

## Chemical shift variations in common metabolites

Fatema Bhinderwala<sup>a,b,c,1</sup>, Heidi E. Roth<sup>a,1</sup>, Hannah Noel<sup>a</sup>, Dennis Feng<sup>a</sup>, Robert Powers<sup>a,b,\*</sup>

<sup>a</sup> Department of Chemistry, University of Nebraska-Lincoln, Lincoln, NE 68588-0304, United States

<sup>b</sup> Nebraska Center for Integrated Biomolecular Communication, University of Nebraska-Lincoln, Lincoln, NE 68588-0304, United States

<sup>c</sup> University of Pittsburgh School of Medicine, Department of Structural Biology, Pittsburgh, PA 15260, United States<sup>2</sup>



### ARTICLE INFO

#### Article history:

Received 8 August 2022

Revised 31 October 2022

Accepted 9 November 2022

Available online 13 November 2022

#### Keywords:

pH

Temperature

<sup>1</sup>H NMR

Automated metabolite assignment

Biofluids

### ABSTRACT

The reliability and robustness of metabolite assignments in <sup>1</sup>H NMR is complicated by numerous factors including variations in temperature, pH, buffer choice, ionic strength, and mixture composition that led to peak overlap and spectral crowding. As sample conditions fluctuate, peak drift and line broadening further complicate peak deconvolution and subsequent chemical assignment. We present a collection of 1D <sup>1</sup>H NMR spectra of 54 common metabolites at varied pH (6.0 to 8.0 in 0.5 step increments) and temperature (290 K to 308 K) to quantify chemical shift variability to facilitate automated metabolite assignments. Our results illustrate the fundamental challenges with accurately assigning NMR peaks under varied environmental conditions prevalent in complex mixtures. Phosphorylated metabolites showed a larger variation in chemical shifts due to pH, whereas; amino acids showed a higher variation due to temperature. Mixtures of phosphorous compounds showed a consistently poor reliability in achieving an accurate assignment. Phosphorylated cholines, amino acids, and glycerols yielded a 40 % false negative rate for 7 out of 9 mixture conditions. Amino acids had a false negative rate of 57 % at 298 K and pH 8. Our results demonstrate that the automated assignments of complex biofluid mixtures require an expert to intervene to confirm the accuracy of metabolite assignments. Our analysis also indicates the need for reference databases to include spectra under a variety of conditions that includes mixtures and a range of pH and temperature to improve the accuracy and reproducibility of metabolite assignments.

© 2022 Elsevier Inc. All rights reserved.

### 1. Introduction

Metabolomics is a rapidly expanding field reliant on different analytical techniques and variable sample origins that create unique challenges to accurately annotate spectral data [1]. One-dimensional proton (1D) <sup>1</sup>H NMR offers many advantages in metabolomics profiling, namely the relative simplicity of sample preparation [2], a robust method to accurately quantify the most abundant compounds in a metabolome [3], the ready identification of compounds that are difficult to ionize by mass spectrometry [4], the identification of compounds of identical mass, and studies in living organisms [4,5]. Thus, 1D <sup>1</sup>H NMR is ubiquitously performed for a variety of qualitative and quantitative metabolomics studies [6]. A major challenge commonly encountered with 1D <sup>1</sup>H NMR metabolomics is the increase in spectral crowding and peak over-

lap that occurs with an increase in sample complexity [7]. Accurate metabolite assignments become ambiguous, complicated, and potentially, impossible to complete with an increase in spectral complexity [3,8]. In addition to peak overlap, external influences that includes variations in pH, temperature, ionic strength, buffer choice, metabolite concentrations, and mixture composition also present challenges to accurate assignments [9]. These and other factors may induce significant chemical shift deviations from standard reference spectra.

One approach to limiting these issues is the use of two-dimensional (2D) experiments such as 2D <sup>1</sup>H-<sup>1</sup>H TOCSY and 2D <sup>1</sup>H-<sup>13</sup>C HSQC to validate the 1D <sup>1</sup>H NMR assignments [10]. While homonuclear and heteronuclear 2D correlation experiments can help resolve overlapped regions and account for chemical shift variability (*i.e.*, lower sensitivity of <sup>13</sup>C chemical shifts to sample conditions), but they also incur an increase in overall experiment time and may potentially lead to a loss of information. For example, peaks may be missing in 2D spectra due to a lower signal-to-noise, because of the lower sensitivity of <sup>15</sup>N/<sup>13</sup>C resonances in natural abundance samples, or because only a selective subset of the

\* Corresponding author at: University of Nebraska-Lincoln, Department of Chemistry, 722 Hamilton Hall, Lincoln, NE 68588-0304, United States.

E-mail address: [rpowers3@unl.edu](mailto:rpowers3@unl.edu) (R. Powers).

<sup>1</sup> These authors contributed equally to the work.

<sup>2</sup> Current address.

metabolites were  $^{15}\text{N}$ - or  $^{13}\text{C}$ -labeled from a  $^{15}\text{N}/^{13}\text{C}$ -tracer (e.g.,  $^{13}\text{C}_6$ -glucose).

Sample handling and preparation strategies greatly influence the final spectral quality and ease of metabolite assignments [11]. Sample condition variability, such as differences in pH and temperature, are critical factors that can negatively influence assignment accuracy and should be carefully considered during the study design [12,13]. Unfortunately, the large variation in sample preparation and data collection protocols employed by the NMR metabolomics community greatly complicates the metabolite assignment strategy. Metabolite assignment software or protocols range from fully automated to entirely manual approaches [8,14–16]. Fully automated approaches are quick but can lead to high errors in the assignments especially when variable sample conditions are encountered. Alternatively, manual assignments using public databases such as the human metabolome database (HMDB) [17] or the biological magnetic resonance bank (BMRB) [18] can produce highly accurate results but are extremely time consuming and requires extensive training [16]. Nevertheless, manual approaches to metabolite assignments are also severely challenged by poor spectral matches to reference libraries due to environmentally induced chemical shift changes. Simply, the magnitude and direction of a chemical shift perturbation for each individual NMR resonances is highly variable, and in practice, unpredictable, which makes unambiguous matches extremely difficult. Similar changes in peak linewidths further confound the process of annotating NMR metabolomics spectra.

To start resolving these issues, we have collected a series of 1D  $^1\text{H}$  NMR spectra of 54 common metabolites acquired over a range of pHs and temperatures, and then compiled the chemical shift changes on a per resonance basis. We have catalogued chemical shift uncertainties relative to reference spectra at pH 7 and 298 K for a diversity of functional groups. Further, we demonstrate the challenges faced by variable pH and temperature conditions in the analysis of 1D  $^1\text{H}$  NMR spectra of individual metabolites and mixtures. The magnitude of peak shifts and linewidth changes are metabolite dependent with the NMR spectra of some compounds being more significantly impacted than others. As a result, software designed to automate NMR assignments are severely challenged by these chemical shift uncertainties and, unsurprisingly, by the increasing complexity of the mixtures [8]. Our results clarify the need for spectral databases to expand their content to include pH and temperature variations for a diversity of chemical functional groups and reference spectra for a variety of complex mixtures including biofluids.

## 2. Material and methods

### 2.1. Urine and serum samples

De-identified human serum samples were obtained from the NIH NeuroBioBank, specifically the University of Miami Brain Endowment Bank. De-identified human urine samples were obtained from the Multiple Sclerosis Clinic within the Saunders Medical Center (Wahoo, NE, USA). All materials, biospecimens, and human subjects' data collected, stored, and maintained for and during the conduct of this research was reviewed and approved by the University of Nebraska-Lincoln Institutional Review Board (IRB, IRB#: 20200820533EP, IRB#: 20180517991EPCOLLA).

### 2.2. Preparation of standard NMR samples

The standard samples for 1D  $^1\text{H}$  NMR data collection included 5 mixtures and 54 individual compounds consisting of 33

phosphorous-containing compounds and 21 amino acids. The phosphorous-containing compounds were divided into four groups comprising four NMR samples and consisted of phosphorylated nucleic acid analogs (group 1, 11 compounds), sugars (group 2, 8 compounds), coenzymes (group 3, 6 compounds), and cholines, amino acids and glycerols (group 4, 8 compounds). The 21 amino acids were combined to form one mixture of amino acids. The complete list of individual compounds and mixture compositions are available in Table S1. Stock solutions for amino acids and phosphorous-containing compounds were prepared in  $\text{D}_2\text{O}$  (Sigma Aldrich, 99.8 %) at a concentration of 90 mM and 100 mM, respectively. The stock solutions were then diluted into a 50 mM phosphate buffer with 50  $\mu\text{M}$  of 3-(tetramethylsilane) propionic acid-2,2,3,3- $\text{d}_4$  (TMSP- $\text{D}_4$ ) in 99.8 %  $\text{D}_2\text{O}$  at five different pH values (uncorrected) of 6.0, 6.5, 7.0, 7.5, and 8.0 to prepare NMR samples for each individual compound at a final concentration of 1.5 mM and 15 mM for amino acids and phosphorous-containing compounds, respectively. 1D  $^1\text{H}$  NMR spectra were collected at 298 K. In addition, 1D  $^1\text{H}$  NMR spectra for the pH 7.0 samples were collected at 290 K, 294 K, 304 K, and 308 K. The same protocol was followed for preparing each of the five standard mixtures outlined in Table S1.

