investigated on all other source/column configurations except the iKey. The clusters in this mass range on the  $300 \mu\text{m}$  column are not pure homo-multimer species.

### Cluster Dependence Across Solvent Flow Rates

The ion source, concentration, and analyte dependence can be further investigated as a function of instantaneous analyte concentration linking the behavior of the various experimental configurations together. The instantaneous mass of a bile acid species eluting from the column was calculated by assuming that 100% of the injected mass elutes during the observed peak. Then, the fraction of the total signal at each point defining the chromatographic peak can be used to estimate the mass of compound being eluted. Similarly, the integrated peak areas can be plotted in the same coordinate-space by integrating the solvent flow rate across the entire peak. The number of moles of bile acid injected on column is divided by the integrated solvent volume to provide an average concentration across the peak.

This scan-by-scan approach allows for a much broader range of concentrations to be evaluated compared to only examining integrated peak areas, takes advantage of all the available data, and highlights the changing relative abundance of ions across an eluting chromatographic peak. Figure 5 shows the  $[\text{M-H}]^-$  ion plotted against the instantaneous concentration of bile acid. It should be noted that the assumptions made to calculate instantaneous analyte concentration scan-by-scan may fail at the highest on-column mass loadings due to ESI saturation. At high concentrations, the ESI response to instantaneous changes in concentration becomes non-linear biasing the scan-by-scan concentrations. This is observed in Figure 5 since the integrated peaks show slight deviations from linearity in log-log space compared to the individual points which were calculated assuming signal response is directly proportional to concentration.

The data acquired using 1.0 and 2.1 mm columns fall along a single linear trend line for each source in Figure 5. Interestingly, the  $300 \mu\text{m}$  capillary column and  $150 \mu\text{m}$  iKey exhibit multiple offset linear trend lines for all the species except for

taurocholic acid on the  $300 \mu\text{m}$  capillary column. These linear patterns are highly reproducible and highlighted in the [supplemental information](#). Each individual linear trend line corresponds to a different mass of bile acid injected on column, and all the data presented in Figure 5 are shown in triplicate (i.e., multiple injections at the same concentration all fall on the same trend line). We hypothesize that this behavior arises from two sources: (1) non-linear ESI response at high concentrations biases the calculation of the x-axis, and (2) the iKey and  $300 \mu\text{m}$  capillary column exhibit the highest extent of homo-dimer formation which behaves non-linearly as a function of concentration.

These hypotheses are supported by the fact that taurocholic acid ( $[\text{M-H}]^-$ ) falls along a single trend line on the  $300 \mu\text{m}$  capillary column, produces few ion-neutral clusters, and shows little deviation from linear behavior across the entire calibration range. Conversely, glycocholic acid ( $[\text{M-H}]^-$ ) on the same column shows multiple trend lines, produces more ion-neutral clusters than taurocholic acid, and also shows little deviation from linear behavior across the entire calibration range. Thus, cluster formation in competition with ( $[\text{M-H}]^-$ ) appears to also produce these trend lines.

The other ions in the compound spectrum can be investigated by examining the ratio of some ion-neutral cluster of interest divided by the corresponding  $[\text{M-H}]^-$  signal intensity. Two examples are provided following the data displayed in Figure 4: the ratio of  $[\text{2M-H}]^-/[\text{M-H}]^-$  (Figure 6) and  $[\text{M-2H+FA+Na}]^-/[\text{M-H}]^-$  (Figure 7). These two species represent the general phenomenon occurring in these experiments. The  $[\text{2M-H}]^-/[\text{M-H}]^-$  ratio represents the dominant phenomenon occurring at lower solvent flow rates where conditions lead to a higher concentrations of bile acid relative to solvent vapors, buffer additives, and contaminants. Conversely,  $[\text{M-2H+FA+Na}]^-/[\text{M-H}]^-$  represents contaminant influence that becomes more important at higher flow rate operation with standard analytical columns. This ion-neutral cluster contains sodium, which is a contaminant species present in higher relative abundances at higher solvent flow rates. The  $[\text{2M-H}]^-$  ratio linearly increases with the concentration of M because the formation of this species is dependent on M (note the log-scale). The abundance of the  $[\text{M-2H+FA+Na}]^-$  relative to  $[\text{M-H}]^-$  should decrease at lower solvent flow rates which limit the abundance of  $\text{Na}^+$ . This trend is observed in the data. These results across instrument configurations show that the concentration of various species in the solvent and the various available ion-neutral adduction pathways determine the observed ion distribution.

