

Accelerating Cross-Modality Reaction Optimization via Robotically Automated Vacuum Enabled Direct-inject Mass Spectrometry (RAVE MS)

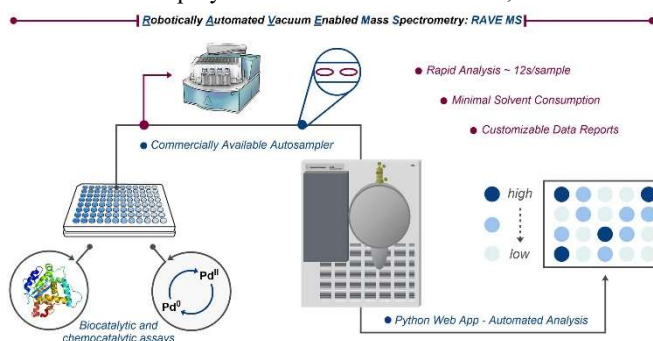
Daniel A. Holland-Moritz^{1*}, Sarah R. Moor^{1*}, Joseph B. Parry¹, Elliot J. Medcalf³, Claire M. Eberle¹, Andrew C. Strakham², Shane T. Grosser¹, Hang Hu¹, Noah P. Dunham¹, Maximillian Gantz³

1. Process Research & Development, Merck & Co., Inc., Rahway, New Jersey 07065, United States

2. Analytical Research and Development, Merck & Co., Inc., Rahway, New Jersey 07065, United States

3. Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, CB2 1GA Cambridge, United Kingdom

ABSTRACT: In this report, we detail direct inject mass spectrometry via a Robotically Automated Vacuum Enabled (RAVE) interface that utilizes commercially available capillary electrophoresis hardware to directly inject samples for mass spectrometry (MS) at a sampling rate of approximately 12s per sample. This system enables direct electrospray ionization from standard 48, 96 or 384-well plates with minimal investment in hardware and software, utilizing a custom developed web app interface that provides autosampler control and analysis of raw extracted data from the mass spectrometer. We show a high level of correlation among results obtained with RAVE coupled MS, acoustic ejection (Echo) MS, and liquid chromatography coupled MS (LCMS) on 384 biocatalytically driven reactions. We additionally utilize RAVE MS on an array of 96 chemocatalytic reaction conditions to show that, while direct MS analysis can be challenging in complex mixtures, simple dilution followed by direct injection is often sufficient for analysis. With these results, we demonstrate the potential for RAVE MS to be utilized as a low-cost, low barrier to entry tool for rapid direct-inject MS analysis.



Introduction

Successful process development in the pharmaceutical industry requires researchers to design, build, and test thousands of chemical reactions, optimizing for catalyst and reagent selection, concentration, and reaction conditions across a broad spectrum of chemical and engineering design space. The high volume of experiments conducted and the size of the available design space have driven the necessity for data-rich experimentation, especially, high throughput screening tools, to keep up with the experimental demands of process development scientists.¹⁻³ High throughput screening (HTS) tools, such as robotic automation and plate based reaction parallelization, are key enabling technologies in the exploration of this extensive chemical space, and the rising ubiquity of automation in industrial labs has begun to shift the limiting step in HTS from sample generation to sample analysis. Chemocatalytic and biocatalytic syntheses are examples of fields where the generation of samples often outstrips the capacity of current analytical tools, and while robotic automation has significantly accelerated the process of generating these chemical and biological samples, sample analytics remain a key bottleneck.⁴

A growing body of research and development has focused on developing faster analytical methodologies to match

the pace with which samples are generated.^{5, 6} In screening workflows where photometric methods of detection are available and unambiguous, spectroscopy remains the dominant technology for rapid analytics in HTS.⁷⁻⁹ However, when assays are either too complex or challenging to read directly with fluorescence or absorption (as is often the case in pharmaceutical chemistry), liquid chromatography (LC) predominates as the analytical tool of choice.¹⁰ Unfortunately, chromatographic separations greatly increase the time required to read samples generated in HTS, as each reaction must be sampled, injected, eluted, and the column regenerated for the next analytical run to begin. When scaled across hundreds or thousands of reactions, this analytical burden becomes a critical bottleneck, and a growing need for selective, rapid, and direct analysis of complex reactions has become evident.¹¹

Direct analysis by mass spectrometry has emerged as a key technique that bridges the gap between the higher analytical throughput of spectroscopic interrogation and the richness of analytical data provided by chromatographic separations.¹²⁻¹⁴ The growing capabilities of mass spectrometry to both isolate and monitor individual species in complex reaction mixtures enable direct sample analysis without the prerequisite of

chromatographic separation.¹⁵⁻¹⁷ The speed with which direct MS analysis may be implemented has in turn necessitated the development of new methods for sample introduction directly to the MS source.^{13, 18, 19}

Rapid sample introduction to mass spectrometry has taken shape in two major formats – analysis via direct sampling (e.g. Echo MS), and analysis via reformatting into an MS-interfaceable format (e.g. a MALDI plate or a mesh for Direct Analysis in Real Time).²⁰⁻²⁹ In the latter methodology, sample reformatting often becomes the bottleneck, and utilization is hindered by the speed and reliability of the reformatting process, as well as the complexity of sample preparation necessary to interrogate reactions.^{26, 29, 30} These tools are also costly analytical options due to the distinctive (and often, non-standard) methods by which they achieve rapid ionization of their reformatting samples requiring specialized ionization sources and expensive additional hardware.

The alternative to reformatting samples for rapid MS analysis typically is the direct injection of samples into electrospray ionization sources. The simplest way to automate this for HTS is to take advantage of existing auto sampling systems to perform serial direct injections.²⁵ The drawbacks to this approach are sample-to-sample carryover, sample matrix effects, and the low throughput at which a traditional LC autosampler can send liquid to the analyzer and reset for the next injection (1-2 min).^{31, 32} Matrix effects can be reduced or eliminated using online solid phase extraction at the front end of a direct inject system, and employment of an autosampler optimized for direct inject workflows significantly reduces the timescales of direct-inject MS. The Agilent RapidFire MS is a commercialized platform for achieving these goals.³³ The Echo MS (commercialized by Sciex) can be used to analyze samples as fast as 1/s by coupling heavy dilution in an open port interface to acoustic ejection from a well plate to deliver sample to a continuous dilution stream.^{34, 35} Due to the low carryover achieved by acoustic ejection, the minimization of matrix effects via dilution, and the rapid sample introduction, Echo MS has rapidly become the preferred method for direct injection mass spectrometry.³⁰ ³⁶Despite these benefits, the cost of both the RapidFire and Echo MS systems is high, rendering it beyond the reach of many academic and small-scale laboratories and necessitating the development of more accessible alternatives to achieve similar efficiency.