### 2.3. Preparation of serum NMR samples

A 150  $\mu\text{L}$  aliquot of serum was added to 300  $\mu\text{L}$  of methanol to precipitate proteins. The 1:2 mixture of serum:methanol was vortexed for 10 s and then incubated at 4  $^\circ\text{C}$  for 10 min followed by centrifuging at 14,000 g for 20 mins at 4  $^\circ\text{C}$  to pellet the proteins. The supernatant was collected and centrifuged again at 14,000 g for 5 min at 4  $^\circ\text{C}$  to pellet any remaining proteins. Supernatants were then collected, snap frozen in liquid nitrogen, and dried by speed vacuum centrifugation (SpeedVac R Plus, Savant) followed by lyophilization using FreeZone<sup>TM</sup> (Labconco, Kansas City, MO) for 24 h. Samples were then stored at  $-80$   $^\circ\text{C}$  until NMR analysis. At time of data collection, samples were reconstituted using 150  $\mu\text{L}$  of a 50 mM phosphate buffer in "100 %"  $\text{D}_2\text{O}$  at pH 7.2 (uncorrected) with 50  $\mu\text{M}$  of TMSP- $\text{D}_4$  as a chemical shift reference. The samples were centrifuged at 14,000 g for 20 min at 4  $^\circ\text{C}$  to remove any precipitant. The supernatants were placed in a 3 mm NMR tube for data acquisition [19].

### 2.4. Preparation of urine NMR samples

A 150  $\mu\text{L}$  aliquot of urine was centrifuged at 14,000 g for 10 min to pellet debris. 135  $\mu\text{L}$  of the supernatant was combined with 15  $\mu\text{L}$  of a 50 mM phosphate buffer in "100 %"  $\text{D}_2\text{O}$  at pH 7.2 (uncorrected) with 50  $\mu\text{M}$  of TMSP- $\text{D}_4$  as a chemical shift reference [20]. The samples were then transferred to a 3 mm NMR tube for data acquisition.

### 2.5. NMR data collection and processing for standard samples

1D  $^1\text{H}$  NMR spectra were collected on a Bruker Avance III-HD 700 MHz spectrometer equipped with a quadruple resonance QCI-P cryoprobe ( $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^{31}\text{P}$ ) with z-axis gradients. A SampleJet automated sample changer system, automated tune and match device, and Bruker ICON NMR software was used to automate the NMR data collection. 1D  $^1\text{H}$  NMR spectra were collected with 32 K data points, 64 scans, 4 dummy scans, and a spectral width of 9,090 Hz using an excitation sculpting pulse sequence [21]. 1D  $^1\text{H}$  NMR spectra were processed using Bruker TopSpin 3.6 and MestreNova 12.0.2 (<https://mestrelab.com/>) in NMRBox [22]. The spectra were processed with exponential multiplication and twice zero filled followed by Fourier transform. Spectra were manually phased and referenced to TMSP- $\text{D}_4$ . Peaks were

picked, integrated, and then exported to obtain peak centers and linewidths. The NMR data set is available by request from the authors and will be made available through the National Metabolomics Data Repository at Metabolomics Workbench (<https://www.metabolomicsworkbench.org/>).

### 2.6. 1D $^1\text{H}$ NMR and 2D $^1\text{H}$ - $^{13}\text{C}$ HSQC data collection and processing for biofluids

1D  $^1\text{H}$  NMR spectra were collected with 64 K data points, 138 scans, 4 dummy scans, and a spectral width of 9,090 Hz using an excitation sculpting pulse sequence [21]. Natural abundance 2D  $^1\text{H}$ - $^{13}\text{C}$  HSQC, and  $^1\text{H}$ - $^1\text{H}$  TOCSY spectra were collected at 25 % sparsity using our deterministic non-uniform sampling (NUS) schedule [23]. 2D  $^1\text{H}$ - $^1\text{H}$  TOCSY spectra were collected with 8 scans, 16 dummy scans, and 2,048 data points and a spectral width of 7,003 Hz in the direct dimension and 256 data points with a spectral width of 7,003 Hz in the indirect dimension. 2D  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectra were collected with 128 scans, 16 dummy scans, and 1,024 data points and a spectral width of 11,160 Hz in the direct dimension and 256 data points and a spectral width of 29,060 in the indirect dimension. Both 1D and 2D NMR spectra were processed using Topspin 3.6 and were Fourier transformed, phased, and referenced to TMS. NUS spectra were reconstructed using the multi-dimensional decomposition method (MDD) [24].

### 2.7. Metabolite assignment strategy

The metabolites in the standard and biofluid mixtures were assigned using Chenomx NMR Suite version 8.51 (<https://www.chenomx.com/>), the Human Metabolome Database (<https://hmdb.ca/>) [17] and the Chenomx 700 MHz reference library. Spectra labeled 'automated' were assigned using the batch fit algorithm with reference libraries containing the set of 33 phosphorous compounds and 21 amino acids. Automated assignments were then triaged, or manually inspected and validated for accuracy. An overview of the protocol is shown in Fig. 1A.

## 3. Results and discussion

### 3.1. Need for cataloging chemical shift variations across metabolite classes

1D  $^1\text{H}$  NMR is a routine choice for untargeted metabolomics since it provides information rich data that is easy to acquire and amenable to high throughput analysis [25]. However, the process of annotating NMR spectral data and assigning metabolites to their corresponding spectral peaks is a highly time consuming process that is prone to errors (Fig. 1A) [26]. Spectral complexity and the resulting high degree of peak overlap further complicates the assignment workflow. This is clearly evident from the histogram plot of chemical shift distributions for metabolites deposited in the HMDB (Fig. 1B) [17]. The low spectral resolution also leads to a loss in observable peak multiplicities, which are equally important to the assignment process and would result in more ambiguous peaks.

2D NMR experiments that include 2D  $^1\text{H}$ - $^1\text{H}$  TOCSY (Fig. 1C) and 2D  $^1\text{H}$ - $^{13}\text{C}$  HSQC (Fig. 1D) can be acquired to supplement the 1D  $^1\text{H}$  NMR data and partly resolve the peak overlap problem. Of course, the significant increase in experiment time, the decrease in signal-to-noise, especially for natural abundant samples, and the potential loss of low abundant metabolites may diminish the utility of 2D NMR experiments compared to 1D NMR experiments. Further, the use of a  $^{15}\text{N}/^{13}\text{C}$  tracer chemical to enrich the  $^{15}\text{N}$ - or  $^{13}\text{C}$ -labeling of the metabolome is only selective for metabolites