Close examination of Figure 6 shows that clustering under different flow regimes is not entirely explained by the instantaneous concentration of bile acid eluting from the column. This is due to numerous other processes occurring in the electrospray across flow regimens during gradient elution. Not only does the concentration profile of each compound differ with each column, the ESI sensitivity is also changing due to different flow rates. Examining the two species in Figures 6 and 7 also tells an

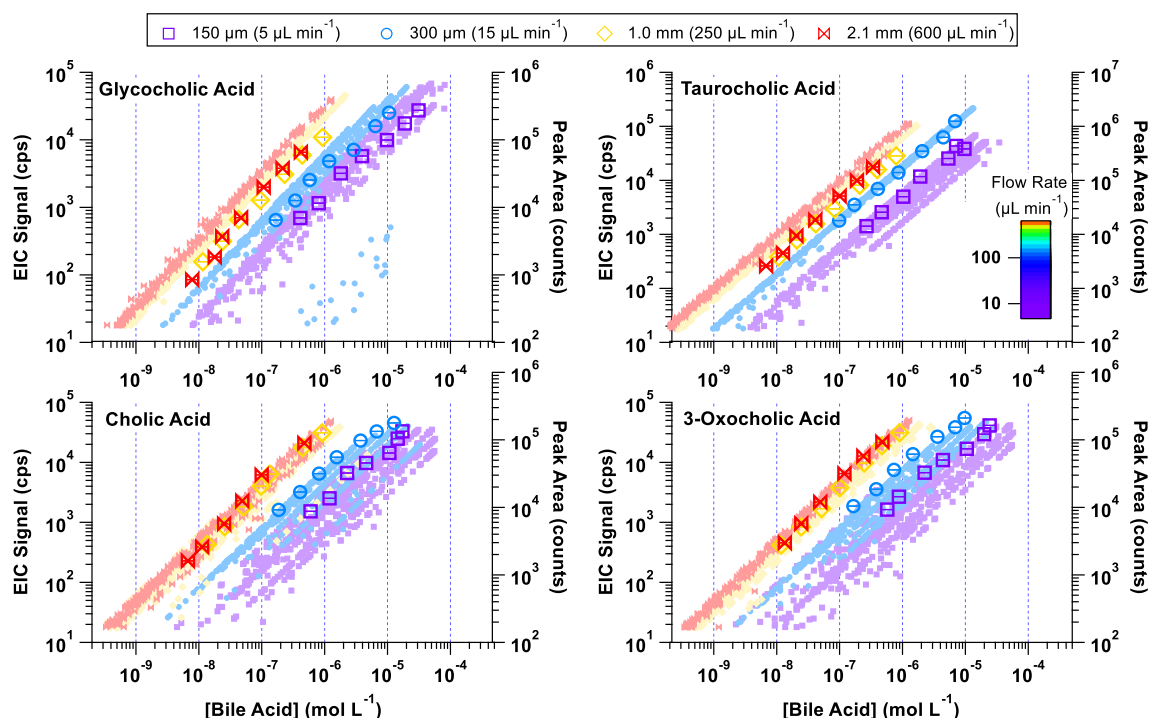


Figure 5.  $[M-H]^-$  bile acid signal intensity plotted against the calculated concentration of bile acid eluting from the column for both integrated peak areas (large markers, right axis) and extracted ion chromatograms (small markers, left axis) for all instrument configurations. The integrated peak areas are the average of triplicate injections with error bars representing the standard deviation of the average. Extracted ion chromatograms are plotted spectra by spectra but are also displayed in triplicate

incomplete story since the distribution of clusters is different with each column. This again arises from numerous possible factors including the composition of the solvent during the eluting peak [9], the solute concentration [43,

44], concentration of various contaminants [9], gas-phase chemistry arising from neutral species in thermal equilibrium with the produce ions [22], and solvent flow rate though the electrospray emitter [40, 41].

