SHIMADZU
Excellence in Science

Quadrupole Time-of-Flight
Liquid Chromatograph Mass Spectrometer

LCMS-9030

Effortless Performance

Conduct Critical Qualitative and Quantitative Analysis with Genuine Confidence and Ease

Shimadzu's research-grade **LCMS-9030 quadrupole time-of-flight (Q-TOF) mass spectrometer** combines the engineering DNA from our proven triple quadrupole (LC-MS/MS) platform with powerful, new TOF architecture to transform high mass accuracy workflows. The result is a system that delivers high-resolution, accurate-mass detection with incredibly fast data acquisition rates.

Learn more at **www.ssi.shimadzu.com**

Shimadzu Scientific Instruments Inc., 7102 Riverwood Dr., Columbia, MD 21046, USA

Received: 30 November 2018 | Revised: 27 March 2019 | Accepted: 17 April 2019

DOI: 10.1002/rcm.8469

RESEARCH ARTICLE

WILEY Rapid Communications in Mass Spectrometry

Compositional characterization of complex protopeptide libraries *via* triboelectric nanogenerator Orbitrap mass spectrometry

Marcos Bouza[1,2] | Anyin Li[1,2] | Jay G. Forsythe[3] | Anton Petrov[1,2] | Zhong Lin Wang[4,5] | Facundo M. Fernández[1,2]

[1] NSF/NASA Center for Chemical Evolution, Atlanta, GA 30033, USA

[2] School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA 30332, USA

[3] Department of Chemistry and Biochemistry, College of Charleston, Charleston, SC 29424, USA

[4] School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

[5] Beijing Institute of Nanoenergy and Nanosystems, Chinese Academy of Sciences, Beijing 100083, China

Correspondence
F. M. Fernández, NSF/NASA Center for Chemical Evolution.
Email: facundo.fernandez@chemistry.gatech.edu

Present Address
Anyin Li, Department of Chemistry, University of New Hampshire, Durham, NH 03824, USA

Funding information
NSF/NASA, Grant/Award Number: CHE-1504217

Rationale: Understanding of the molecular processes that led to the first biomolecules on Earth is one of the key aspects of origins-of-life research. Depsipeptides, or polymers with mixed amide and ester backbones, have been proposed as plausible prebiotic precursors for peptide formation. Chemical characterization of depsipeptides in complex prebiotic-like mixtures should benefit from more efficient ion sources and ultrahigh-resolution mass spectrometry (UHR-MS) for elemental composition elucidation.

Methods: A sliding freestanding (SF) Triboelectric Nanogenerator (TENG) was coupled to glass nanoelectrospray emitters for the analysis of a depsipeptide library created using 11 amino acids and 3 alpha-hydroxy acids subjected to environmentally driven polymerization. The TENG nanoelectrospray ionization (nanoESI) source was coupled to an UHR Orbitrap mass spectrometer operated at 1,000,000 resolution for detecting depsipeptides and oligoesters in such libraries. Tandem mass spectrometry (MS/MS) experiments were performed on an Orbitrap Q-Exactive mass spectrometer.

Results: Our previous proteomics-like approach to depsipeptide library characterization showed the enormous complexity of these dynamic combinatorial systems. Here, direct infusion UHR-MS along with *de novo* sequencing enabled the identification of 524 sequences corresponding to 320 different depsipeptide compositions. Van Krevelen and mass defect diagrams enabled better visualization of the chemical diversity in these synthetic libraries.

Conclusions: TENG nanoESI coupled to UHR-MS is a powerful method for depsipeptide library characterization in an origins-of-life context.

1 | INTRODUCTION

Understanding how molecular and environmental processes led to the origins of life on Earth is one of the greatest scientific questions of our time.[1] With decades of research since the original Miller-Urey experiment,[2] the non-enzymatic origin of biopolymers such as nucleic acids, peptides, and glycans has been studied extensively,[3-5] and it is likely that cyclic dry-wet environments were central to their formation.[6] Of all biopolymers, proteins play major roles in terms of enzymatic catalysis, membrane transport, signaling, and energy conversion.[7] Therefore, understanding the mechanisms involved in abiotic peptide and protein synthesis is vital to developing a plausible model for the origins of life.