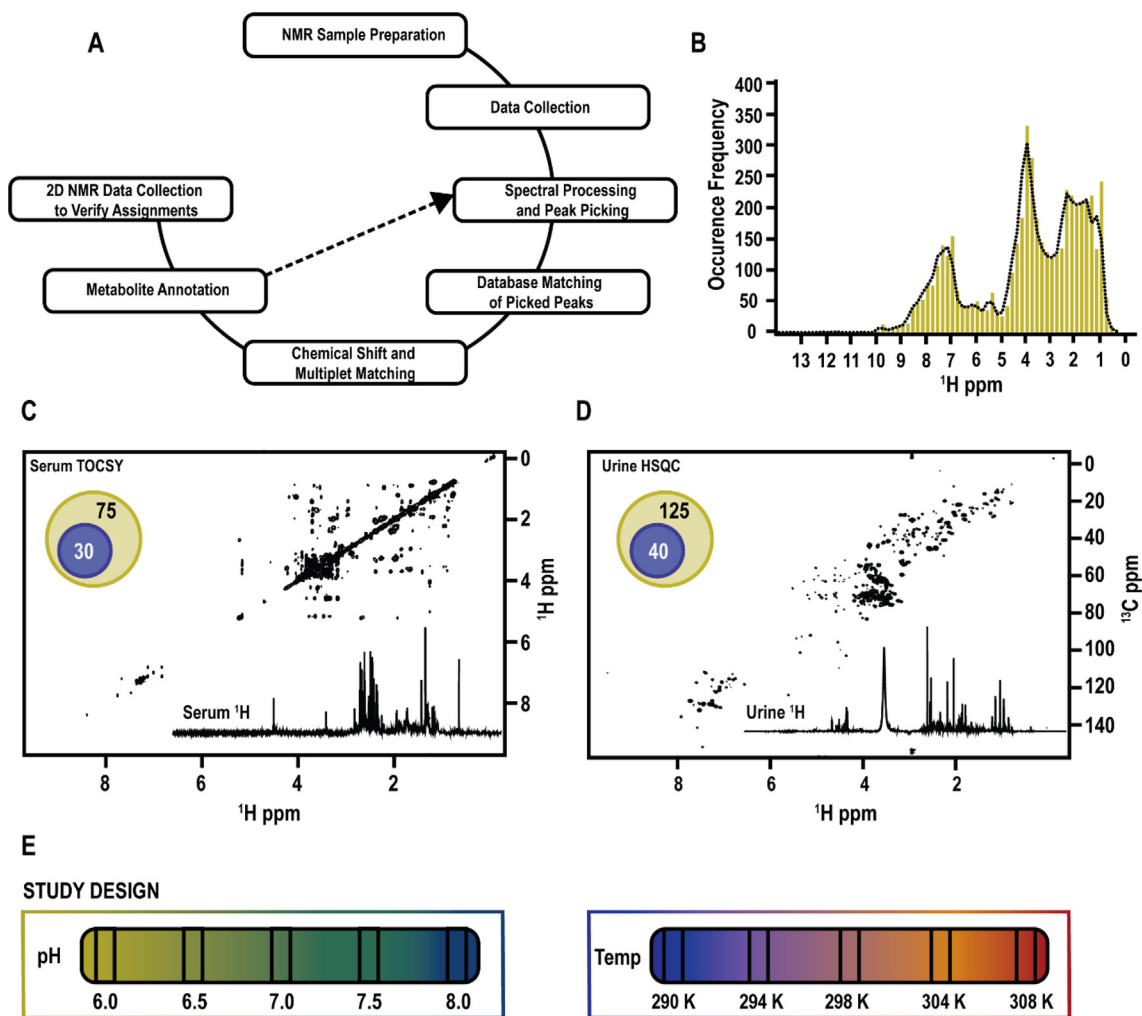
enzymatically derived from the  $^{15}\text{N}/^{13}\text{C}$  tracer. In this regard, metabolites detected by the 1D NMR experiment may be missing in the 2D NMR spectra of a  $^{15}\text{N}/^{13}\text{C}$ -labeled metabolome. Chemical shift variability resulting from differences in pH, temperature, ionic strength, buffer, metabolite concentrations, and mixture composition further confounds the metabolite assignment process [9]. The uncertainties or errors in chemical shifts are unknown values especially on a per peak basis. Therefore, both manual and automated approaches to metabolite assignments tend to use a single large error range to ensure matches between reference and experimental NMR spectra. A higher false positive rate is the likely outcome with generous, uniform error-bars, while alternatively, a higher false negative rate will occur with tighter error-bars. Further, the choice of chemical shift error ranges is completely reliant on uninformed guesses. To partly address these issues, a series of 1D  $^1\text{H}$  NMR spectra were collected at various temperatures and pH values to better define chemical shift error-bars for matching spectra (Fig. 1E). Our spectral data also illustrates the broad impact of pH and temperature changes on peak linewidths. This demonstrates the need for metabolomics databases to augment 1D  $^1\text{H}$  NMR data sets with spectra collected over a range of temperatures and pH values to increase assignment efficiency and reliability.

### 3.2. pH dependent chemical shift variation across metabolite groups

A set of 1D  $^1\text{H}$  NMR spectra were collected for each of the common metabolites comprising amino acids and phosphorous-containing compounds as detailed in Table S1. The 21 common amino acids provided enough structural diversity to characterize trends in chemical shift variations based on functional group. Additionally, the common structural features for the amino acids highlight annotation challenges due to chemical shift overlap. The 33 phosphorous containing compounds provided a similar framework while exhibiting a higher sensitivity to pH and a likely upper-bound to chemical shift variability. The sample pH was only varied from 6 to 8 to bracket the typical target pH of 7–7.4 and to capture the likely maximal pH-induced chemical shift changes encountered in routine metabolomics samples. Fig. 2 summarizes the pH dependent chemical shift variations per functional group for the 21 amino acids and the 33 phosphorous containing compounds. Full profiles of the pH-induced  $^1\text{H}$  chemical shifts changes for the 21 amino acids are presented in Fig. S1A–D and tabulated in Table S2.

Fig. 2A–D displays  $^1\text{H}$  chemical shift variations compared to pH 7 for the selected amino acids.  $^1\text{H}$  chemical shift variations as a function of pH were grouped according to the common  $\text{H}\alpha$ ,  $\text{H}\beta$ ,  $\text{H}\gamma$ , and  $\text{H}\delta$  notations for amino acids (Fig. 2E). Most NMR resonances exhibited a downfield shift as the pH increased from 6 to 8. The  $^1\text{H}$  chemical shift changes ( $\Delta^1\text{H}$  ppm) ranged from  $-0.008$  ppm to  $+0.1$  ppm. The magnitude and spread of these chemical shift changes generally increased in order from  $\text{H}\alpha$  to  $\text{H}\delta$  with  $\text{H}\gamma$  and  $\text{H}\delta$  having the greatest pH-induced chemical shift changes.  $\text{H}\gamma$  and  $\text{H}\delta$  chemical shifts ranged from  $-0.002$  ppm to  $+0.005$  ppm and from  $-0.008$  ppm to  $+0.008$  ppm, respectively. An example of this trend can be easily seen by comparing the  $\text{H}\alpha$  of Asp (Fig. 2A) with its  $\text{H}\beta$  (Fig. 2B). Four notable outliers from this trend were the  $\text{H}\beta$ s of leucine and histidine,  $\text{H}\gamma$  of glutamate, and the  $\text{H}\delta$  of isoleucine, which shifted upfield with variable magnitude changes as the pH was increased.

The amino acids with some of the largest pH-induced chemical shift changes were plotted separately for clarity in Fig. 2F–I. Specifically, the presence of exchangeable sidechain  $^1\text{H}$  resonances in arginine (pKa 12.48), cysteine (pKa 8.37), histidine (pKa 6.04), and lysine (pKa 10.54) resulted in sizable pH induced chemical shift changes for  $\text{H}\alpha$  (Fig. 2F) and  $\text{H}\delta$  (Fig. 2I) due to the sidechain pKa values. Notably, the  $\text{H}\beta$  and  $\text{H}\gamma$  shifts (Fig. 2G, H) were rela-



**Fig. 1.** Study Design to Assess Chemical Shift Variation Based on Temperature and pH (A) Workflow of NMR metabolite assignment protocol. The dashed arrow shows the iterative steps in the assignment process. (B) Bar graph demonstrating the distribution of  $^1\text{H}$  chemical shifts for metabolites in HMDB. A typical (C)  $2\text{D } ^1\text{H}-^1\text{H}$  TOCSY spectrum collected from a human serum sample and (D) a typical  $2\text{D } ^1\text{H}-^{13}\text{C}$  HSQC spectrum collected from a human urine sample. The corresponding  $1\text{D } ^1\text{H}$  NMR spectra are in the lower right-hand corner of each figure. Each Venn diagram represents the number of metabolites generally identified from the  $1\text{D } ^1\text{H}$  NMR (yellow) and  $2\text{D}$  NMR (blue) experiments for the indicated biofluids. (E) Study design color scheme that is used throughout the study to indicate a specific pH or temperature value. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