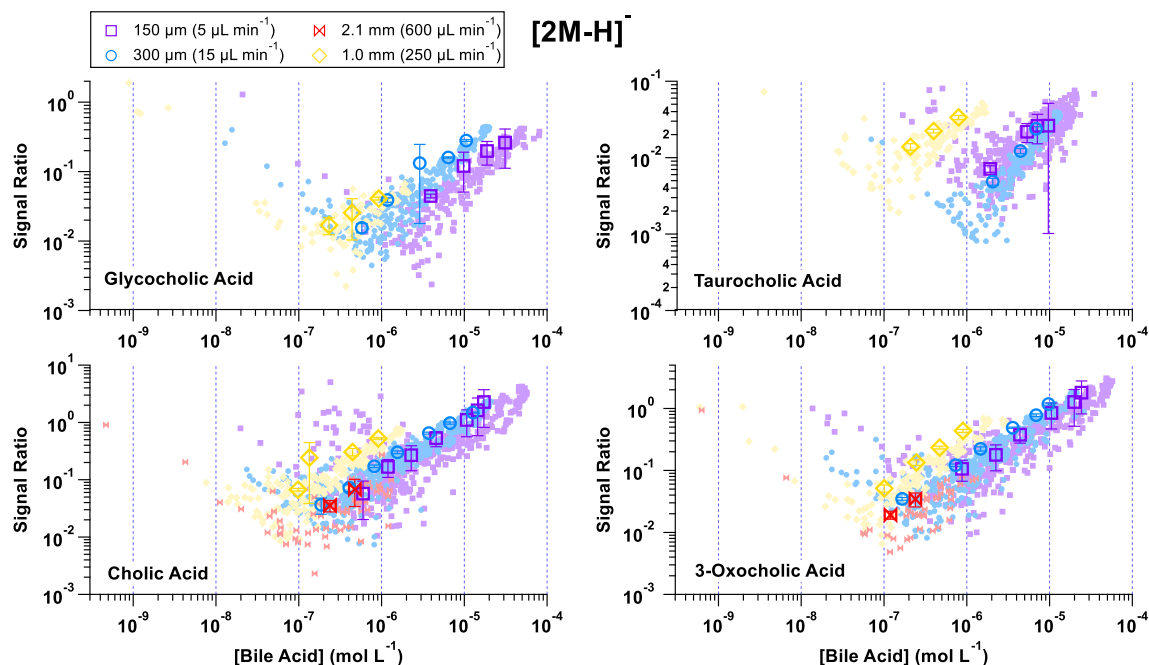


Figure 6. Ratio plots of the  $[2M-H]^-$  ion signal intensity relative to the  $[M-H]^-$  ion signal intensity plotted against the calculated concentration of bile acid eluting from the column. Integrated peak areas (large markers, right axis) and extracted ion chromatograms (small markers, left axis) are displayed for all instrument configurations