Both endogenous[2] and exogenous[8] sources of amino acids (AAs), the building blocks of proteins, have been identified as contributors to the early-Earth chemical inventory. Plausible prebiotic routes for

polymerization of such AAs, however, have been less understood until recently.[9] AA polymerization is thermodynamically unfavorable in aqueous solution,[10] and often produces cyclic compounds that act as thermodynamic "traps".[11] Mechanisms ranging from chemical activation[12] to heating with inorganic catalysts[3] have been proposed for abiotic peptide synthesis, but have only yielded short chain lengths,[13] or have been demonstrated for only a few AA types.[14,15]

We recently proposed a prebiotic mechanism for the formation of peptides involving depsipeptides, polymers with mixed amide and ester backbones.[16] In this process, depsipeptides are formed by the condensation of α-AAs and α-hydroxy acids (AHAs), also believed to have been present on the prebiotic Earth.[17] We previously showed depsipeptides form *via* ester-amide exchange reactions in simulated day-night environments that "ratchet" the polymer chain to progressively more amide-enriched species. The kinetics of such chemistry is now being explored, with the goal of finding optimum conditions that will lead to functional protopeptides.[18]

Condensation polymerization of AA and AHA monomers leads to complex combinatorial depsipeptide libraries with large numbers of esters and depsipeptides. In recent work, we explored the complexity of such systems by *de novo* sequencing in combination with liquid chromatography (LC), ion mobility (IM), and high-resolution mass spectrometry (HR-MS).[19] Three reactions, one AHA and three AAs each, produced libraries with more than 600 unique depsipeptide sequences in total. It is likely that hundreds more species were undetected due to hydrolysis during the LC step, insufficient ion generation efficiency, and/or limited resolution of the time-of-flight mass analyzer.

Ultrahigh-resolution (UHR)-MS has become a key tool in the elemental composition elucidation of complex organic samples.[20] It has seen widespread use in petroleomics,[21] natural organic matter analysis,[22] and proteomics,[23] but much less in the field of prebiotic chemistry, with one notable exception being the in-depth investigation of organics in the Murchison meteorite.[24] Such experiments, typically requiring Fourier Transform Ion Cyclotron Resonance (FT-ICR) instrumentation, may now become more commonplace due to advances in UHR Orbitrap MS.[25]

Sensitive characterization of combinatorial protopeptide mixtures also requires improvements in ion generation. Nanoelectrospray ionization is a more suitable ion source for direct infusion experiments, with reduced ion suppression and softer ionization expected to produce wider sequence coverage compared with standard electrospray analysis. Added to that, in recent work we have introduced triboelectric nanogenerators (TENGs) as a powerful, sensitive, and reproducible alternative to generate ions via nanoelectrospray.[26] TENGs are power sources that transform mechanical motion into electricity,[27] with one of their most attractive features being the on-demand generation of discrete amounts of charge in the nano-Coulomb (nC) range during each operation cycle. A TENG in combination with MS enables zeptomole detection limits,[26] which could prove essential in observing species that may be missed with less sensitive ionization approaches.

Here, we explore the chemical diversity of the most complex protopeptide mixture studied to date, consisting of 11 AAs and 3 AHAs, subject to four cycles of day-night polymerization condensation. We characterize this mixture using TENG nanoelectrospray ionization combined with UHR Orbitrap MS at 1,000,000 resolution (m/Δm at full width at half maximum or FWHM). Since depsipeptides are not present in protein databases, modifications were included on a *de novo* sequencing software for depsipeptide sequences annotation. This approach resulted in the identification of hundreds of sequences without the need for lengthy chromatographic separations.

2 | EXPERIMENTAL

2.1 | Materials

L-Alanine (A), glycine (G), L-histidine (H), L-leucine (L), L-methionine (M), L-proline (P), L-glutamine (Q), L-arginine (R), L-serine (S), L-threonine (T), L-valine (V), L-lactic acid (a), L-malic acid (d) and 2-hydroxyisobutyric acid or 2-HIBA (z) were obtained from Sigma-Aldrich. The AAs were selected in terms of water solubility, and the majority are considered potentially prebiotic compounds.[28] The three AHAs selected (L-lactic acid, L-malic acid and 2-hydroxyisobutyric acid) are all thought to be prebiotic.[29] Structures for all these building blocks, as well as a generic peptide and depsipeptide structures, are provided in Figures S1 and S2 (supporting information). Ultrapure deionized water (18.2 MΩ cm) was obtained from a Barnstead Nanopure Diamond system (Van Nuys, CA, USA). Methanol (LC/MS grade) was obtained from Sigma-Aldrich.