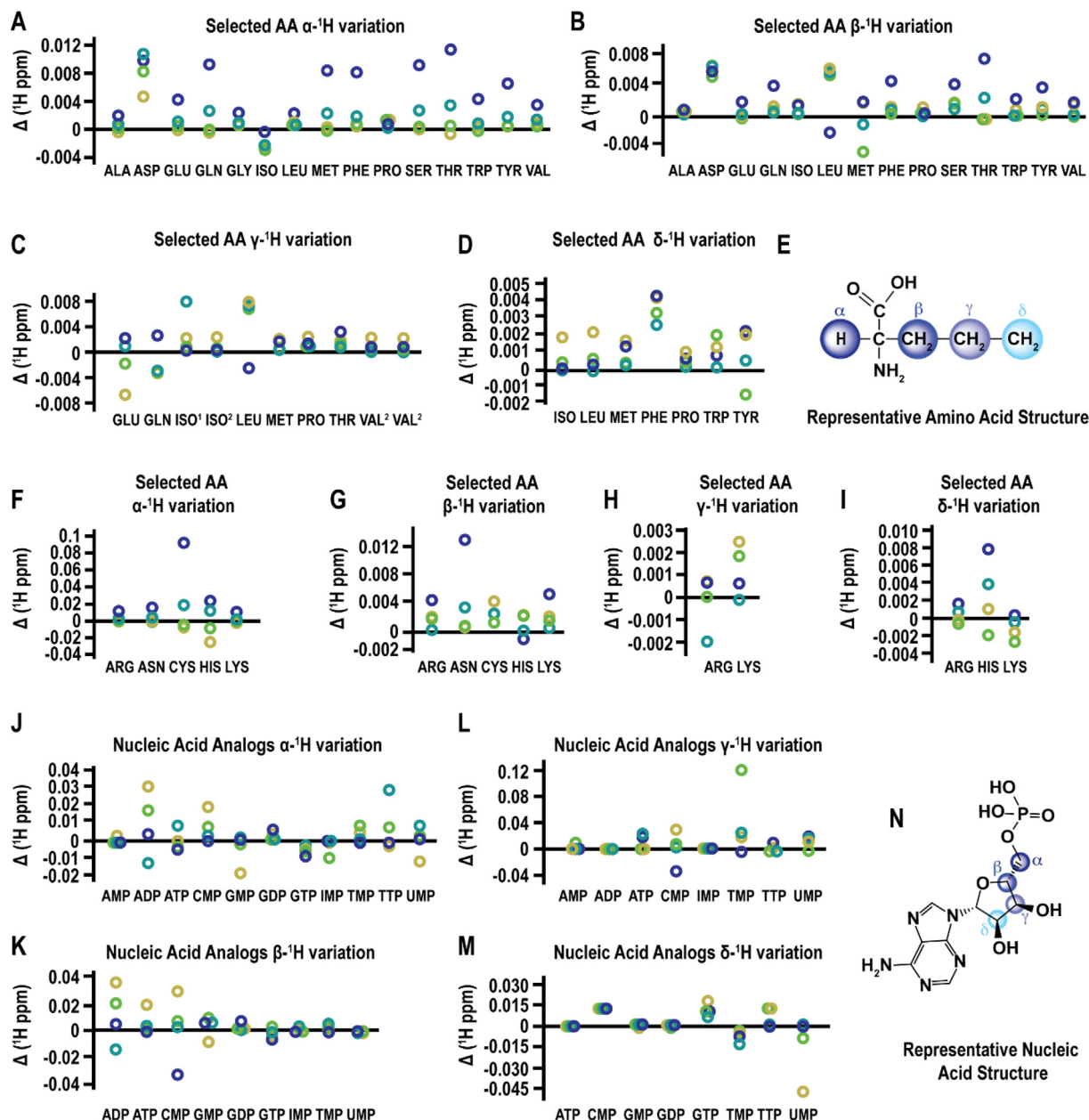
tively unaffected and similar to the changes observed for the other amino acids (Fig. 2B,C). A likely impact of hydrogen bonding in pH induced chemical shift changes was observed for serine and threonine. A relatively large variation of 0.004 to 0.008 ppm was observed for the  $\text{H}\beta$  chemical shift.

The phosphorous containing compounds were assigned to one of four groups that included (1) nucleic acid analogs, (2) phosphorylated analogs of sugars, (3) coenzymes, and (4) choline analogs, amino acids, and phosphoglycerols. Fig. 2J-M highlights the chemical shifts of nucleic acid analogs. Fig. 2N shows a generic nucleic acid structure and an illustration of the  $\text{H}\alpha$ ,  $\text{H}\beta$ ,  $\text{H}\gamma$ , and  $\text{H}\delta$  naming convention for the  $^1\text{H}$  chemical shifts. Full profiles of the  $^1\text{H}$  chemical shifts changes for the remaining phosphorous containing compounds are presented in Fig. S2A-J. Full profiles of the pH-induced  $^1\text{H}$  chemical shift changes for the 33 phosphorous containing compounds are tabulated in Table S3. The larger variance in pKa values and the considerable structural complexity of the phosphorous compounds compared to amino acids resulted in greater chemical shift variations. The largest  $\Delta^1\text{H}$  ppm for the amino acids was 0.013 for the  $\text{H}\delta$  of asparagine at pH 8.0 (Fig. 2G). Conversely, the largest  $\Delta^1\text{H}$  ppm for the phosphorous compounds was 0.078 for the  $\text{H}\delta$  of glucose 1,6-bisphosphate (GBP) at pH 6.0 (Fig. S2D).

There was also a considerable deviation from the pH-dependent downfield shift observed with the amino acids. In particular, the  $\text{H}\alpha$  (Fig. 2J) shifted drastically upfield (ADP) or downfield (TTP, UMP) at pH 7.5 compared to the other pH values. This inconsistency in a pH-dependent trend in  $\Delta^1\text{H}$  ppm for the phosphorous containing compounds was seen for all four groups and for each  $^1\text{H}$  resonance type and can again be attributed to pKa differences. It should be noted that the group with the highest chemical shift variation was the phosphorylated sugars, which when combined with their complex  $1\text{D } ^1\text{H}$  NMR spectra will likely lead to a higher uncertainty in obtaining a correct assignment from complex mixtures.

### 3.3. Temperature dependent chemical shift variation across metabolite groups

The effect of temperature on the chemical shifts for the 21 amino acids was also evaluated by varying the sample temperature from 290 K to 308 K in 4 K increments. The sample temperature was only varied from 290 K to 308 K to bracket the typical target temperature of 298 K and to capture the likely maximal temperature-induced chemical shift changes encountered in rou-



**Fig. 2. Chemical Shift Variation as a Function of pH and Metabolite Group.** (A-I) Dot plots summarizing the chemical shift variation across five different pH values for select amino acids. Amino acids with relatively low to moderate chemical shift changes are plotted in panels A-D. Amino acids with relatively high chemical shift changes are plotted in panels F-I. The  $^1\text{H}$  chemical shift changes ( $\Delta^1\text{H}$  ppm) relative to pH 7 for (A,F)  $\text{H}\alpha$ , (B,G)  $\text{H}\beta$ , (C,H)  $\text{H}\gamma$ , and (D,I)  $\text{H}\delta$  are plotted per amino acid. Each circle is colored according to pH as outlined in Fig. 1E. (E) General amino acid structure depicting the standard nomenclature used throughout the manuscript. (A-I) Dot plots summarizing the chemical shift variation across five different pH values for select phosphorus containing compounds (Group 1). The  $^1\text{H}$  chemical shift changes ( $\Delta^1\text{H}$  ppm) relative to pH 7 for (J)  $\text{H}\alpha$ , (K)  $\text{H}\beta$ , (L)  $\text{H}\gamma$ , and (M)  $\text{H}\delta$  are plotted per nucleic acid analog. (N) General nucleic acid structure depicting the standard nomenclature used throughout the manuscript. pH values are represented by colored circles: 6.0 (yellow), 6.5 (green), 7.5 (teal), 8.0 (blue). Table S1 lists the three-letter code for each metabolite. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

tine metabolomics samples. Temperature-dependent chemical shift changes were measured relative to 298 K. Full profiles of the temperature-induced  $^1\text{H}$  chemical shifts changes for the 21 amino acids are tabulated in Table S2. Temperature was found to have a greater impact on chemical shift and the  $\Delta^1\text{H}$  values presented an opposite trend compared to pH (Fig. S1E-F). The  $\Delta^1\text{H}$  values tended to increase moving along the amino acid sidechain from the  $\text{H}\alpha$  to the  $\text{H}\delta$  position. Specifically,  $\Delta^1\text{H}$  ppm ranged from  $-0.005$  to  $+0.01$  for  $\text{H}\alpha$ ,  $-0.005$  to  $+0.025$  for  $\text{H}\beta$ ,  $-0.01$  to  $+0.025$

for  $\text{H}\gamma$ , and  $-0.005$  to  $+0.020$  for  $\text{H}\delta$ . These chemical shift ranges were much broader than previously observed for the pH changes.

The effect of temperature on chemical shifts was also observed for the phosphorous containing compounds, and in contrast to the amino acids, temperature had less of an impact on the distribution of chemical shifts compared to pH (Fig. S2K-N). Full profiles of the temperature-induced  $^1\text{H}$  chemical shifts changes for the 33 phosphorous containing compounds are tabulated in Table S3. Most of the phosphorous containing compounds had

temperature-induced  $\Delta^1\text{H}$  values within  $\pm 0.05$  ppm. Conversely, the amino acids ranged from  $-0.01$  to  $+0.025$  ppm and were primarily shifted downfield. Peak shifts due to temperature and pH can vary drastically, where the exact range and magnitude of these chemical shift variations are distinct to each structure and NMR resonance. Consequently, chemical shift variabilities present unique challenges when interpreting NMR spectra of complex mixtures and creates serious uncertainties in obtaining correct metabolite assignments.