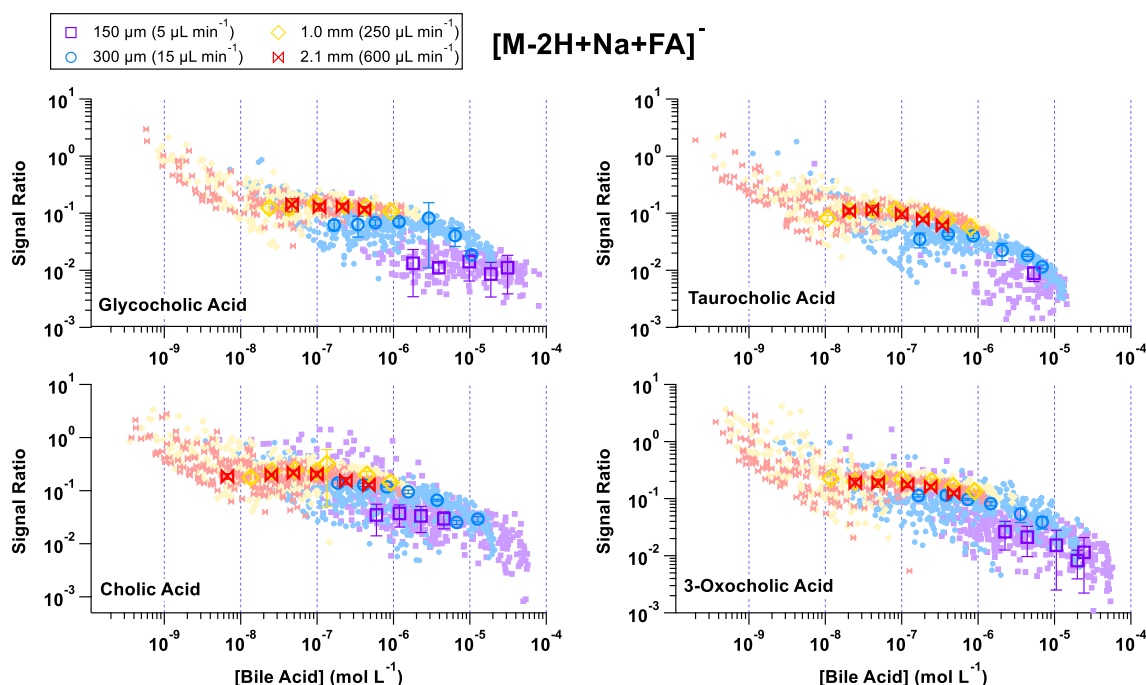


Figure 7. Log scale ratio plots of the  $[M-2H+Na+FA]^-$  ion signal intensity relative to the  $[M-H]^-$  ion signal intensity plotted against the calculated concentration of bile acid eluting from the column. Integrated peak areas (large markers, right axis) and extracted ion chromatograms (small markers, left axis) are displayed for all instrument configurations.

## Discussion

### *Ionization and Clustering Mechanisms*

The various observations above cannot be reconciled with standard ESI models [40, 41, 44–46] (CRM and IEM) and by the potential for the gas-phase formation (or transformation) of clusters during transit from the source through the API without knowing which processes are ultimately controlling the observed ion distribution [8]. Lower solvent flow rates will generate smaller solvent droplets with higher concentrations of the retained analyte species in each droplet if the same mass of analyte is loaded on the column. As droplet evaporation, droplet fission, and/or ion-evaporation occur, the likelihood of multiple analyte species forming dimers and higher-order clusters increases under low-flow conditions relative to higher-flow systems due to increased analyte concentration. High-flow rate systems will produce larger droplets with lower analyte concentrations. As these droplets undergo fission and/or ion-evaporation, the ratio of contaminant species and buffer additives to analyte species will be larger than in the low-flow systems leading to more contaminant and buffer containing clusters.

Similarly, in the gas phase, one would expect all the thermalized ions and ion-neutral clusters produced in the electrospray to rapidly cool during transfer through the high pressure stages of the API reaching thermal equilibrium [15, 17, 19–21, 47]. At constant contaminant concentrations in the mobile phase (e.g.,  $Na^+$ ), increased solvent flow rates result in increased contaminant and buffer additive gas-phase concentrations relative to the eluting compound. This shifts the equilibrium towards contaminant and buffer-additive-containing

clusters and away from the homodimer (retained species) population. Conversely, low-flow LC conditions will produce much higher concentrations of bile acid and lower concentrations of contaminant species leading to a relative increase in homodimers or retained species clusters.

All of the observed ions arise from complex processes where various ions/clusters formed in the ion source are being transmitted to the detector to various extents and fragmenting and/or clustering into other potentially observable ions and clusters even under low collision energy MS mode [19–21, 47, 48]. The cluster distribution in the ion source or in the initial stages of the API is unknown, but the observed ions (the mass spectrum) depend on the initial ion distribution and configuration of the mass spectrometer (i.e., internal temperature of ions in the source, collisional/translational energies through each optic, thermal equilibration during adiabatic expansion between pressure stages, etc.) (see [19] and [47] for a detailed discussion of clusters and the processes occurring in APIs). This makes understanding the exact formation mechanism of each member of these compound spectra very difficult. More importantly, the behavior and general trends that are observed for each species can be investigated to provide some constraints to better understand these complex systems. This becomes valuable for groups attempting to extract the maximum amount of chemical information from the mass spectrum. These data suggest that the matrix effects and differences across different platforms can actually be understood and investigated for method optimization. Sample matrix changes the chemical environment during the ionization process as well as during the ion's initial transfer into the API leading to changes in

sensitivity, ionized species, and analytical performance. Many of these reactions are observable if the mass spectrum is carefully examined and statistical grouping tools are applied to track the observed ions back to the individual compound.

### Linearity and Concentration Response

The increased linear behavior observed for the summed MS compound spectra makes sense when examining the trends of the various ion-neutral clusters. For example,  $[M-2H+Na+FA]^-$  saturates at the highest mass loadings while the  $[2M-H]^-$  cluster begins to increase exponentially. This behavior was also observed by Zook et al. [8] when infused analyte concentration became sufficiently high to saturate the electrospray ( $\sim 10^{-5}$  M). It is likely that this system is  $Na^+$  limited leading to the saturation of the  $[M-2H+Na+FA]^-$  cluster while the  $[2M-H]^-$  cluster depends only on the concentration of the bile acid species, M.