In this work, we describe the principles of operation for Robotically Automated Vacuum Enabled (RAVE) MS, a simple direct injection mass spectrometry tool. We detail its experimental usage in both bio- and chemo-catalysis screening and demonstrate its utility as a rapid and low-cost tool with a low barrier to entry. We directly compare the operation of RAVE MS on a single quadrupole mass spectrometer to the performance of commercial direct-inject tools such as Echo MS, demonstrating strong agreement on a flexible, standard, and accessible platform. RAVE MS enables rapid qualitative assessment of relative reaction yields, and throughputs that, prior to this point, came only with significant capital investment for most laboratories. This workflow can be implemented using commercially available ionization sources and autosamplers and is operated using open-source Python libraries. We share here two use cases for RAVE MS and envision additional utilization across a wide range of MS-enabled reaction monitoring and screening applications.

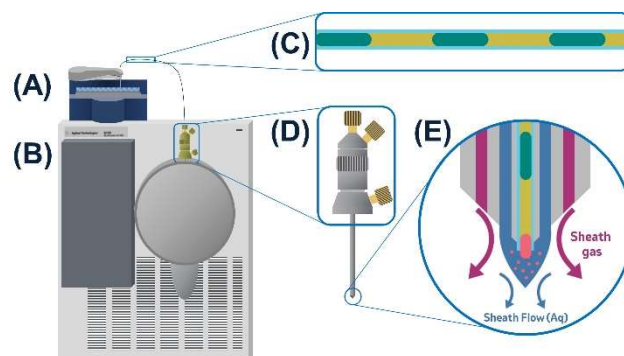


Figure 1. The key components of RAVE MS. An autosampler (A) moves a capillary from well to well on a plate, flushing with fluorinated oil after each sample. These samples are drawn into the mass spectrometer (B) as a series of oil spaced plugs (C). The suction necessary to draw samples into the system is generated at the source in a CE-MS triple tube needle (D), where the high flow rate of the sheath gas exiting the needle drives a localized reduction in pressure at the needle tip (E).

RAVE MS Principles of Operation:

RAVE MS leverages the hydrodynamic draw in a commercial capillary electrophoresis (CE-ESI) source to directly aspirate samples from a well plate, without the use of pumps, valves, or flow control. All sampling is driven entirely by the vacuum effect produced at the tip of the source within the MS itself, allowing sample plugs to be aspirated into the capillary by dipping the free end into the solution to be sprayed and subsequently analyzed. This can be performed manually, or, ideally, with the assistance of a robotic autosampler (Figure 1A).

In our work, we utilized Agilent's CE-ESI (G1607-60002) triple tube source (Figure 1D) that utilizes a two layered co-axial flow of sheath buffer and drying gas around an inserted capillary (Figure 1E) to achieve electrospray ionization.³⁷ The sheath flow ensures a steady stream of liquid for Taylor Cone formation at the emitter tip, irrespective of the relatively low flow rates typical of capillary electrophoretic flow.³⁸ The combination of liquid sheath flow and heated drying gas has been previously reported to exert a hydrodynamic force on the liquid within CE capillaries, an effect postulated to be due to the Bernoulli effect caused by the high velocity of gas flow at the source tip.³⁹ This hydrodynamic force drives samples into the mass spectrometer through vacuum suction, enabling continuous and reliable infusion of samples.

Experimental Section

Sampling Capillary Setup

To adapt the CE-ESI source for RAVE MS, the typical glass capillaries used in capillary electrophoresis were replaced with a 150µm ID, 360µm OD high purity perfluoroalkoxy (HPFA) capillary (Irex Health Sciences, Oak Harbor, WA). This change was made to reduce the surface affinity of samples with the capillary and minimize sample-to-sample carryover. We utilized a wash with a fluorinated oil to separate samples, prime the capillary, initiate flow, and clear the capillary after each sample (SI Figure 6, Capillary Priming).⁴⁰ The fluorinated oil (Novec7500) is both volatile and omniphobic, which enables it to keep both aqueous and organic dissolved samples from mixing as they are carried to the ion source while

also readily evaporating in the stream of heated gas at the source.

The capillary was cut to a length of ≤ 65 cm to minimize capillary back pressure and the resistance of the system to flow when filled with sample. The robotic autosampler was placed on top of the mass spectrometer, as close to the source as possible (Figure 1A). The capillary was threaded through the sampling needle on the Sielc Alltesta autosampler and held in place by a capillary fitting in the modified robot arm (SI Figure 5). The robot was then used to immerse the sampling needle with the protruding end of the capillary first in fluorinated oil to provide sample separation, and then in sample, alternating between the two liquids during operation.

Mass Spectrometer

The commercial mass spectrometer used in this work was an Agilent single quadrupole (6120); we have achieved successful RAVE MS sampling in both triple quadrupole and Q-TOF instruments as well. The sheath gas temperature and flow rate have been reported to be important contributors to the hydrodynamic forces experienced by the inserted capillary in these types of triple tube sprayers.³⁷ In our experiments, we operated with the highest temperature (350 °C), drying gas flow rate (13 L/min), and nebulizer pressure (60 psig) available to us with the aim of maximizing these effects, and we did not optimize or modify these run-to-run. In early exploration, we observed that the nebulizer pressure had a visible effect on the flow rate for samples injected onto the mass spectrometer. The capillary ionization voltage was set at 4000V and could likely be optimized from analyte to analyte to tune the response, although we did not do so in this work. LC grade water, infused at 30 μ L/min was used for the sheath flow, as had been previously described in droplet electrospray work with the triple tube source.^{40, 41}

RAVE Single Ion Monitoring Operation

RAVE auto sampling was performed by dipping the free end of the sampling capillary into each well to allow a brief infusion of sample into the source. The Alltesta autosampler used in this work was operated using custom software written in Python and incorporated into a web app interface (SI Figures 1-4). In order to minimize run time, dwell time in the bottom of the sample wells was intentionally limited to 1 s, and most of the 12 s/sample cycle time of RAVE was a function of the time it took the autosampler arm to rise from the well, move to the oil reservoir, sample it, and return to the next well. Under steady operation with 50 μ L of sample in a well, we observed samples were sprayed for only ~ 1.6 s. We initially chose to read samples using full scans but found that extracted-ion-chromatograms from full scan calibrations exhibited reduced sensitivity relative to single ion monitoring (SIM) of those same analytes (SI Figure 8).⁴² SIM maximized the time our instrument spent collecting signal for the ions of interest and helped ensure sufficient signal from the components most relevant to reaction monitoring.