2.2 | Depsipeptides synthesis through dry–wet cycling

Oligomers were prepared in 1.5 mL Eppendorf vials with 200 µL aqueous solution containing 10 mM of each AA and 35 mM of each AHA. Following our previous report,[16] a model prebiotic 24-h cycle was applied to nine replicates. Eppendorf tubes were dried-down at 85°C for 18 h with the vial cap open (dry/ "day" phase) and then rehydrated and heated at 65°C for 5.5 h with the vial cap closed (wet/ "night" phase). Samples were divided in three groups that were subjected to 1, 2, and 4 cycles, respectively. Samples were diluted 1:200 using methanol/water (1:1). For depsipeptide naming purposes, AA residues were capitalized (e.g., alanine = A) and the analogous AHAs were in lower case (e. g., lactic acid = a). For 2-HIBA, which does not have a matching proteinogenic amino acid, the lower case letter "z" was used.

2.3 | TENG nanoelectrospray ion generation

A sliding freestanding (SF) TENG power source, similar to the one used in previous work,[26] was utilized for all experiments. Figure 1 presents a schematic of the TENG-UHR Orbitrap setup. The SF-

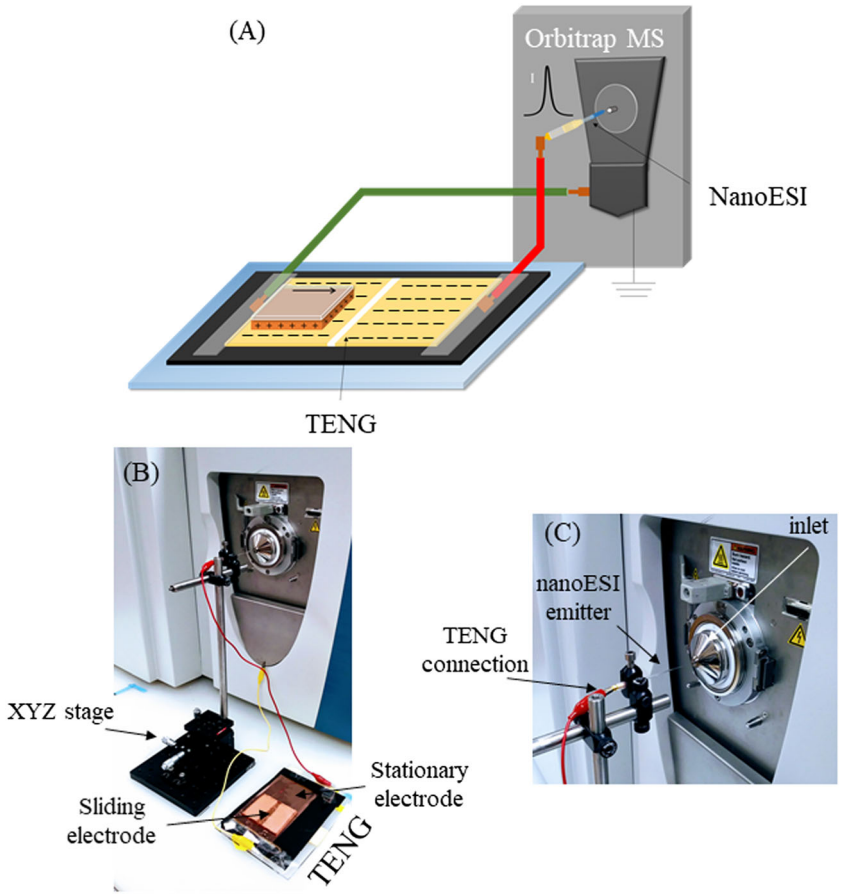

FIGURE 1 TENG UHR-MS setup. A, General scheme of the sliding freestanding TENG power source connected to an Orbitrap mass spectrometer (orange represents copper, yellow fluorinated ethylene propylene). The black arrow represents the direction of movement of the top copper electrode on the static electrode. The pulse represents the flow of generated charges. B, Picture showing the TENG power source coupled to the atmospheric pressure interface inlet of the mass spectrometer. The nanoESI emitter is mounted on a XYZ stage that allows its position to be controlled in respect to the spectrometer inlet. C, Zoomed-in view of the inlet region [Color figure can be viewed at wileyonlinelibrary.com]