### 3.4. pH dependent changes to chemical shift also affects peak linewidths

In addition to chemical shift variabilities, changes in peak linewidths due to pH changes can further increase the challenges in assigning metabolites within complex mixtures. To quantify pH-induced linewidth changes, individual peak regions were fitted using MestreNova to estimate linewidths at each pH. Again, changes in linewidths were measured relative to pH 7. Full profiles of the temperature and pH-induced changes in peak multiplicity and linewidths for the 21 amino acids and 33 phosphorus containing compounds are tabulated in **Tables S2 and S3**, respectively. A series of stack plots of expanded 1D  $^1\text{H}$  spectra for select metabolites and NMR resonances over the pH range of 6 to 8 are shown in **Fig. 3A-I**. Each stack plot has a chemical structure of the metabolite that highlights the selected NMR resonances and a line plot indicating the linewidth changes at each pH. Selected peaks for the amino acids arginine (**Fig. 3A**), histidine (**Fig. 3B**), and lysine (**Fig. 3C**) show that peaks broaden variably as a function of pH. Arginine and histidine exhibited broader linewidths at lower pH, while lysine showed broader linewidths at higher pH. The effect of pH on the apparent peak multiplicity was also demonstrated with histidine (**Fig. 3B**) where the  $\text{H}\alpha$  peak shifted from a triplet at pH 6.0 to quartet at pH 6.5–8.0.

Similar results were observed for selected peaks from the phosphorous compounds pyridoxal 5'-phosphate (**Fig. 3D**), fructose 6-phosphate (**Fig. 3E**), glucose 1,6-bisphosphate (**Fig. 3F**), glucose 6-phosphate (**Fig. 3G**), cytidine monophosphate (**Fig. 3H**), and thymidine monophosphate (**Fig. 3I**). Like the amino acids, the selected peaks broaden differently as a function of pH. The NMR peaks for fructose 6-phosphate, glucose 1,6-bisphosphate, glucose 6-phosphate, and cytidine monophosphate became sharper at lower pH, but pyridoxal 5'-phosphate became broader. Apparent peak multiplicities again changed as a function of pH in fructose 6-phosphate and thymidine monophosphate. At pH 7, the fructose 6-phosphate  $\text{H}\beta$  displays a single complex multiplet of low intensity that separates into two distinct multiplets at both higher and lower pH values. Similarly, decreasing the pH causes the thymidine monophosphate  $\text{H}\gamma$  multiplet to collapse into a neighboring multiplet. The presence of exchangeable protons further complicates pH-dependent linewidth changes leading to variable peak broadening. Both thymidine monophosphate and glucose 1,6-bisphosphate demonstrate this effect by greater peak broadening at pH 6.5 and 7.5 relative to pH 6.0 and 8.0.

### 3.5. pH dependent linewidth changes lead to spectral crowding

Variable changes in peak positions, splitting patterns and linewidths makes it difficult to analyze a 1D  $^1\text{H}$  NMR spectrum and assign a metabolite. The metabolite assignment problem is further complicated by complex mixtures that leads to peak overlap and spectral crowding. To clearly illustrate the problem, various bubble plots are shown in **Fig. 4** where each circle's center and diameter corresponds to the peak center and the chemical shift range, respectively, that the peak occupies at a given pH. The distribution of chemical shifts and linewidths results in spectral regions that

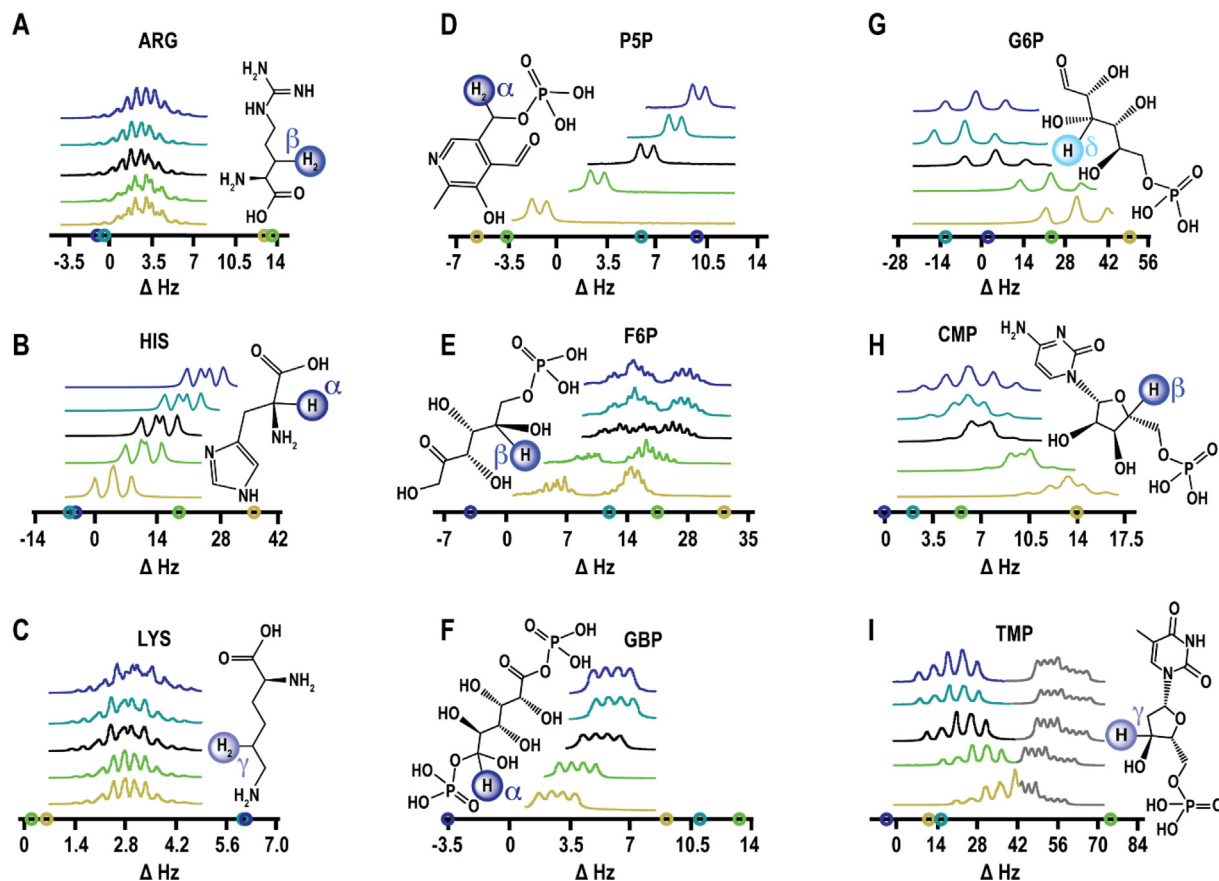
are highly populated with extensive peak overlap. Similar results were seen for the peak distribution for amino acids (**Fig. 4A**) and coenzymes (**Fig. 4D**). The expanded spectral regions at 0–5 ppm (**Fig. 4B**) and 6.5–8.5 ppm (**Fig. 4C**) further demonstrate the high degree of peak overlap for the amino acids at a single pH of 6.5. A similar result was obtained with the phosphorous compounds. The expanded spectral regions of 1–3 ppm (**Fig. 4E**), 3–4 ppm (**Fig. 4F**), 4–5 ppm (**Fig. 4G**), and 5–11 ppm (**Fig. 4H**) at a pH of 6.5 again highlights the spectral crowding that occurs due to the variability in chemical shifts and linewidths. This problem would be further exasperated if the data from the amino acids and phosphorous compounds were combined into a single bubble plot.

The spectral region with the highest peak density is 2–4 ppm. Accordingly, this spectral crowding would likely cause difficulties in identifying peak splitting patterns, which would hinder metabolite assignments as individual peaks lose their unique, identifying characteristics. As demonstrated in **Fig. 3A-I**, pH can alter apparent peak splitting patterns, peak locations, and peak resolutions. Factoring in these variations will only worsen peak crowding and overlap, and the challenges of annotating 1D  $^1\text{H}$  NMR spectra becomes more difficult, if not impossible to achieve. It should be noted that the amino acid mixture (**Fig. 4A-C**) contains only 21 compounds, the coenzyme phosphorous mixture (**Fig. 4D**) contains only 6 compounds, and the full phosphorous compound mixture (**Fig. 4E-H**) contains only 33 compounds. The size of these mixtures is significantly smaller than metabolomics samples from biofluids or cell/tissue extracts, which can have upwards of a hundred detectable metabolites.