Additionally, the lack of separation on the front end of the analysis in RAVE made it important to carefully account for differences in reaction composition across samples, which affected the ionization efficiency and suppression of the target analyte.^{43, 44} Samples with significantly different background

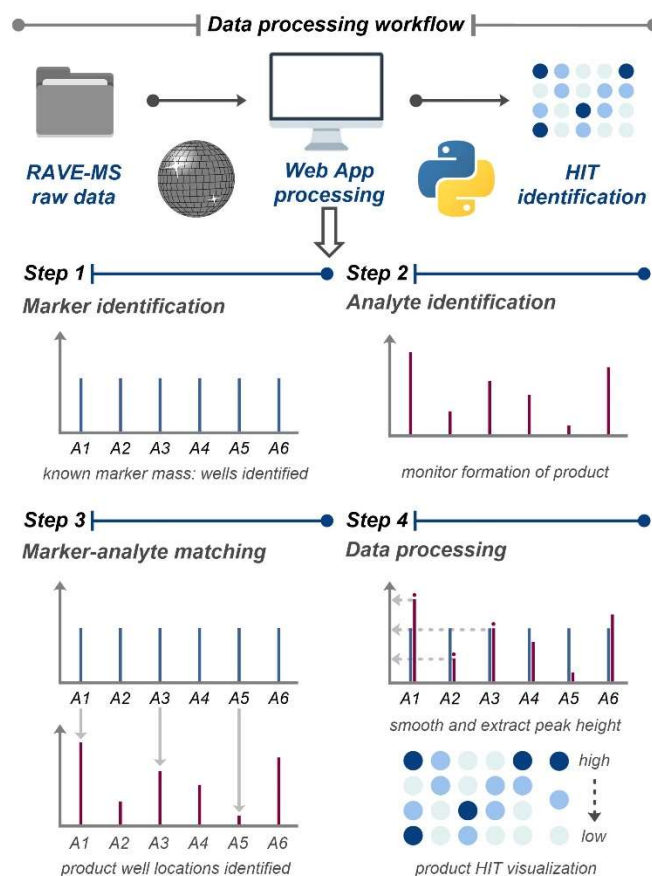


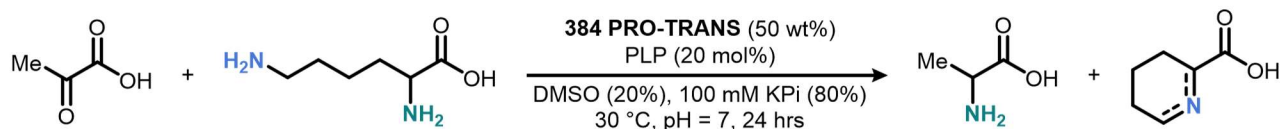
Figure 12. RAVE MS data extraction and analysis is performed by Python code integrated into a web app interface. Step 1: A marker analyte is chosen to aid in the delineation of start and end of each injection. Step 2: the analytes of interest are concurrently monitored via SIM. Subsequently, these raw SIM traces from the mass spectrometer are directly uploaded to the RAVE app which utilizes signals from the marker ion to establish the retention time for each peak and correlate it to a well on a plate. Step 3: these retention times are then applied to each monitored ion to extract the peak max and average for each analyte and report out relative intensities for the samples monitored. Step 4: after the marker-analyte matching, the app performs a series of data processing steps to allow the user to visualize and export the data.

matrices were therefore challenging to compare directly to one another due to the differences in ionization efficiency across samples. We resolved this by tuning our dilution factor for each sample matrix screened.

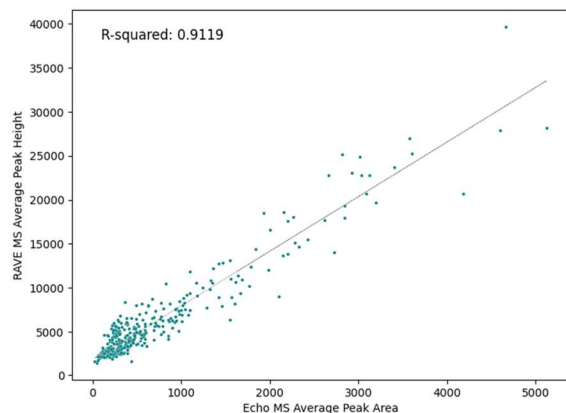
Data Extraction and Analysis

For all examples in this work, RAVE MS was run in single ion monitoring mode on the mass spectrometer, with up to four possible analytes monitored for any given reaction. The total number of analytes monitored was primarily limited by the Agilent 6120 SIM capabilities. Upon screen completion, the data was extracted as raw signal counts over time and analyzed using our custom Python application. Whilst one can use Chemstation to extract peak height and area from chromatograms, the variable retention times from screen to screen (a by-product of the RAVE auto sampling and injection operating

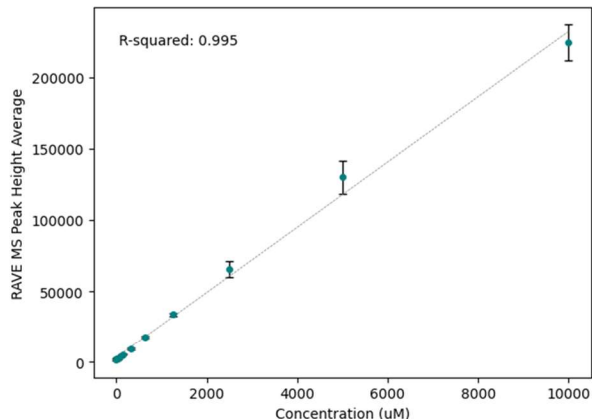
Biocatalysis: Transaminase Screen



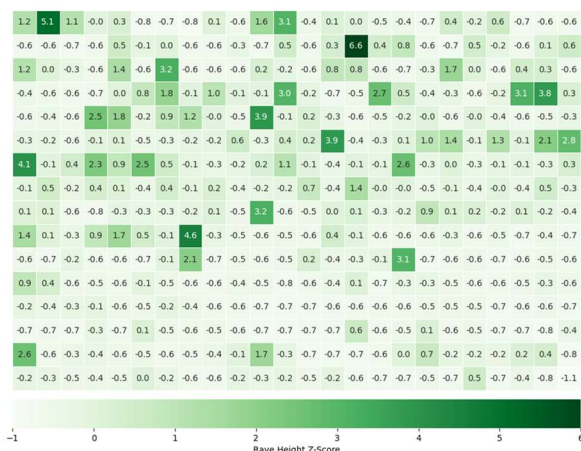
(A) Echo MS vs RAVE MS Parity Plot



(B) RAVE MS Alanine Calibration



(C) RAVE MS Normalized Height



(D) Echo MS Normalized Area

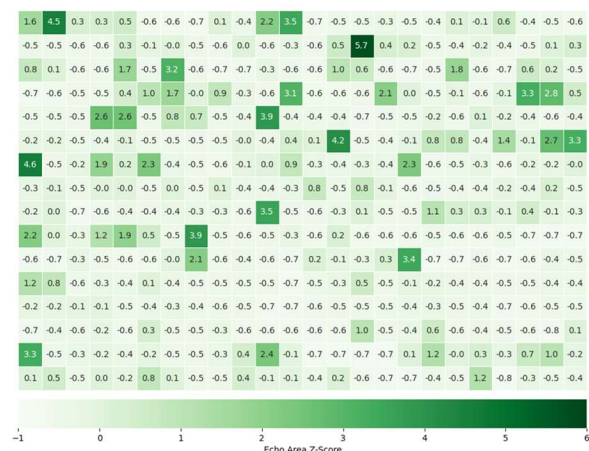


Figure 3. Implementation of RAVE MS in the HTE screening of a transaminase reaction. (A) Correlation shows strong agreement between RAVE MS peak height and Echo MS peak area data for Alanine measurements in the transamination reaction. (B) An alanine calibration plot created using RAVE MS demonstrating linear response to alanine in the transaminase quench buffer. (C) Heat map illustrating the Z-score normalized product peak height extracted from RAVE MS for the 384 screened reaction conditions. (D) Heat map illustrating the Z-score normalized product peak area extracted from Echo MS for the 384 screened reaction conditions.

independently from the MS spectrum initialization) made it difficult to establish a methodology for extracting and processing screen data. We therefore developed our own data analysis methodology for raw data processing.