TENG device consisted of two separate pieces, the static and sliding electrodes. The static electrode consisted of two $75 \times 60\,mm^2$ rectangles of Cu film deposited onto fluorinated ethylene propylene (FEP) film. These rectangles were separated by a $75 \times 1\,mm^2$ uncoated rectangular section and were mounted on an acrylic support. The movable SF-TENG electrode was made of Cu foil ($55 \times 65\,mm^2$) and was mounted onto a rectangular piece of acrylic board also. To operate the TENG, the movable electrode was placed on top of the static part of the device, which was fixed onto an XYZ stage placed underneath the inlet to the mass spectrometer. Both parts were oriented so that the two metal layers were separated by the FEP layer (Figure 1). The movable electrode was slid between the two positions corresponding to the two Cu-coated regions of the static electrode, with a travel distance of 60 mm. The two electrical contacts are represented as green (ground) and red (connected to the ion source). The TENG power source was actuated manually at a frequency of ~0.5 Hz, generating positive ions when the sliding electrode was moved from the ground contact to the contact connected to the nanoESI emitter, and negative ions when moved in the opposite direction. Glass nanoESI emitters (Econo 12, New Objective) were used in all experiments. Each of the samples (10 µL) was loaded into the emitters using a 10 µL pipette. A Pt wire was used to provide electric contact between the sample solutions loaded on the emitters and the TENG power supply.

2.4 | Mass spectrometry experiments

Synthetic protopeptide libraries were studied using TENG nanoESI in direct infusion mode coupled to a prototype Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Scientific, San Jose, CA, USA) at a resolving power of ~1,000,000. External mass calibration was employed in all cases, as the internal lock mass function was not yet implemented. Results were compared with a Xevo G2 QTOF mass spectrometer (Waters Corp., Manchester, UK), and a Q-Exactive hybrid quadrupole-Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA, USA). All tandem mass spectrometry (MS/MS) experiments were performed on the Q-Exactive. Operational parameters for all three mass spectrometers are provided in the supporting information. Negative ionization mode was used for all MS and MS/MS experiments, as depsipeptides ionize well as $[M-H]^-$ species with little-to-no adduct formation. Positive ion mode experiments were performed to investigate arginine-containing depsipeptides.

An overview of the workflow used for depsipeptide identification is presented in Figure S3 (supporting information). Accurate masses were obtained with the Orbitrap Fusion Lumos Tribrid mass spectrometer. A custom MATLAB algorithm was used for accurate mass matching as described in the supporting information. Accurate mass matching was followed by a custom MS/MS data-dependent acquisition (DDA) method implemented in the Q-Exactive Orbitrap

mass spectrometer, also described in detail in the supporting information. Data produced by DDA MS/MS experiments were processed by using PEAKS 8.0 software to yield depsipeptide and polyester sequences. In order to account for the presence of AHA residues instead of AAs, custom "post-translational modifications" were defined in PEAKS, as also described in the supporting information (Table S1 and Figure S4).

3 | RESULTS AND DISCUSSION

3.1 | Effect of mass resolution

Figure 2A shows the results from the TENG MS analysis of a four-cycle mixture using three different mass spectrometers, each with increasing resolving power. As expected, the QTOF mass spectrometer had the poorest resolution of all three instruments, being unable to resolve