Metwally et al. differentiate the peak splitting due to adduct formation and ion suppression in their work investigating salt-induced signal suppression in protein detection using electrospray [6]. The authors define ion suppression as the point at which salt concentrations reduce the total analyte signal. This is mostly consistent with the data presented herein, but requires some clarification in light of this work and earlier studies examining these effects [8, 9]. Summing the ion intensities of each ion in the MS level compound spectra improves linearity compared to only using the most abundant signal,  $[M-H]^-$ ; this is clearly consistent with peak splitting. At sufficiently high analyte (or salt) concentrations where charge competition begins to dominate and reduce ionization efficiency saturation as defined by Metwally *et al.* occurs. Simultaneously, the formation of homo-multimers ( $[nM-H]^-$  or  $[nM+H]^+$ ) increases when the saturation is driven by the analyte concentration. Thus, a substantial increase in homo-multimer formation under analyte saturating conditions becomes a dominant feature of the mass spectrum. This differs from salt-induced ion suppression in that the analyte is inducing the suppression effects. Thus, peak splitting continues to occur under high concentrations, which leads to ion suppression. Below concentrations leading to ion suppression, the intensity of various adducts should be summed to increase linearity because the signal is simply being divided into different ion  $m/z$  bins.

The role of instantaneous analyte concentration across a chromatographic peak is important for the formation of and the competition between various ions and ion-neutral adducts. One needs to be careful to not oversimplify this generalization; previous work has shown that increasing the buffer concentration (i.e., salt) does not directly lead to increased buffer-containing adducts [49, 50]. Again, this is likely due to the complex relationships between various ion formation pathways and subsequent chemistry occurring in the API. Increasing salt or buffer concentrations reduces the extent to which analyte species are initially ionized due to charge competition. Thus, this immediately reduces the concentration of analyte ions

preventing the formation of analyte-ion containing clusters, regardless of where they form.

## Conclusions

The complexity and diversity of ions observed for bile acids detected in the negative mode is extensive, and it significantly deviates from the canonical model of molecular or “pseudo-molecular” ions ( $[M-H]^-$  or  $[M+H]^+$ ) dominating the mass spectrum. Decades of work have attempted to understand the behavior and formation mechanisms of ion-neutral clusters generated from electrosprays, highlighting the fact that this is a ubiquitous phenomenon of electrospray ionization and atmospheric pressure ionization techniques in general. Understanding the relationship of the ions and ion-neutral clusters generated in electrosprays and their subsequent transformations during transmission through API mass spectrometers remains a significant challenge to both experimentalists and modelers. The results presented herein provide qualitative and quantitative descriptions of the behavior of these clusters in the context of LC-ESI-MS, but further studies are needed to fully explore the mechanisms of cluster formation.

The chemical diversity of the molecules under investigation for metabolomics workflows, particularly non-targeted workflows, necessitates that these processes are given more attention as quantification and robust annotation remain ongoing challenges [1, 3]. Quantification by summing all signals arising from a compound under low collisional energy operation improves linearity and increases the sensitivity. The complexity observed with this structurally similar set of molecules should not be thought of as rare, isolated, or worst case scenarios but as ubiquitous phenomenon [1, 3, 19, 51]. Many of these adducts are energetically stable with sufficiently high binding energies that collisional dissociation of all the clusters would inevitably cause unwanted and substantial fragmentation of other species.

The distribution of ions changes across LC flow regimes following the availability of ions and neutral species. Standard analytical columns (2.1 and 1.0 mm) operated under typical UPLC conditions lead to higher levels of contaminant-containing ion-neutral adducts compared to micro-flow regimes. The micro-flow regimes produce ion distributions with an increased likelihood of observing homodimers and higher order clusters with retained analytes driven primarily by instantaneous solute concentration.

The complexity of high-resolution UPLC-ESI-MS data require data-driven workflows that make fewer assumptions and operate on both individual data files and data sets. Most models and statistical methods would be incapable of dealing with the compound spectra where the relative abundance of each analyte is changing across the concentration range (often non-linearly). Spectral similarity scores developed for matching experimental spectra to reference spectra show limitations for these compound spectra due to the ions shifting in relative abundance as a function of concentration. When taken as a

whole and considering the amount of data produced on high-resolution mass spectrometers, integration of all available data for a single sample set or individual experiment remains a distant goal, but we have reached a level of understanding where the benefits of such approaches are undeniable for improving annotation and compound identification.