To analyze RAVE data, we first extracted the SIM trace on all utilized channels. This produced up to four files with a time and an intensity read for the entirety of the run. To detect all peaks in each run, it was critical that at least one of these four files showed peaks for every sample injection to ensure that wells without target analyte present were accounted for. This signal trace, which we referred to as the marker trace, could be extracted from a read of the substrate signal because all reactions contained this analyte. However, in instances where substrate signal proved challenging to monitor,

background ions from the remaining matrix of injected samples or internal standards in the dilution matrix could be used instead to monitor the injection time for each sample. With the help of this marker trace, well sampling times could be identified and applied to the other three extracted spectra to measure target analyte signal intensity during screen analysis (Figure 2 Steps 1 & 2).

Using custom-built software incorporated into the RAVE MS autosampler control user interface, we analyzed extracted files for each RAVE run. The marker trace was uploaded to the application, where the user could adjust the peak width, minimum intensity, plateau size, and the minimum distance between peaks (SI Figure 4). These arguments were then used to identify each of the peaks from a screen by utilizing the

SciPy peak picking algorithm.⁴⁵ To aid in peak identification, the data could be smoothed prior to peak picking by calculating the moving average of the peaks or applying a Savitzky-Golay filter.⁴⁶ This generated a list of retention times that were then applied to each of the other uploaded traces to mark the location of each sample. The average ion intensity in the region identified from the marker file was then used as the reported value for the analyte of interest at each marked retention time, generating a list of values for each sample. These values were then used directly as a measure of relative intensity (Figure 2 Steps 3 & 4).

It is important to note that, due to the on/off nature of sampling in RAVE MS, peak area alone was a poor measure of analyte response because injection volumes were controlled only through the hydrodynamic pull to the source, and the volumes of RAVE injection can therefore vary from sample to sample. A larger injection volume will spray for longer than a smaller one, and raw peak areas will therefore unintentionally bias the response towards larger injections. We addressed this by both reporting an average peak height for each peak (corrected for peak width at half maximum) and a maximum. We have found that both correlated well to the relative performance of individual reactions (SI Figure 7).

Results and Discussion

Case Study 1: RAVE MS for Transaminase Panel Screening for Lysine Amine Donor Tolerance

Transaminases are pyridoxal 5'-phosphate (PLP)-dependent enzymes that catalyze the transfer of an amine from a donor substrate to a ketone or aldehyde receptor. This class of enzymes is synthetically valuable for its ability to catalyze production of enantioenriched primary amines under mild conditions. In these reactions, the nitrogen in the generated amine is derived from the donor molecule. For example, in the commercial production of sitagliptin, isopropylamine is used as the donor, leading to the formation of acetone as a byproduct. The directed evolution of transaminases for industrially relevant chemistry has been well-demonstrated, most notably in the commercial production of sitagliptin.⁴⁷ These directed evolution efforts frequently involve the screening of thousands of enzyme variants for improved activity on their non-native substrates, and thus rapid analytical tools like direct-inject mass spectrometry can vastly reduce the analytical burdens associated with these screens.

Lysine has recently been reported as a smart amine donor for transaminase reactions due to the thermodynamically favored cyclization of the 6-amino-2-oxohexanoic acid side product to the corresponding piperidine-2-carboxylic acid imine.⁴⁸ In their original publication, our colleagues utilized LCMS to assess a panel of 384 commercially available wild-type enzymes for their compatibility with lysine as an amine donor. Their work looked at the transamination of pyruvate to alanine to assess the preference of each panel member for L- and D-lysine in the reaction. This initial analytical endeavor required >27 h of LCMS run time, even with an optimized LC method less than 4 min in length. The stereoselectivity of each panel member was also assessed by running the reaction separately with a feed of L- and D-Lysine, doubling the analytical burden. We therefore set out to demonstrate the utility of

RAVE MS for the same screen to reduce analytical run time and achieve comparative quantitative analysis.

We repeated the work of our colleagues with L-alanine to generate samples that could be directly compared across analytical platforms. Reactions were performed (SI Section 4b) in a total solvent volume of 50 μ L within a 384 well plate pre-plated with 1 mg of fermentation powder from each enzyme variant. After a 24-h reaction time, the plate was sampled and diluted (5% v/v) into a quench solution of 1:1 MeCN/100 mM potassium phosphate buffered to pH 7.0. Quenched samples were centrifuged to spin down any particulate and reduce the opportunity for capillary clogging during RAVE MS analysis.

To rapidly assess the analytical sensitivity of RAVE MS to alanine and pyruvate concentrations, stock solutions in the quench and diluent solution were created. These were diluted further in MeCN/potassium phosphate buffer to create calibration curves for alanine and pyruvate that approximate the background matrix of the panel samples (Figure 3A). Prior to obtaining the calibration curve, direct injections of both alanine and pyruvate (1.25mM) were read in full scan mode from 50-160 m/z in positive and negative ion mode to determine the optimal SIM trace to monitor. These experiments showed that negative ion mode monitoring resulted in clear signal for the target analytes while positive mode did not (SI – SIM target Selection). Given this, negative ion mode was used to monitor alanine production (88 M/Z) and pyruvate (87 M/Z) consumption.

The calibration plot for alanine can be seen in Figure 3B. A strong linear correlation ($R^2 = 0.995$) between RAVE MS peak height and concentration was obtained, indicating that average peak height can be used to accurately compare analyte concentrations across samples. Limits of detection for alanine and pyruvate in RAVE injections were found to be 99 μ M and 68 μ M, respectively (SI Figure 11).

Following calibration, we analyzed the 384 well plate of transaminases using RAVE MS, LCMS, and Echo MS. After completion of RAVE MS analysis, peak heights for each monitored ion were extracted using our custom data analysis software. The extracted RAVE MS peak height averages and Echo MS peak areas are plotted in Figure 3A and exhibit a strong correlation ($R^2 = 0.912$). Correlations between RAVE MS and LCMS were also in alignment ($R^2 = 0.922$, SI Figure 10), further validating the RAVE MS assay for HTS and hit selection. The correlation between pyruvate signal among the three analytical methods was similarly robust, with R^2 values above 0.970 for all comparisons (SI Figure 10). These observations highlight the effectiveness of RAVE MS as a rapid and reliable analytical tool, offering significant time savings and strong agreement with established mass spectrometry techniques and enabling high-throughput screening in biochemical assays and providing an exciting path forward to address a critical bottleneck in evaluating the many unique engineered enzymes accessible through directed evolution.