### 3.6. Matrix effects in known mixtures

The Chenomx batch fit algorithm, the 700 MHz Chenomx metabolite library (338 compounds), and two individualized metabolite libraries (*i.e.*, amino acids and phosphorous containing compounds) were used to demonstrate the challenges of assigning metabolites from complex mixtures, and to evaluate the robustness of an automated metabolite assignment approach without accounting for chemical shift variability. The two individualized metabolite libraries were derived from the complete 700 MHz Chenomx metabolite library. Please note that six phosphorous containing compounds that included inosine monophosphate, thymidine monophosphate, glucose 1,6-bisphosphate, dihydroxyacetone phosphate, O-phospho-L-tyrosine, and cyclophosphamide compounds were not available in the 700 MHz Chenomx metabolite library. So, the individualized library for the phosphorous containing compounds only contained 27 reference spectra instead of 33. Importantly, the 6 compounds missing from the metabolite reference libraries were not included in the calculations of true positives, false positives, and false negative rates. These three metabolite libraries were then used to assign the 1D  $^1\text{H}$  NMR spectra obtained for our known amino acid (**Fig. 5A**) and phosphorous containing compounds (**Fig. 5B**) mixtures. The complete list of true positives, false positives, and false negative rates are provided in **Table S4**.

For the amino acid mixture of 21 compounds at pH 7.0 and 298 K, the batch fit algorithm assigned a total of 90 metabolites. Of these 90 metabolites, 20 % (18) were true positives, 80 % (72) were false positives. Three amino acids (14 %), asparagine, cysteine, and histidine, were false negatives. Repeating the batch assignment for the five pHs and four temperatures yielded an average true positive rate of  $17.4 \pm 0.2$  %. The batch assignment of the 1D  $^1\text{H}$  NMR spectra for the amino acid mixture was repeated with the individualized metabolite library containing only the 21 amino acids present in the mixture. A true positive rate of 81 % was achieved for the amino acid mixture at pH 7.0 and 298 K and an average true positive rate of  $73 \pm 1.1$  % was achieved across the



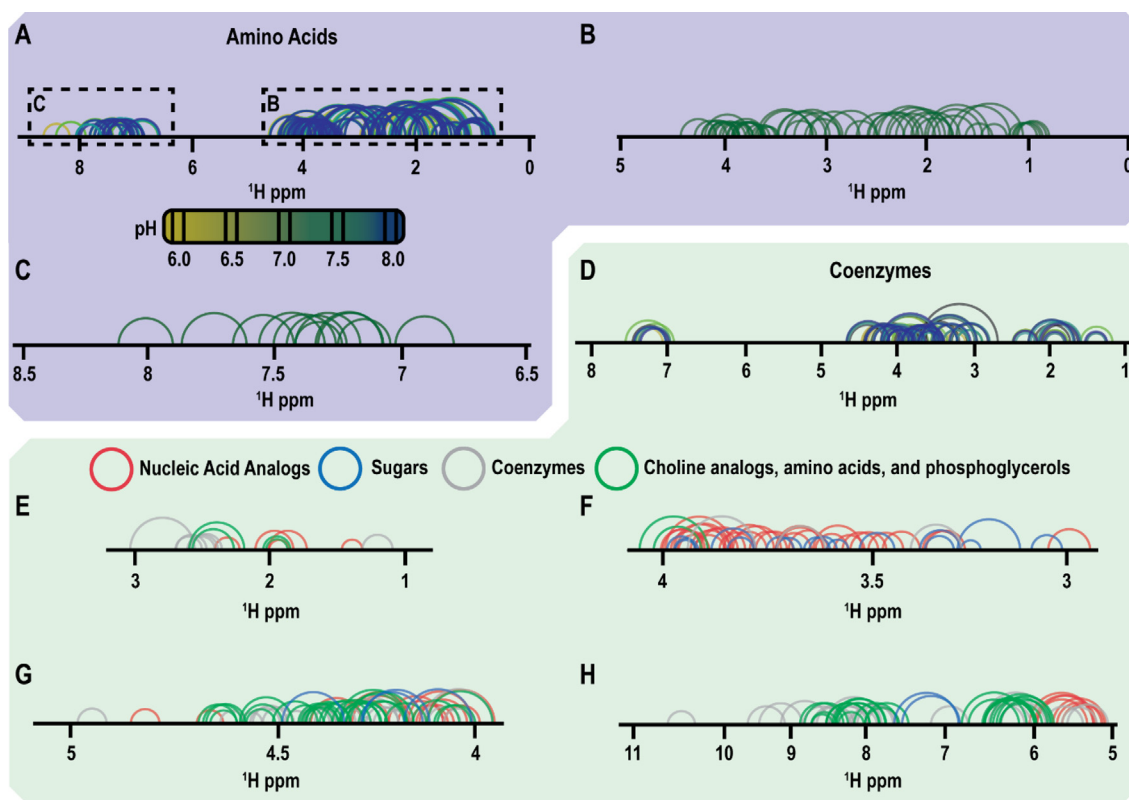
**Fig. 3.** NMR Spectral Variations as a Function of pH and Metabolite Group.  $1D^1H$  NMR stack plots and a corresponding line plot showing the distribution of peaks compared to pH 7 for a select set of amino acids: (A) arginine (ARG), (B) histidine (HIS), (C) lysine (LYS), and phosphorylated compounds: (D) pyridoxal 5-phosphate (P5P), (E) fructose 6-phosphate (F6P), (F) glucose 1,6-bisphosphate (GBP), (G) glucose 6-phosphate (G6P), (H) cytidine monophosphate (CMP), and (I) thymidine monophosphate (TMP).  $1D^1H$  spectra traces and associated circles are colored according to pH: 6.0 (yellow), 6.5 (green), 7.0 (black), 7.5 (teal), and 8.0 (blue). The region of the  $1D^1H$  spectra colored grey in I corresponds to the peaks that merge with the  $H_\gamma$  peak of TMP. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

entire set of pH and temperature conditions. Using either the 700 MHz Chenomx metabolite library or the amino acid library, the true positive assignment rate decreased as pH varied from 7.0. Notably, temperature deviations from 298 K impacted true assignment rates far less. Four amino acids (*i.e.*, glycine, phenylalanine, serine, and tryptophan) were correctly assigned 100 % of the time regardless of the pH or temperature condition. Conversely, histidine, glutamine, cysteine, and cystine were correctly assigned less than 50 % of the time. Cysteine failed to be assigned at all by the batch fit algorithm, and histidine was only assigned once. This outcome is consistent with the chemical shift and linewidth variations where cysteine and histidine showed consistently larger chemical shift deviations from pH 7.0 (Figs. 2F-I, 3B). The expanded spectral regions in Fig. 5 demonstrates where Chenomx accurately assigned (dashed line) and incorrectly assigned (solid line) amino acids within the mixture. The red spectral line shows the sum fit of the reference spectra for the metabolites identified by Chenomx that match the black experimental  $1D^1H$  NMR spectrum. The higher intensity of the fit line (red) compared to the true spectral intensity (black) is indicative of overfitting by the software.

The  $1D^1H$  NMR spectra for the mixtures of phosphorous containing compounds were assigned using the same 700 MHz Chenomx metabolite library and then the individualized library of phosphorous containing compounds. The batch fit algorithm performed significantly worse for the phosphorylated compounds compared to the amino acids. The best results were for glucose-

1-phosphate, which was correctly assigned for all conditions except pH 6.0, and for glucose 6-phosphate that was correctly assigned in all spectra except at pH 6.0 and 8.0. The group 2 mixture of phosphorylated sugars at pH 7.0 and 298 K produced only a 7.7 % (2) true positive rate, a 76.9 % (20) false positive rate, and a 15.4 % (4) false negative rate. The automated assignments improved when the targeted phosphorous containing compound library was used. Chenomx correctly assigned 50 % of the compounds with 3 true positives, 3 false negatives, and zero false positives. Overall, the four mixtures of phosphorous containing compounds had true positive rates predominantly less than 50 % for all pH and temperature conditions. The group 2 and 4 mixtures performed slightly better with an average true positive rate of  $44.4\% \pm 14.2\%$  and  $43.3\% \pm 17.6\%$ , respectively. The poor results may be partially attributed to the greater peak crowding in the sugar region of the NMR spectra and the lack of chemical shift dispersion. The expanded spectral region shown in Fig. 5B clearly demonstrates the poor fit between the red spectral line comprising the sum fit of reference metabolite spectra identified by Chenomx that match the black experimental  $1D^1H$  NMR spectrum. The random fluctuations in the true positive rates across pH and temperature were correlated with the chemical shift variability detailed in Figs. 2 and S2.