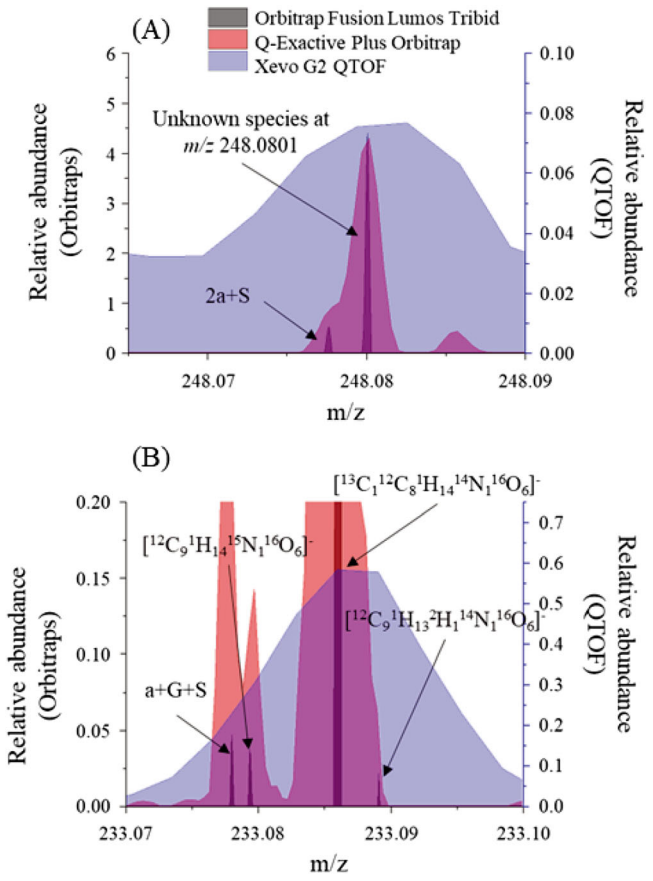

FIGURE 2 Effect of mass spectrometer resolving power on depsipeptide mixture analysis. TENG negative ion mode mass spectra for A) region between *m/z* 248.065 and 248.09 that includes the 2a + S depsipeptide with *m/z* 248.0777, resolved only with the Orbitrap Fusion Lumos Tribrid. B) Signals in the region between *m/z* 233.07 and 233.10, which includes the a + G + S depsipeptide that typically overlaps with isotopes from the species at *m/z* 232.0821 (2a + a or a + z + G). The a + G + S depsipeptide was only resolved with the Orbitrap Fusion Lumos Tribrid [Color figure can be viewed at wileyonlinelibrary.com]

many of the species in the studied samples. In the *m/z* 248.07–248.09 region, for example, the QTOF instrument only detected one broad spectral feature centered around *m/z* 248.0825, whereas more species were resolved by the two Orbitrap instruments. The Q-Exactive instrument, operated at a resolving power of 140,000 FWHM (full width at half maximum), was able to resolve many of the species present, but its resolving power was not sufficient in some cases. As seen in Figure 2A, the 2a + S depsipeptide was only observed as an unresolved shoulder of an unknown, and more abundant, ionic species at *m/z* 248.0801. The Orbitrap Fusion Lumos Tribrid, operated at 1,000,000 resolution, effectively resolved the most challenging cases of spectral overlap. Figure 2A shows a species at *m/z* 248.0777 baseline resolved from the above-mentioned spectral interference, with effective resolving powers of 630,600 and 871,302 (FWHM), respectively. Within a 2 ppm mass error, the composition of this species was assigned to 2a + S (2 lactic acid + serine).

Numerous interferences between the various isotopologues from the various depsipeptides generated by wet–dry cycling were also observed in these very complex mixtures. This is better illustrated by examining the spectral region between 233.075 and 233.09 *m/z*. In this region, the Orbitrap Fusion Lumos Tribrid was able to resolve completely several closely related species. For example, a depsipeptide composed of lactic acid, glycine and serine (a + G + S, *m/z* 233.0774) was resolved from the ^{15}N isotopologue of a depsipeptide with *m/z* of 232.0821 (2a + A and a + z + G, $[M-H]^- = {}^{12}C_9{}^1H_{14}{}^{15}N_1{}^{16}O_6$, 233.0792 *m/z*). The fine isotopic structure of 2a + A and a + z + G, which included the ^{13}C ($^{13}C_1{}^{12}C_8{}^1H_{14}{}^{14}N_1{}^{16}O_6$) and the ^2H isotopologues ($^{12}C_9{}^1H_{13}{}^2H_1{}^{14}N_1{}^{16}O_6$), was resolved also, making spectral assignments more straightforward.