Case Study 2: RAVE MS for Rapid Assessment of Pd-catalyzed C-N Cross-Coupling Conditions

C-N bond formation is of fundamental importance within organic synthesis, as showcased by its widespread application toward the synthesis of pharmaceuticals, agrochemicals and privileged ligand systems.⁴⁹ A diverse array of transition metals, including nickel, copper, ruthenium, and palladium

Chemo-Catalysis: Buchwald-Hartwig C-N Coupling

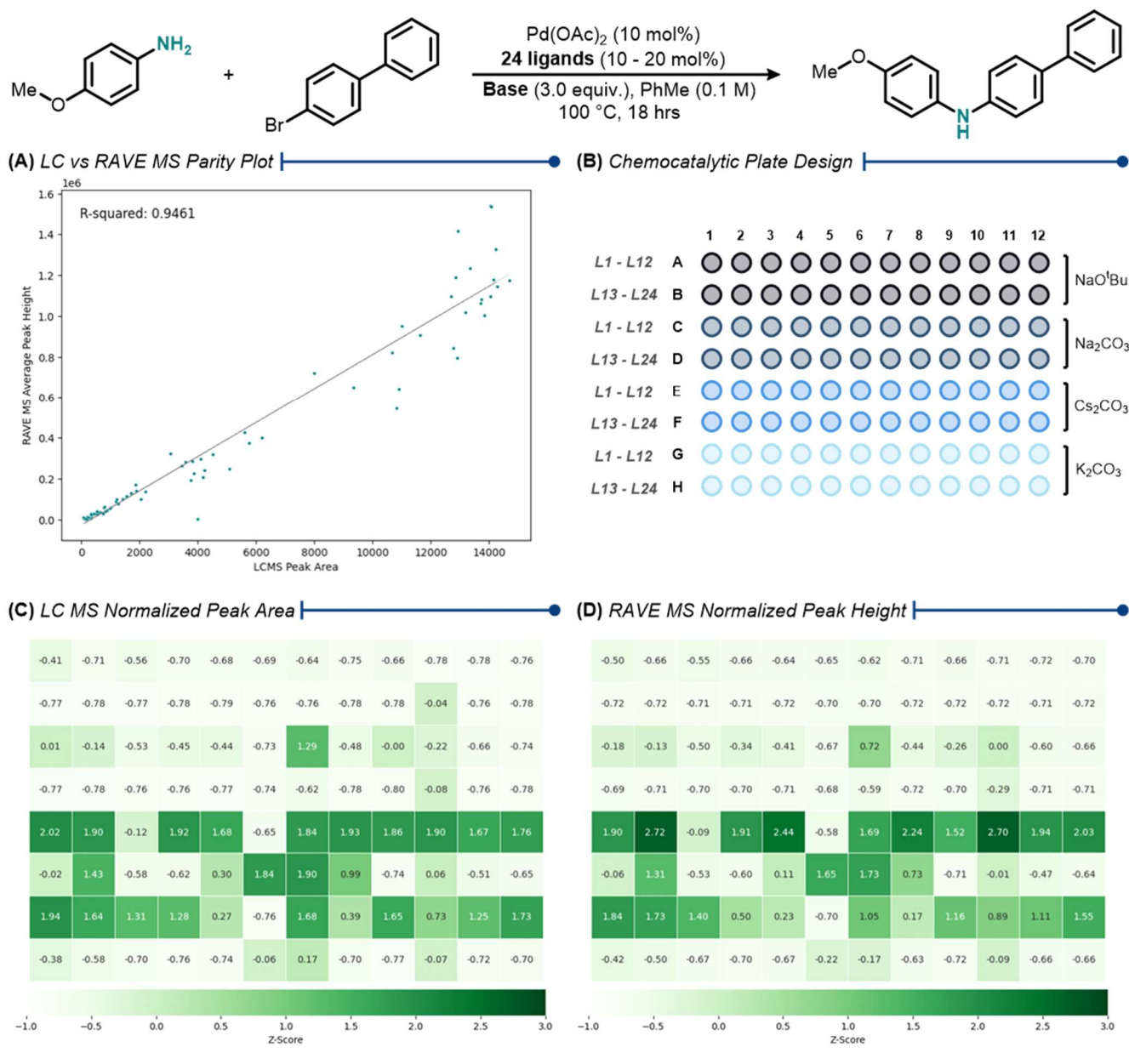


Figure 4. Implementation of RAVE MS in the HTE screening of a Pd-catalyzed Buchwald-Hartwig coupling. (A) Correlation between RAVE MS peak height and LCMS peak area data. (B) Distribution of ligands and inorganic bases across a 96-well reaction plate. (C) Heat map illustrating the Z-score normalized product peak height extracted from RAVE MS for the 96 screened reaction conditions. (D) Heat map illustrating the Z-score normalized product peak area extracted from LCMS for the 96 screened reaction conditions.

have been used to catalyze C-N bond formation.⁵⁰⁻⁵⁴ Notably, the Buchwald-Hartwig coupling, which involves the coupling of an aryl halide or pseudo halide with a suitable nitrogen-based nucleophile in the presence of Pd/ligand and a base, is well-established as the preferred reaction methodology to forge C-N bonds.⁵⁵

In this investigation, we used RAVE MS to assess the performance of 24 phosphine ligands, encompassing both monodentate and bidentate scaffolds, in combination with 4 bases, to identify the optimal reaction conditions for the C-N cross-coupling of 4-methoxyaniline and 4-bromo-1,1'-bi-phenyl. Our aim was to demonstrate utilization of RAVE MS to rapidly triage reactions for catalyst/ligand selection and

optimize the reaction with minimal method development and analytical run time.

24 pre-plated phosphine ligands were combined with 4 pre-plated bases (NaO^tBu, Na₂CO₃, Cs₂CO₃ and K₂CO₃), resulting in the generation of 96 reaction conditions in a plate format, as depicted in Figure 4B. Following the completion of the reactions (see SI, Pd-Catalyzed Buchwald-Hartwig coupling), two plates were prepared for analysis: one for RAVE MS and the other for LCMS, serving as an orthogonal analytical approach for the validation of our RAVE MS results. Analytical plates were prepared by diluting reaction plates with a solution of internal standard. For analysis via RAVE MS, the peak heights corresponding to the molecular ions formed by the

product ($[M+H]^+ = 276.14$) and lysine ($[M+H]^+ = 147.11$) internal standard from the dilution were tracked in positive single-ion monitoring mode. For LCMS analysis, the extracted ion chromatogram for the target product was integrated over its elution.

To determine the most suitable concentration range for plate dilution and to perform similar calibration studies to those described in our first case study, we synthesized an authentic product standard and subjected it to analysis via RAVE MS in the expected dilution matrix. At concentrations between 0 – 60 μM , we observed a linear response (SI Figure 18). We subsequently diluted our reaction samples to a concentration of 25 μM product at 100% conversion and compared the RAVE response to LCMS conducted at an analyte concentration of 125 μM at 100% conversion. A strong linear correlation of $R^2 = 0.9461$ was achieved (Figure 4A).

The results obtained from both assays were visualized as 96 well heatmaps to evaluate the performance of the screen and to determine any assay-to-assay variability. Prior to visualization, the product peak height and peak area were Z-Score normalized to facilitate cross assay comparison. Reactions analyzed on RAVE MS show high levels of product formation when Cs_2CO_3 is used as the base in concert with the following ligands: XPhos (L1), SPhos (L2), BrettPhos (L7), CataCXium A (L8) and $[(^t\text{Bu})_3\text{P}][\text{HBF}_4]$ (L10), which is echoed in the LCMS data. This indicates that RAVE MS is suitable for hit identification, even in systems that feature a high degree of variation in reaction matrix composition.