Biofluid samples were used to further demonstrate the challenges that chemical shift variability imposes on an automated approach to annotate complex  $1D^1H$  NMR spectra. The NMR spectra were fit automatically then manually triaged using the



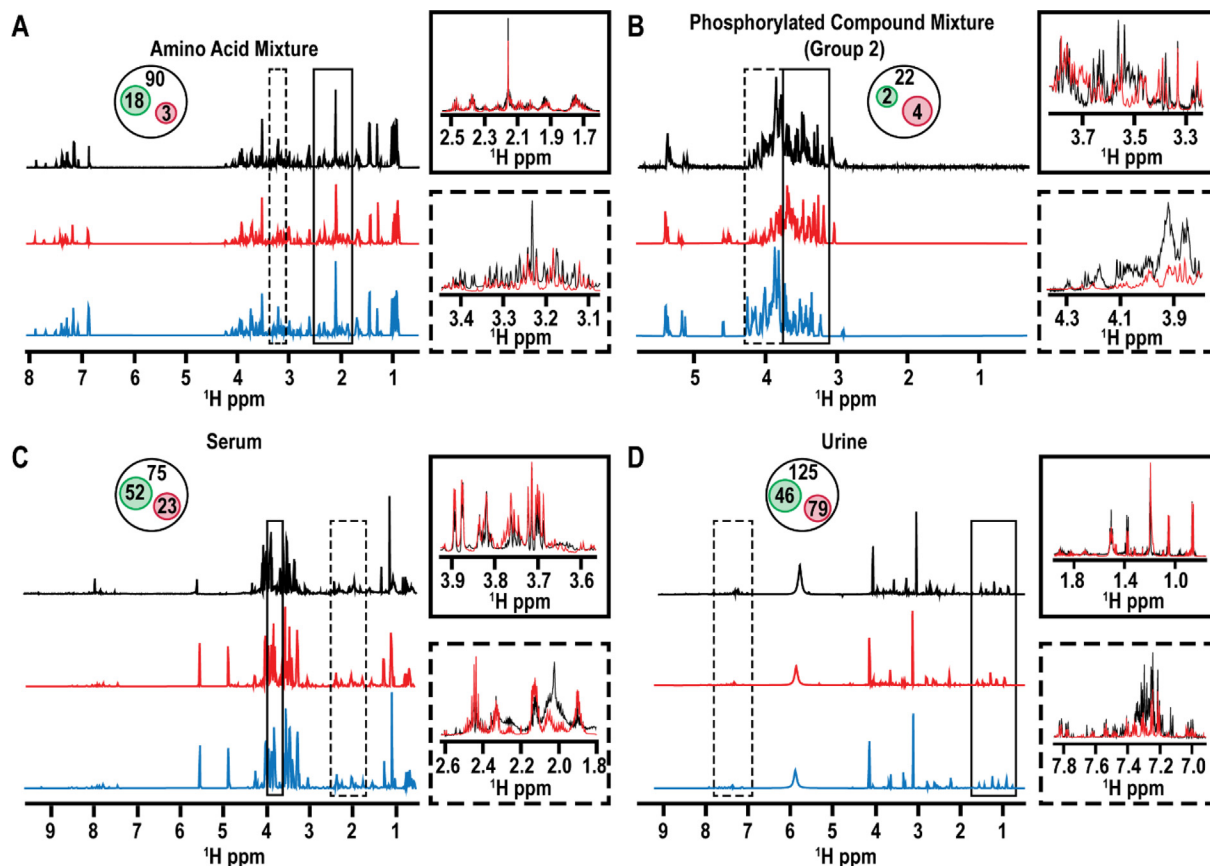
**Fig. 4. Spectral Crowding Due to pH-Induced Linewidth Changes in Complex Metabolic Mixtures.** Bubble plots with circle diameters corresponding to peak width and color corresponding to pH 6.0 (yellow), 6.5 (green), 7.5 (teal), 8.0 (blue) for (A) amino acids and (D) coenzymes. Expanded views of the amino acid bubble plots corresponding to the (B) 1.0 to 5.0 ppm region and the (C) 6.50 to 8.5 ppm region. (E–H) Expanded views of the bubble plots corresponding to the four groups of phosphorylated compounds consisting of nucleic acid analogs (red), sugars (blue), coenzymes (gray), and cholines, amino acids and glycerols (green). Expanded views correspond to the (E) 1 to 3 ppm region, (F) the 3 to 4 ppm region, (G) the 4 to 5 ppm region, and the (H) 5 to 11 ppm region. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

700 MHz Chenomx metabolite library. For the 1D  $^1\text{H}$  NMR spectra obtained from the serum sample shown in Fig. 5C, an automated assignment fit a total of 75 compounds where a manual triage reduced the number of compounds that fit the spectra to 52. The dashed box highlights an expanded spectral region with a poor fit from the automated assignments, while the solid box highlights an expanded region that appears to be well fit by the software. Nevertheless, the solid box includes a crowded region that is still overfit by the software while the dashed boxed region appears underfit. For the 1D  $^1\text{H}$  NMR spectra obtained from the urine sample shown in Fig. 5D, an automated assignment fit a total of 125 compounds where a manual triage reduced the number of compounds that fit the spectra to 46. The expanded spectral region highlighted by a solid box appears to be well fit by the software but upon closer inspection shows the region is actually overfit while the dashed region is again underfit. Overall, the automated assignments falsely fit numerous compounds because it can't accurately account for the variability in both peak position and linewidths, which is consistent with a similar conclusion made by Tredwell et al (2011) [16]. Most assignment software can only make minor adjustments to account for peak fluctuations so the accurate annotation of a 1D  $^1\text{H}$  NMR spectrum still requires a manual intervention [8]. It is important to note that the results reported with the Chenomx batch fit algorithm are not necessarily representative of all metabolomics assignment software. Instead, our analysis was intended to highlight the high failure rate that is likely to occur if chemical shift and linewidth variability are not properly accounted for by either a manual or automated assignment protocol.

#### 4. Conclusion

While the automated assignment of 1D  $^1\text{H}$  NMR spectra is the ideal scenario for metabolomics analyses, it remains fraught with problems. Variations in pH, temperature, ionic strength, buffer, mixture composition, sample preparation, and magnetic field strength, among other issues, pose serious challenges when annotating an experimental NMR spectrum with a database of reference spectra. Current databases are data limited and do not account for various environmental factors that contribute to peak position uncertainties. We have shown that the accuracy of an automated assignment decreases when peak variability is not adequately accounted for, which also confounds a manual assignment [16]. Adherence to pH, temperature, ionic strength, and magnetic field strengths that are consistent with the conditions used to collect the NMR reference data used by automated assignments software improves assignment accuracy, but of course it doesn't account for chemical shift variabilities due to other factors like differences in mixture compositions or metabolite concentrations [9]. The expansion of reference databases to include matrix and sample condition effects would allow for improved data analyses. Specifically, knowing the *true* chemical shift uncertainty associated with each individual NMR resonance or functional group would precisely define search parameters (chemical shift range) to allow for a reliable match between experimental and reference peak while avoiding overfitting of the data [8]. To partly address this need, we experimentally measured the chemical shift variability for 54 common metabolites using modest differences in pH (6 to 8) and temperature (290 K to 308 K) as a surrogate for the natural variability





**Fig. 5. The Accuracy of Automated NMR Assignments are Affected by Chemical Shift Variabilities.** 1D  $^1\text{H}$  NMR stack plots of (A) the amino acid mixture, (B) a phosphorylated compound mixture (Group 2), and human (C) serum and (D) urine samples. Experimental 1D  $^1\text{H}$  NMR spectra are colored black. The spectral line comprising the sum fit of the reference NMR spectra for the metabolites identified by Chenomx are colored red. The manually assigned 1D  $^1\text{H}$  NMR spectra are colored blue. The solid and dashed boxes overlaid onto the 1D  $^1\text{H}$  NMR spectra identify the expanded spectral regions shown to the right of each spectrum. The solid box identifies the expanded spectral region that matches well to the sum of reference spectra, while the dashed box identifies the spectral region that was poorly fitted to the sum of reference spectra. The Venn diagrams summarize the total number of metabolites (white circle), the correctly assigned metabolites (green circle), and the incorrectly assigned metabolites (red circle) identified by Chenomx. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

encountered in complex metabolomics mixtures. It is important to note that other factors such as differences in ionic strength, buffers, metabolite concentrations, and mixture composition, will also modulate chemical shifts. Our analysis of the pH and temperature dependency of chemical shifts is only the first step to completely characterizing the true range of chemical shift errors. Incorporating known chemical shift variability on a per resonance basis into an automated metabolite assignment software may improve the overall accuracy for one of the most crucial aspects of a metabolomics project [8].