## Acknowledgments

The authors acknowledge Waters Corporation for postdoctoral research funding and support for this project. The authors also acknowledge Angela Doneanu of Waters Corporation for technical support.

## References

1. Broeckling, C.D., Ganna, A., Layer, M., Brown, K., Sutton, B., Ingelsson, E., Peers, G., Prenni, J.E.: Enabling efficient and confident annotation of LC-MS metabolomics data through MSI spectrum and time prediction. *Anal. Chem.* **88**, 9226–9234 (2016)
2. Alechaga, É., Moyano, E., Galceran, M.T.: Ion-molecule adduct formation in tandem mass spectrometry. *Anal. Bioanal. Chem.* **408**, 1269–1277 (2016)
3. Keller, B.O., Sui, J., Young, A.B., Whittall, R.M.: Interferences and contaminants encountered in modern mass spectrometry. *Anal. Chim. Acta* **627**, 71–81 (2008)
4. Krueve, A., Kaupmees, K.: Adduct formation in ESI/MS by mobile phase additives. *J. Am. Soc. Mass Spectrom.* **28**, 887–894 (2017)
5. Li, X.-F., Ma, M., Scherban, K., Tam, Y.K.: Liquid chromatography-electrospray mass spectrometric studies of ginkgolides and bilobalide using simultaneous monitoring of proton, ammonium and sodium adducts. *Analyst* **127**, 641–646 (2002)
6. Metwally, H., McAllister, R.G., Konermann, L.: Exploring the mechanism of salt-induced signal suppression in protein electrospray mass spectrometry using experiments and molecular dynamics simulations. *Anal. Chem.* **87**, 2434–2442 (2015)
7. Gamero-Castaño, M., Fernández De La Mora, J.: Mechanisms of electrospray ionization of singly and multiply charged salt clusters. *Anal. Chim. Acta* **406**, 67–91 (2000)
8. Zook, D.R., Bruins, A.P.: On cluster ions, ion transmission, and linear dynamic range limitations in electrospray (ionspray) mass spectrometry. *Int. J. Mass Spectrom. Ion Processes* **162**, 129–147 (1997)
9. Meng, C.K., Fenn, J.B.: Formation of charged clusters during electrospray ionization of organic solute species. *Org. Mass Spectrom.* **26**, 542–549 (1991)
10. Dole, M., Mack, L.L., Hines, R.L., Mobley, R.C., Ferguson, L.D., Alice, M.B., Malcolm Dole, L.L.M., Marck, L.L., Hines, L., Mobley, R.C., Ferguson, L.D., Alice, M.B.: Molecular beams of macroions. *J. Chem. Phys.* **49**, 2240–2249 (1968)
11. Thomson, B.A.: Atmospheric pressure ionization and liquid chromatography/mass spectrometry—together at last. *J. Am. Soc. Mass Spectrom.* **9**, 187–193 (1998)
12. Thomson, B.A., Iribarne, J.V., Dziedzic, P.J.: Liquid ion evaporation/mass spectrometry/mass spectrometry for the detection of polar and labile molecules. *Anal. Chem.* **54**, 2219–2224 (1982)
13. Thomson, B.A., Iribarne, J.: V: Field induced ion evaporation from liquid surfaces at atmospheric pressure. *J. Chem. Phys.* **71**, 4451–4463 (1979)
14. Iribarne, J.V., Thomson, B.A.: On the evaporation of small ions from charged droplets. *J. Chem. Phys.* **64**, 2287 (1976)
15. Hayes, J.M., Small, G.J.: Supersonic jets, rotational cooling, and analytical chemistry. *Anal. Chem.* **55**, 565A–574A (1983)
16. Proctor, C.J., Todd, J.F.J.: Atmospheric pressure ionization mass spectrometry. *Org. Mass Spectrom.* **18**, 509–516 (1983)
17. Carroll, D.I., Dzidic, I., Horning, E.C., Stillwell, R.N.: Atmospheric pressure ionization mass spectrometry. *Appl. Spectrosc. Rev.* **17**, 337–406 (1981)
18. Whitehouse, C.M., Dreyer, R.N., Yamashita, M., Fenn, J.B.: Electrospray interface for liquid chromatographs and mass spectrometers. *Anal. Chem.* **57**, 675–679 (1985)
19. Klee, S., Derpmann, V., Wißdorf, W., Klopotoski, S., Kersten, H., Brockmann, K.J., Benter, T., Albrecht, S., Bruins, A.P., Dousty, F., Kauppila, T.J., Kostiaimen, R., O'Brien, R., Robb, D.B., Syage, J.A.: Are clusters important in understanding the mechanisms in atmospheric pressure ionization? Part 1: Reagent ion generation and chemical control of ion populations. *J. Am. Soc. Mass Spectrom.* **25**, 1310–1321 (2014)