Analysis of the 96 well plate took > 570 min to complete when using a 6-min per-well LCMS method. In contrast, analysis using RAVE MS took only 3.5% of that analytical time at only 20 min per plate. We found that the screening data provided by RAVE MS enabled facile identification of reaction conditions of interest for further investigation. The rapid and efficient analysis provided by RAVE MS on these chemocatalytic reactions underscores its potential as an accessible analytical tool across chemistries for high-throughput screening, with its impact only likely to increase with the size of screen conducted.

CONCLUSIONS

Herein, we introduced RAVE MS as a simple and easy-to-implement tool for high throughput direct-inject mass spectrometry. RAVE MS takes advantage of the hydrodynamic forces generated by a CE-MS electrospray emitter to pull samples directly into a mass spectrometer for analysis. To automate data acquisition and analysis, we developed a Python web app interface that allows for autosampler control as well as raw data processing and visualization to accelerate hit identification from high-throughput screening. Following this workflow, we applied the RAVE MS assay to screen enzyme variant libraries for a biocatalytic transaminase reaction and in reaction condition screening for a Pd-catalyzed Buchwald-Hartwig coupling. In both cases, the data obtained validate the use of RAVE MS as a simple, rapid analytical tool that is compatible with plate-based chemistries, with the demonstrated potential to reduce screening times by as much as 28-fold for our case study screens. RAVE MS-generated data was in excellent agreement

with state-of-the-art direct inject analytical platforms such as Echo MS and orthogonal methods like LCMS at a fraction of the total operating cost, rendering it an attractive and accessible analytical tool for use across academia and industry.

The analytical pipeline described in this work outlines the development and implementation of RAVE MS as an attractive platform that can be readily incorporated into high-throughput screening assays. At our company, we have already begun to use it to accelerate hit identification within the process research and development space – deploying it for rapid screening on tight timelines to power our complex protein engineering portfolio. With the expansion into chemocatalytic screening, we anticipate further proliferation across small molecule process development in the coming years.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Supplemental Information (PDF)

Protocol for RAVE setup (PDF)

AUTHOR INFORMATION

Corresponding Authors

Daniel Holland-Moritz - Process Research & Development, Merck & Co., Inc., Rahway, New Jersey 07065, United States; orcid.org/0000-0003-4634-6122; Email: holladan@merck.com

Sarah R. Moor - Process Research & Development, Merck & Co., Inc., Rahway, New Jersey 07065, United States; orcid.org/0000-0001-5579-407X; Email: sarah.moor@merck.com

Authors

Joseph Parry - Process Research & Development, Merck & Co., Inc., Rahway, New Jersey 07065, United States; orcid.org/0000-0001-5979-6814; Email: joseph.parry@merck.com

Elliot Medcalf - Department of Biochemistry, University of Cambridge, Cambridge, United Kingdom; orcid.org/0009-0004-3124-0346; Email: ejm208@cam.ac.uk

Claire Eberle - Process Research & Development, Merck & Co., Inc., Rahway, New Jersey 07065, United States; orcid.org/0009-0005-3335-5259; Email: claire.eberle@merck.com

Andrew Strakham - Analytical Research and Development, Merck & Co., Inc., Rahway, New Jersey 07065, United States; orcid.org/0009-0001-1890-3457; Email: andrew.strakham@merck.com

Shane T. Grosser - Process Research & Development, Merck & Co., Inc., Rahway, New Jersey 07065, United States; orcid.org/0000-0002-9268-7545; Email: shane_grosser@merck.com

Hang Hu - Analytical Research and Development, Merck & Co., Inc., Rahway, New Jersey 07065, United States; orcid.org/0000-0002-8590-9787; Email: hang.hu@merck.com

Noah Dunham - Process Research & Development, Merck & Co., Inc., Rahway, New Jersey 07065, United States; orcid.org/0000-0001-8006-9566; Email: noah.dunham@merck.com

Notes

The authors declare no competing financial interest.