#### Authors Contributions

RP conceived of the study. FB, HR, HN, and DF performed the experiments; FB, HR, and RP designed the experiments; FB, HR, and RP analyzed the data and wrote the manuscript.

#### Data availability

Data will be made available on request.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgments

We would like to thank the NIH NeuroBioBank, the University of Maryland Brain and Tissue Bank, and the Saunders Medical Center for providing the serum and urine samples used in this study. We would also like to thank Dr. Martha Morton, Director of the Molecular Analysis and Characterization Facility within the Department of Chemistry at the University of Nebraska-Lincoln for her assistance in acquiring NMR data. This material is based upon work supported by the National Science Foundation under Grant Number (1660921). This work was supported in part by funding from the Nebraska Center for Integrated Biomolecular Communication (P20 GM113126, NIGMS). The research was performed in facilities renovated with support from the National Institutes of Health (RR015468-01). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

#### Appendix A. Supplementary material

The following are the Supplementary data to this article: An Excel file containing: **Table S1**: List of compounds and mixture compositions; **Table S2**: Complete List of Chemical Shift Changes for the 21 Amino Acids; **Table S3**: Complete List of Chemical Shift Changes for the 33 Phosphorus Containing Compounds; and

**Table S4:** List of True Positives, False Positives, and False Negative Rates. A PDF file containing: **Figure S1:** Chemical Shift Variation in Amino Acids; **Figure S2.** Chemical Shift Variation in Phosphorous Containing Compounds; and **Figure S3.** Automated 1D  $^1\text{H}$  NMR Phosphorous Mixtures. Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jmr.2022.107335>.

## References

- [1] G.A. Nagana Gowda, D. Raftery, Can NMR solve some significant challenges in metabolomics?, *J Magn. Reson.* 260 (2015) 144–160.
- [2] A.-H.-M. Emwas, The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research, in: J.T. Bjerrum (Ed.), *Metabonomics: Methods and Protocols*, Springer, New York, New York, NY, 2015, pp. 161–193.
- [3] A.H. Emwas, R. Roy, R.T. McKay, L. Tenori, E. Saccenti, G.A.N. Gowda, D. Raftery, F. Alahmari, L. Jaremko, M. Jaremko, D.S. Wishart, NMR spectroscopy for metabolomics research, *Metabolites* 9 (2019).
- [4] D.D. Marshall, R. Powers, Beyond the paradigm: Combining mass spectrometry and nuclear magnetic resonance for metabolomics, *Prog. Nucl. Magn. Reson. Spectrosc.* 100 (2017) 1–16.
- [5] J.L. Markley, R. Brüschweiler, A.S. Edison, H.R. Eghbalnia, R. Powers, D. Raftery, D.S. Wishart, The future of NMR-based metabolomics, *Curr. Opin. Biotechnol.* 43 (2017) 34–40.
- [6] D.S. Wishart, Quantitative metabolomics using NMR, *TrAC, Trends Anal. Chem.* 27 (2008) 228–237.
- [7] G.A. Nagana Gowda, D. Raftery, NMR-based metabolomics, *Adv Exp Med Biol* 1280 (2021) 19–37.
- [8] T. Vu, Y. Xu, Y. Qiu, R. Powers, Shifting-corrected regularized regression for  $^1\text{H}$  NMR metabolomics identification and quantification, *Biostatistics* (2022).
- [9] P.G. Takis, H. Schäfer, M. Spraul, C. Luchinat, Deconvoluting interrelationships between concentrations and chemical shifts in urine provides a powerful analysis tool, *Nat. Commun.* 8 (2017) 1662.
- [10] A.C. Dona, M. Kyriakides, F. Scott, E.A. Shephard, D. Varshavi, K. Veselkov, J.R. Everett, A guide to the identification of metabolites in NMR-based metabonomics/metabolomics experiments, *Comput. Struct. Biotechnol. J.* 14 (2016) 135–153.
- [11] O.A. Snytnikova, A.A. Khlichkina, R.Z. Sagdeev, Y.P. Tsentalovich, Evaluation of sample preparation protocols for quantitative NMR-based metabolomics, *Metabolomics* 15 (2019) 84.
- [12] E.K. Nawrocka, M. Urbańczyk, K. Koziński, K. Kazimierczuk, Variable-temperature NMR spectroscopy for metabolite identification in biological materials, *RSC Adv.* 11 (2021) 35321–35325.
- [13] F. Bhinderwala, P. Evans, K. Jones, B.R. Laws, T.G. Smith, M. Morton, R. Powers, Phosphorus NMR and its application to metabolomics, *Anal. Chem.* 92 (2020) 9536–9545.
- [14] R. Spicer, R.M. Salek, P. Moreno, D. Cañueto, C. Steinbeck, Navigating freely-available software tools for metabolomics analysis, *Metabolomics* 13 (2017) 106.
- [15] K. O'Shea, B.B. Misra, Software tools, databases and resources in metabolomics: updates from 2018 to 2019, *Metabolomics* 16 (2020) 36.
- [16] G.D. Tredwell, V. Behrends, F.M. Geier, M. Liebeke, J.G. Bundy, Between-person comparison of metabolite fitting for NMR-based quantitative metabolomics, *Anal. Chem.* 83 (2011) 8683–8687.
- [17] D.S. Wishart, A. Guo, E. Oler, F. Wang, A. Anjum, H. Peters, R. Dizon, Z. Sayeeda, S. Tian, B.L. Lee, M. Berjanskii, R. Mah, M. Yamamoto, J. Jovel, C. Torres-Calzada, M. Hiebert-Giesbrecht, V.W. Lui, D. Varshavi, D. Varshavi, D. Allen, D. Arndt, N. Khetarpal, A. Sivakumaran, K. Harford, S. Sanford, K. Yee, X. Cao, Z. Budinski, J. Liigand, L. Zhang, J. Zheng, R. Mandal, N. Karu, M. Dambrova, H.B. Schiöth, R. Greiner, V. Gautam, HMDB 5.0: the human metabolome database for 2022, *Nucleic Acids Res* 50 (2022) D622–D631.
- [18] E.L. Ulrich, H. Akutsu, J.F. Doreleijers, Y. Harano, Y.E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, E. Nakatani, C.F. Schulte, D.E. Tolmie, R. Kent Wenger, H. Yao, J.L. Markley, BioMagResBank, *Nucleic Acids Res.* 36 (2008).
- [19] F. Bhinderwala, S. Lei, J. Woods, J. Rose, D.D. Marshall, E. Riekeberg, A.D.L. Leite, M. Morton, E.D. Dodds, R. Franco, R. Powers, *Metabolomics Analyses from Tissues in Parkinson's Disease*, *Methods Mol. Biol. (Clifton,N.J.)* (2019, 1996,) 217–257.
- [20] F. Bhinderwala, R. Powers, NMR metabolomics protocols for drug discovery, *Methods Mol. Biol.* 2037 (2019) 265–311.
- [21] B.D. Nguyen, X. Meng, K.J. Donovan, A.J. Shaka, SOGGY: Solvent-optimized double gradient spectroscopy for water suppression. A comparison with some existing techniques, *J. Magnetic Reson.* 184 (2007) 263–274.
- [22] M.W. Maciejewski, A.D. Schuyler, M.R. Gryk, P.R. Moraru II, E.L. Romero, H.R. Ulrich, M. Eghbalnia, F. Livny, J.C.H. Delaglio, NMRbox: A resource for biomolecular NMR computation, *Biophys. J.* 112 (2017) 1529–1534.
- [23] B. Worley, R. Powers, Deterministic multidimensional nonuniform gap sampling, (2015).
- [24] V.Y. Orekhov, I. Ibraghimov, M. Billeter, Optimizing resolution in multidimensional NMR by three-way decomposition, *J. Biomol. NMR* 27 (2003) 165–173.
- [25] A. Vignoli, V. Ghini, G. Meoni, C. Licari, P.G. Takis, L. Tenori, P. Turano, C. Luchinat, High-throughput metabolomics by 1D NMR, *Angew. Chem. Int. Ed. Engl.* 58 (2019) 968–994.
- [26] A. Amberg, B. Riefke, G. Schlotterbeck, A. Ross, H. Senn, F. Dieterle, M. Keck, NMR and MS Methods for Metabolomics, *Methods Mol. Biol.* 1641 (2017) 229–258.