Overall, the higher the instrument resolving power the larger the number of *m/z* values that can be accurately detected, performance that is better described by the mass spectrometer's peak capacity. As proposed by Coombes et al,[30] calculation of the maximum number of peaks (*N*) detectable by a mass spectrometer can be estimated based on the mass range of interest range and the relative error associated with the measurement. Using a *m/z* range of 100–1000, and establishing the relative mass error threshold as the absolute value of the highest mass error for a specific mass spectrometer, the peak capacity of each instrument was calculated. The relative mass errors were $1.98*10e^{-6}$ for the Orbitrap Fusion Lumos Tribrid, $2.65*10e^{-6}$ for the Orbitrap Q-Exactive, and $1.23*10e^{-5}$ for the QTOF, resulting in a maximum *N* of: $1.2*10e^6$, $8.7*10e^5$, and $1.9*10e^5$, respectively. As expected, UHR-MS allowed a greater number of potential species to be resolved, as also reflected in Figure 2.

3.2 | Depsipeptide library characterization

Throughout the various oligomerization conditions tested, it was observed that as the number of day-night cycles increased, so did the number of peptide, oligoester, and depsipeptide species detected. Figure 3 shows the direct infusion TENG nanoESI mass spectrum of a

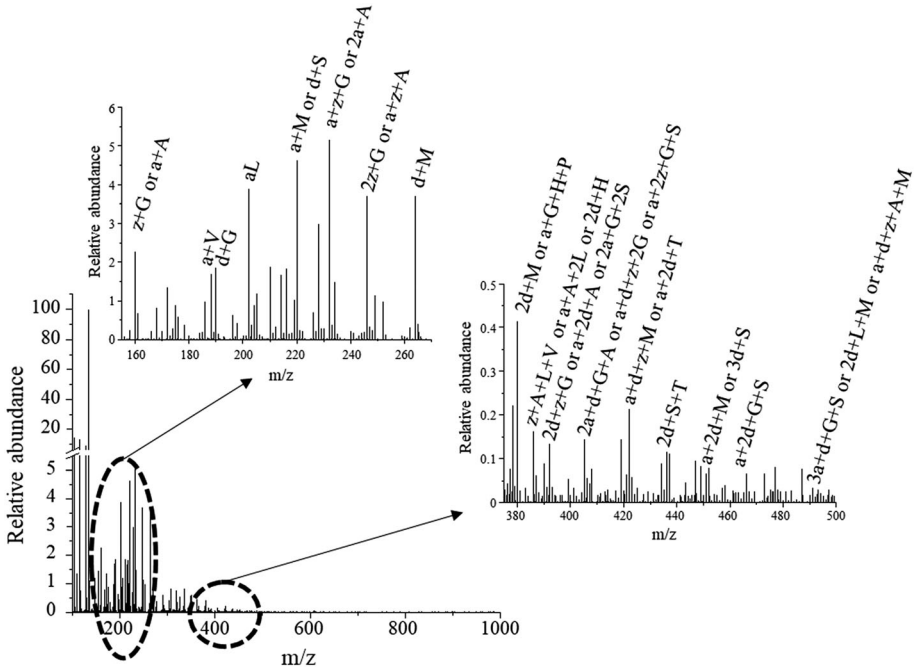

FIGURE 3 Negative ion mode TENG full mass spectrum collected with the Orbitrap Fusion Lumos Tribrid for a four-cycle mixture. As some sequences are isobaric, several candidate species are listed for a number of signals

four-cycle sample, showing the presence of thousands of resolved peaks within the 100–1000 *m/z* range. The total number of depsipeptide and oligoester molecular compositions confirmed by accurate mass Orbitrap MS were 258, within a mass error of ±2 ppm. This value was acceptable for the external mass calibration used with this instrument. As shown in Figure S5 (supporting information), most of the mass errors were between −1 and 1 ppm (78%), with a remaining 22% between 1 and 2 ppm. The more abundant species were at *m/z* = 133.0142 and 103.0400, corresponding to the unreacted or hydrolyzed malic and 2-HIBA $[M-H]^-$ ions. The depsipeptide with the lowest *m/z* value detected was aG at 146.0458. Depsipeptide chain lengths varied from 2 to 7, with the majority between 3 and 4 under the studied conditions. These conditions promote sequence diversity due to the hydrolysis step carried out in between each drying step, and the limiting concentration (10 mM) of each of the monomers used. Longer chains can be achieved with continuous monomer feeding,[31] but such an approach was not utilized in the present study for simplicity.