REFERENCES

- (1) Mennen, S. M.; Alhambra, C.; Allen, C. L.; Barberis, M.; Berritt, S.; Brandt, T. A.; Campbell, A. D.; Castañón, J.; Cherney, A. H.; Christensen, M.; et al. The Evolution of High-Throughput Experimentation in Pharmaceutical Development and Perspectives on the Future. *Organic Process Research & Development* **2019**, *23* (6), 1213-1242. DOI: 10.1021/acs.oprd.9b00140.
- (2) Shevlin, M. Practical High-Throughput Experimentation for Chemists. *ACS Medicinal Chemistry Letters* **2017**, *8* (6), 601-607. DOI: 10.1021/acsmchemlett.7b00165.
- (3) Broach, J. R.; Thorner, J. High-throughput screening for drug discovery. *Nature* **1996**, *384* (6604 Suppl), 14-16. From NLM.
- (4) Grainger, R.; Whibley, S. A Perspective on the Analytical Challenges Encountered in High-Throughput Experimentation. *Organic Process Research & Development* **2021**, *25* (3), 354-364. DOI: 10.1021/acs.oprd.0c00463.
- (5) Mayr, L. M.; Fuerst, P. The future of high-throughput screening. *J Biomol Screen* **2008**, *13* (6), 443-448. DOI: 10.1177/1087057108319644 From NLM.
- (6) Pereira, D. A.; Williams, J. A. Origin and evolution of high throughput screening. *British Journal of Pharmacology* **2007**, *152* (1), 53-61. DOI: 10.1038/sj.bjp.0707373.
- (7) Moor, S. R.; Howard, J. R.; Herrera, B. T.; Anslyn, E. V. High-throughput screening of α -chiral-primary amines to determine yield and enantiomeric excess. *Tetrahedron* **2021**, *94*, 132315. DOI: <https://doi.org/10.1016/j.tet.2021.132315>.
- (8) Blay, V.; Tolani, B.; Ho, S. P.; Arkin, M. R. High-Throughput Screening: today's biochemical and cell-based approaches. *Drug Discovery Today* **2020**, *25* (10), 1807-1821. DOI: <https://doi.org/10.1016/j.drudis.2020.07.024>.
- (9) Eggeling, C.; Brand, L.; Ullmann, D.; Jäger, S. Highly sensitive fluorescence detection technology currently available for HTS. *Drug Discovery Today* **2003**, *8* (14), 632-641. DOI: [https://doi.org/10.1016/S1359-6446\(03\)02752-1](https://doi.org/10.1016/S1359-6446(03)02752-1).
- (10) Sheng, H.; Kim, D.; Chin, A. S.; Zhao, Y.; Liu, Y.; Katwaru, R.; Bateman, K. P.; Abend, A.; Wuelfing, W. P. Development of an automated and High throughput UHPLC/MS based workflow for cleaning verification of potent compounds in the pharmaceutical manufacturing environment. *Journal of Pharmaceutical and Biomedical Analysis* **2020**, *188*, 113401. DOI: <https://doi.org/10.1016/j.jpba.2020.113401>.
- (11) Carr, P. W.; Wang, X.; Stoll, D. R. Effect of Pressure, Particle Size, and Time on Optimizing Performance in Liquid Chromatography. *Analytical Chemistry* **2009**, *81* (13), 5342-5353. DOI: 10.1021/ac9001244.
- (12) Dueñas, M. E.; Peltier-Heap, R. E.; Leveridge, M.; Annan, R. S.; Büttner, F. H.; Trost, M. Advances in high-throughput mass spectrometry in drug discovery. *EMBO Molecular Medicine* **2023**, *15* (1). DOI: 10.15252/emmm.202114850.
- (13) Pu, F.; Elsen, N. L.; Williams, J. D. Emerging Chromatography-Free High-Throughput Mass Spectrometry Technologies for Generating Hits and Leads. *ACS Medicinal Chemistry Letters* **2020**, *11* (11), 2108-2113. DOI: 10.1021/acsmchemlett.0c00314.
- (14) Bensch, M.; Schulze Wierling, P.; von Lieres, E.; Hubbuch, J. High Throughput Screening of Chromatographic Phases for Rapid Process Development. *Chemical Engineering & Technology* **2005**, *28* (11), 1274-1284. DOI: <https://doi.org/10.1002/ceat.200500153>.
- (15) Li, X.; Jiang, G.; Luo, C.; Xu, F.; Wang, Y.; Ding, L.; Ding, C.-F. Ion Trap Array Mass Analyzer: Structure and Performance. *Analytical Chemistry* **2009**, *81* (12), 4840-4846. DOI: 10.1021/ac900478e.
- (16) Barnidge, D. R.; Dasari, S.; Botz, C. M.; Murray, D. H.; Snyder, M. R.; Katzmann, J. A.; Dispenzieri, A.; Murray, D. L. Using Mass Spectrometry to Monitor Monoclonal Immunoglobulins in Patients with a Monoclonal Gammopathy. *Journal of Proteome Research* **2014**, *13* (3), 1419-1427. DOI: 10.1021/pr400985k.
- (17) Pitt, J. J. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *Clin Biochem Rev* **2009**, *30* (1), 19-34. From NLM.
- (18) McLaren, D. G.; Shah, V.; Wisniewski, T.; Ghislain, L.; Liu, C.; Zhang, H.; Saldanha, S. A. High-Throughput Mass Spectrometry for Hit Identification: Current Landscape and Future Perspectives. *SLAS Discov* **2021**, *26* (2), 168-191. DOI: 10.1177/2472555220980696 From NLM.
- (19) Dueñas, M. E.; Peltier-Heap, R. E.; Leveridge, M.; Annan, R. S.; Büttner, F. H.; Trost, M. Advances in high-throughput mass spectrometry in drug discovery. *EMBO Mol Med* **2023**, *15* (1), e14850. DOI: 10.15252/emmm.202114850 From NLM.
- (20) Fuh, M. M.; Heikaus, L.; Schlüter, H. MALDI mass spectrometry in medical research and diagnostic routine laboratories. *International Journal of Mass Spectrometry* **2017**, *416*, 96-109. DOI: <https://doi.org/10.1016/j.ijms.2016.10.004>.
- (21) Ifa, D. R.; Wiseman, J. M.; Song, Q.; Cooks, R. G. Development of capabilities for imaging mass spectrometry under ambient conditions with desorption electrospray ionization (DESI). *International Journal of Mass Spectrometry* **2007**, *259* (1), 8-15. DOI: <https://doi.org/10.1016/j.ijms.2006.08.003>.
- (22) Karas, M.; Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry* **1988**, *60* (20), 2299-2301. DOI: 10.1021/ac00171a028.
- (23) Wiseman, J. M.; Laughlin, B. C. Desorption electrospray ionization (DESI) mass spectrometry: a brief introduction and overview. *Current Separations and Drug Development* **2007**, *22*, 11-14.
- (24) Zubair, F. MALDI mass Spectrometry based proteomics for drug discovery & development. *Drug Discovery Today: Technologies* **2021**, *40*, 29-35. DOI: <https://doi.org/10.1016/j.ddtec.2021.09.002>.
- (25) Berger, S. A.; Grimm, C.; Nyenhuis, J.; Payer, S. E.; Oroz-Guinea, I.; Schrittwieser, J. H.; Kroutil, W. Rapid, Label-Free Screening of Diverse Biotransformations by Flow-Injection Mass Spectrometry. *ChemBioChem* **2023**, *24* (11), e202300170. DOI: <https://doi.org/10.1002/cbic.202300170>.
- (26) Chang, M. S.; Kim, E. J.; El-Shourbagy, T. A. Evaluation of 384-well formatted sample preparation technologies for regulated bioanalysis. *Rapid Communications in Mass Spectrometry* **2007**, *21* (1), 64-72. DOI: <https://doi.org/10.1002/rcm.2808>.
- (27) Cody, R. B.; Laramée, J. A.; Durst, H. D. Versatile New Ion Source for the Analysis of Materials in Open Air under