20. Albrecht, S., Klopotoski, S., Derpmann, V., Klee, S., Brockmann, K.J., Stroth, F., Benter, T.: Studies of the mechanism of the cluster formation in a thermally sampling atmospheric pressure ionization mass spectrometer. *Rev. Sci. Instrum.* **85**, 14102 (2014)
21. Derpmann, V., Albrecht, S., Benter, T.: The role of ion-bound cluster formation in negative ion mass spectrometry. *Rapid Commun. Mass Spectrom.* **26**, 1923–1933
22. Zook, D.R., Grimsrud, E.P.: Detection mass bias in atmospheric pressure ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2**, 232–239 (1991)
23. Kuhl, C., Tautenhahn, R., Böttcher, C., Larson, T.R., Neumann, S.: CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. *Anal. Chem.* **84**, 283–289 (2012)
24. Broeckling, C.D., Afsar, F.A., Neumann, S., Ben-Hur, A., Prenni, J.E.: RAMClust: a novel feature clustering method enables spectral-matching-based annotation for metabolomics data. *Anal. Chem.* **86**, 6812–6817 (2014)
25. Mahieu, N.G., Spalding, J.L., Gelman, S.J., Patti, G.J.: Defining and detecting complex peak relationships in mass spectral data: the Mz.unify algorithm. *Anal. Chem.* **88**, 9037–9046 (2016)
26. Humbert, L., Maubert, M.A., Wolf, C., Duboc, H., Mahé, M., Farabos, D., Seksik, P., Mallet, J.M., Trugnan, G., Masliah, J., Rainteau, D.: Bile acid profiling in human biological samples: comparison of extraction procedures and application to normal and cholestatic patients. *J. Chromatogr. B* **899**, 135–145 (2012)
27. Sarafian, M.H., Lewis, M.R., Pechlivanis, A., Ralphs, S., Mcphail, M.J.W., Patel, V.C., Dumas, M.-E., Holmes, E., Nicholson, J.K.: Bile acid profiling and quantification in biofluids using ultra-performance liquid chromatography tandem mass spectrometry. *Anal. Chem.* **87**(19), 9662–9670 (2015)
28. Wishart, D.S., Jewison, T., Guo, A.C., Wilson, M., Knox, C., Liu, Y., Djoumbou, Y., Mandal, R., Aziat, F., Dong, E., Bouatra, S., Sinelnikov, I., Arndt, D., Xia, J., Liu, P., Yallou, F., Bjorn Dahl, T., Perez-Pineiro, R., Eisner, R., Allen, F., Neveu, V., Greiner, R., Scalbert, A.: HMDB 3.0—the human metabolome database in 2013. *Nucleic Acids Res.* **41**, (2013)
29. Wolff, J.C., Eckers, C., Sage, A.B., Giles, K., Bateman, R.: Accurate mass liquid chromatography/mass spectrometry on quadrupole orthogonal acceleration time-of-flight mass analyzers using switching between separate sample and reference sprays. 2. Applications using the dual-electrospray ion source. *Anal. Chem.* **73**, 2605–2612 (2001)
30. Eckers, C., Wolff, J.-C., Haskins, N.J., Sage, A.B., Giles, K., Bateman, R., Pharmaceuticals, S.B., Frontiers, N., Park, S., Avenue, T., Cm, E.: Accurate mass liquid chromatography / mass spectrometry on orthogonal acceleration time-of-flight mass analyzers using switching between separate sample and reference sprays . 1. Proof of concept. *Anal. Chem.* **72**, 3683–3688 (2000)
31. Plumb, R.S., Johnson, K.A., Rainville, P., Smith, B.W., Wilson, I.D., Castro-Perez, J.M., Nicholson, J.K.: UPLC/MSE; a new approach for generating molecular fragment information for biomarker structure elucidation. *Rapid Commun. Mass Spectrom.* **20**, 1989–1994 (2006)
32. Madala, N.E., Steenkamp, P.A., Piater, L.A., Dubery, I.A.: Collision energy alteration during mass spectrometric acquisition is essential to ensure unbiased metabolomic analysis. In: *Analytical and bioanalytical chemistry*. pp. 367–372 (2012)
33. Rew, R., Davis, G.: NetCDF: An interface for scientific data access. *IEEE Comput. Graph. Appl.* **10**, 76–82 (1990)
34. Benton, H.P., Want, E.J., Ebbels, T.M.D.: Correction of mass calibration gaps in liquid chromatography–mass spectrometry metabolomics data. *Bioinform. Appl. Note* **26**, 2488–2489 (2010)
35. Smith, C.A., Want, E.J., O'maille, G., Abagyan, R., Siuzdak, G.: XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.* **78**(3), 779–787 (2006)