- Ambient Conditions. *Analytical Chemistry* **2005**, 77 (8), 2297-2302. DOI: 10.1021/ac050162j.
- (28) Kertesz, V.; Van Berkel, G. J.; Vavrek, M.; Koeplinger, K. A.; Schneider, B. B.; Covey, T. R. Comparison of Drug Distribution Images from Whole-Body Thin Tissue Sections Obtained Using Desorption Electrospray Ionization Tandem Mass Spectrometry and Autoradiography. *Analytical Chemistry* **2008**, 80 (13), 5168-5177. DOI: 10.1021/ac800546a.
- (29) Shimma, S.; Sugiura, Y. Effective Sample Preparations in Imaging Mass Spectrometry. *Mass Spectrom (Tokyo)* **2014**, 3 (Spec Issue), S0029. DOI: 10.5702/massspectrometry.S0029 From NLM.
- (30) Hu, H.; Singh, A. N.; Lehnher, D.; Mdluli, V.; Chun, S. W.; Makarewicz, A. M.; Gouker, J. R.; Ukaegbu, O.; Li, S.; Wen, X.; et al. Accelerating Pharmaceutical Process Development with an Acoustic Droplet Ejection-Multiple Reaction Monitoring-Mass Spectrometry Workflow. *Analytical Chemistry* **2024**, 96 (3), 1138-1146. DOI: 10.1021/acs.analchem.3c04211.
- (31) Aretz, I.; Meierhofer, D. Advantages and Pitfalls of Mass Spectrometry Based Metabolome Profiling in Systems Biology. *Int J Mol Sci* **2016**, 17 (5). DOI: 10.3390/ijms17050632 From NLM.
- (32) Yamagaki, T.; Yamazaki, T. Troubleshooting Carry-Over in the LC-MS Analysis of Biomolecules: The Case of Neuropeptide Y. *Mass Spectrom (Tokyo)* **2019**, 8 (2), S0083. DOI: 10.5702/massspectrometry.S0083 From NLM.
- (33) Bretschneider, T.; Ozbal, C.; Holstein, M.; Winter, M.; Buettner, F. H.; Thamm, S.; Bischoff, D.; Luippold, A. H. RapidFire BLAZE-Mode Is Boosting ESI-MS Toward High-Throughput-Screening. *SLAS Technology* **2019**, 24 (4), 386-393. DOI: <https://doi.org/10.1177/2472630318822449>.
- (34) DiRico, K. J.; Hua, W.; Liu, C.; Tucker, J. W.; Ratnayake, A. S.; Flanagan, M. E.; Troutman, M. D.; Noe, M. C.; Zhang, H. Ultra-High-Throughput Acoustic Droplet Ejection-Open Port Interface-Mass Spectrometry for Parallel Medicinal Chemistry. *ACS Medicinal Chemistry Letters* **2020**, 11 (6), 1101-1110. DOI: 10.1021/acsmchemlett.0c00066.
- (35) Van Berkel, G. J.; Kertesz, V. An open port sampling interface for liquid introduction atmospheric pressure ionization mass spectrometry. *Rapid Communications in Mass Spectrometry* **2015**, 29 (19), 1749-1756.
- (36) Kempa, E. E.; Hollywood, K. A.; Smith, C. A.; Barran, P. E. High throughput screening of complex biological samples with mass spectrometry – from bulk measurements to single cell analysis. *The Analyst* **2019**, 144 (3), 872-891. DOI: 10.1039/c8an01448e.
- (37) Rozing, G. B.; Hans; Aldridge, Sibylle; Wenz, Christian. *CE/MS Principles and Practices: A guidebook for novices and practitioner*; 2019.
- (38) Huikko, K.; Kotiaho, T.; Kostianen, R. Effects of nebulizing and drying gas flow on capillary electrophoresis/mass spectrometry. *Rapid Commun Mass Spectrom* **2002**, 16 (16), 1562-1568. DOI: 10.1002/rcm.744 From NLM.
- (39) Henion, J. D.; Mordehai, A. V.; Cai, J. Quantitative Capillary Electrophoresis-Ion Spray Mass Spectrometry on a Benchtop Ion Trap for the Determination of Isoquinoline Alkaloids. *Analytical Chemistry* **1994**, 66 (13), 2103-2109. DOI: 10.1021/ac00085a028.
- (40) Diefenbach, X. W.; Farasat, I.; Guetschow, E. D.; Welch, C. J.; Kennedy, R. T.; Sun, S.; Moore, J. C. Enabling Biocatalysis by High-Throughput Protein Engineering Using Droplet Microfluidics Coupled to Mass Spectrometry. *ACS Omega* **2018**, 3 (2), 1498-1508. DOI: 10.1021/acsomega.7b01973 From NLM.
- (41) Holland-Moritz, D. A.; Wismer, M. K.; Mann, B. F.; Farasat, I.; Devine, P.; Guetschow, E. D.; Mangion, I.; Welch, C. J.; Moore, J. C.; Sun, S.; Kennedy, R. T. Mass Activated Droplet Sorting (MADS) Enables High-Throughput Screening of Enzymatic Reactions at Nanoliter Scale. *Angewandte Chemie International Edition* **2020**, 59 (11), 4470-4477. DOI: <https://doi.org/10.1002/anie.201913203> (accessed 2025/02/25).
- (42) Harvey, D. J. Chapter 14 - Mass spectrometric detectors for gas chromatography. In *Gas Chromatography (Second Edition)*, Poole, C. F. Ed.; Elsevier, 2021; pp 399-424.
- (43) Taylor, P. J. Matrix effects: the Achilles heel of quantitative high-performance liquid chromatography–electrospray–tandem mass spectrometry. *Clinical Biochemistry* **2005**, 38 (4), 328-334. DOI: <https://doi.org/10.1016/j.clinbiochem.2004.11.007>.
- (44) Zhou, W.; Yang, S.; Wang, P. G. Matrix Effects and Application of Matrix Effect Factor. *Bioanalysis* **2017**, 9 (23), 1839-1844. DOI: 10.4155/bio-2017-0214.
- (45) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* **2020**, 17 (3), 261-272. DOI: 10.1038/s41592-019-0686-2.
- (46) Savitzky, A.; Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry* **1964**, 36 (8), 1627-1639. DOI: 10.1021/ac60214a047.
- (47) Savile, C. K.; Janey, J. M.; Mundorff, E. C.; Moore, J. C.; Tam, S.; Jarvis, W. R.; Colbeck, J. C.; Krebber, A.; Fleitz, F. J.; Brands, J.; et al. Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science* **2010**, 329 (5989), 305-309. DOI: 10.1126/science.1188934 From NLM.
- (48) Dunham, N. P.; Winston, M. S.; Ray, R.; Eberle, C. M.; Newman, J. A.; Gao, Q.; Cao, Y.; Barrientos, R. C.; Ji, Y.; Reibarkh, M. Y.; Silverman, S. M. Transaminase-Catalyzed Synthesis of β -Branched Noncanonical Amino Acids Driven by a Lysine Amine Donor. *Journal of the American Chemical Society* **2024**, 146 (23), 16306-16313. DOI: 10.1021/jacs.4c05175.
- (49) Raps, F. C.; Rivas-Souchet, A.; Jones, C. M.; Hyster, T. K. Emergence of a distinct mechanism of C–N bond formation in photoenzymes. *Nature* **2024**. DOI: 10.1038/s41586-024-08138-w.
- (50) Müller, T. E.; Hultsch, K. C.; Yus, M.; Foubelo, F.; Tada, M. Hydroamination: Direct Addition of Amines to Alkenes and Alkynes. *Chemical Reviews* **2008**, 108 (9), 3795-3892. DOI: 10.1021/cr0306788.
- (51) Bariwal, J.; Van der Eycken, E. C–N bond forming cross-coupling reactions: an overview. *Chem Soc Rev* **2013**, 42 (24), 9283-9303. DOI: 10.1039/c3cs60228a From NLM.
- (52) Lee, C.; Kang, H.-J.; Hong, S. NiH-catalyzed C–N bond formation: insights and advancements in hydroamination of unsaturated hydrocarbons. *Chemical Science* **2024**, 15 (2), 442-457. DOI: 10.1039/d3sc05589b.
- (53) Zheng, Y.-N.; Zheng, H.; Li, T.; Wei, W.-T. Recent Advances in Copper-Catalyzed C–N Bond Formation Involving N-Centered Radicals. *ChemSusChem* **2021**, 14 (24), 5340-5358. DOI: <https://doi.org/10.1002/cssc.202102243> (accessed 2024/10/22).

(54) Ruiz-Castillo, P.; Buchwald, S. L. Applications of Palladium-Catalyzed C–N Cross-Coupling Reactions. *Chemical Reviews* **2016**, *116* (19), 12564-12649. DOI: 10.1021/acs.chemrev.6b00512.

(55) Forero-Cortés, P. A.; Haydl, A. M. The 25th Anniversary of the Buchwald–Hartwig Amination: Development,

Applications, and Outlook. *Organic Process Research & Development* **2019**, *23* (8), 1478-1483. DOI: 10.1021/acs.oprd.9b00161.