36. Tautenhahn, R., Bottcher, C., Neumann, S.: Highly sensitive feature detection for high resolution LC/MS. *BMC Bioinformatics*. **9**, 504 (2008)
37. R Core Development Team: R: a language and environment for statistical computing, 3.2.1, <https://www.r-project.org/> (2016)
38. Stein, S.E., Scott, D.R.: Optimization and testing of mass-spectral library search algorithms for compound identification. *J. Am. Soc. Mass Spectrom.* **5**, 859–866 (1994)
39. Wan, K.X., Vidavsky, I., Gross, M.L.: Comparing similar spectra: from similarity index to spectral contrast angle. *J. Am. Soc. Mass Spectrom.* **13**, 85–88 (2002)
40. Juraschek, R., Dülcks, T., Karas, M.: Nanoelectrospray – more than just a minimized-flow electrospray ionization source. *J. Am. Soc. Mass Spectrom.* **10**, 300–308 (1999)
41. Schmidt, A., Karas, M., Dülcks, T.: Effect of different solution flow rates on analyte ion signals in nano-ESI MS, or: When does ESI turn into nano-ESI? *J. Am. Soc. Mass Spectrom.* **14**, 492–500 (2003)
42. Tang, K., Page, J.S., Smith, R.D.: Charge competition and the linear dynamic range of detection in electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **15**, 1416–1423 (2004)
43. Alfaro, C.M., Uwakweh, A.O., Todd, D.A., Ehrmann, B.M., Cech, N.B.: Investigations of analyte-specific response saturation and dynamic range limitations in atmospheric pressure ionization mass spectrometry. *Anal. Chem.* **86**, 10639–10645 (2014)
44. Tang, L., Kebarle, P.: Dependence of ion intensity in electrospray mass spectrometry on the concentration of the analytes in the electrosprayed solution. *Anal. Chem.* **65**, 3654–3668 (1993)
45. Tang, K., Smith, R.D.: Physical/chemical separations in the break-up of highly charged droplets from electrosprays. *J. Am. Soc. Mass Spectrom.* **12**, 343–347 (2001)
46. Kebarle, P., Tang, L.: From ions in solution to ions in the gas phase – the mechanism of electrospray mass spectrometry. *Anal. Chem.* **65**, 972A–986A (1993)
47. Gabelica, V., De Pauw, E.: Internal energy and fragmentation of ions produced in electrospray sources. *Mass Spectrom. Rev.* **24**, 566–587 (2005)
48. Brophy, P., Farmer, D.K.: Clustering, methodology, and mechanistic insights into acetate chemical ionization using high-resolution time-of-flight mass spectrometry. *Atmos. Meas. Tech.* **9**, 3969–3986 (2016)
49. Dziadosz, M.: Influence of buffer concentration on electrospray ionisation of  $\gamma$ -hydroxybutyrate adducts with the components of the mobile phase used in liquid chromatography-tandem mass spectrometry. *J. Chromatogr. B*. **1008**, 240–241 (2016)
50. Kamel, A.M., Brown, P.R., Munson, B.: Effects of mobile-phase additives, solution pH, ionization constant, and analyte concentration on the sensitivities and electrospray ionization mass spectra of nucleoside antiviral agents. *Anal. Chem.* **71**, 5481–5492 (1999)
51. Jaeger, C., Méret, M., Schmitt, C.A., Liseč, J.: Compound annotation in liquid chromatography/high-resolution mass spectrometry based metabolomics: robust adduct ion determination as a prerequisite to structure prediction in electrospray ionization mass spectra. *Rapid Commun. Mass Spectrom.* **31**, 1261–1266 (2017)