Once compositional analysis via accurate mass matching was performed, *de novo* sequencing was carried out following the procedures described in the supporting information. Using a custom DDA acquisition approach and custom modifications to monomers in the PEAKS software, 524 unique sequences corresponding to 320 depsipeptide or polyester compositions (Tables S2 and S3, supporting information) were generated. As seen in these tables, a significant variability in N-terminal residue composition was observed; 305 species contained an AHA residue in that position (166 lactic acid, 108 malic acid, and 31 2-HIBA). A total of 219 species contained an AA at the N-terminus, the majority containing glycine, methionine and serine. These species are likely formed by hydrolysis of longer species, as the ester-amide exchange mechanism primarily supports the presence of AHAs in such a position.[16]

Arginine-containing depsipeptides were not detected in negative ion mode analysis, as this residue is protonated at the pH values used during the synthesis (pH 3–4). When four-cycle libraries were analyzed in positive ion mode, species such as aR, dR, and zR were detected, along with histidine-containing depsipeptides, and a handful of peptides such as GG and GGG (Figure S6, supporting information). Trimer and longer arginine-containing depsipeptides were not observed, likely due to "backbiting" leading to intramolecular ring closure,[32] a reaction favored at acidic pH that further inhibits chain elongation.

By far, glycine was the most common AA incorporated into depsipeptide chains, present in more than 50% of the sequences (Figure S7, supporting information). The longest depsipeptides detected (3a + A + 3G, 2a + z + 4G, 2a + 2z + 2G + H) contained at least 2 glycine residues. Sequences with the highest number of isomers (*n* = 5) were all found to be glycine-rich. The base peak in positive mode experiments was that of di-glycine, and more than half of the sequences obtained contained this AA. Glycine is most efficiently incorporated into depsipeptide sequences, leading to highly complex libraries as found here.

3.3 | Depsipeptide mass mapping

Following compositional and sequence analysis, a number of graphical representations of the dataset were investigated. Van Krevelen diagrams,[33] Kendrick plots,[34] and mass defect filter diagrams[35] have

been used to enhance visualization of chemical diversity in complex datasets. To this end, a modified Van Krevelen plot (Figure 4) enabled grouping of detected species based on the number of nitrogen atoms in their composition. Nitrogen atoms ranged from zero, which corresponded to pure oligoesters, to six. As shown in Figure 4, N/C ratios and the detected *m/z* values allowed separation of the detected species into five depsipeptide groups. The appearance of depsipeptide sequences with higher nitrogen content also followed an increase in the number of species with higher nitrogen/hydrogen ratios at higher *m/z* values. This is due to the incorporation of more histidine (H) and glutamine (Q) residues in longer sequences.

Another modified Van Krevelen diagram was used to group depsipeptides as a function of AA content (Figure S8a, supporting information). Depsipeptides containing shorter side chain AAs, such as A, G, or S, typically fell at higher N/C ratios and low *m/z* values for each nitrogen content series. As seen in Figure S8b (supporting information), for each nitrogen-content group, depsipeptides with lower *m/z* values and higher N/C ratios typically contained serine, followed by threonine, methionine, valine, proline, and leucine as the N/C ratio decreased.

Normalized monoisotopic mass defects (NMDs) and normalized isotopic shifts (NIS), calculated by Supplementary Equations 1 and 2 (supporting information), provided additional information useful in describing complex libraries based on their AA content.[35] Changes in elemental compositions resulted in different trends and patterns in the resulting mass defect plots. Figure 5A depicts the NMDs for all detected depsipeptides plotted against their NIS, and color coded by the specific AA present in the sequence. Depsipeptides containing at least one methionine residue and various other residues were differentiated from the rest of the depsipeptide population, as the presence of a sulfur atom results in much higher NIS values even when other AAs are present.

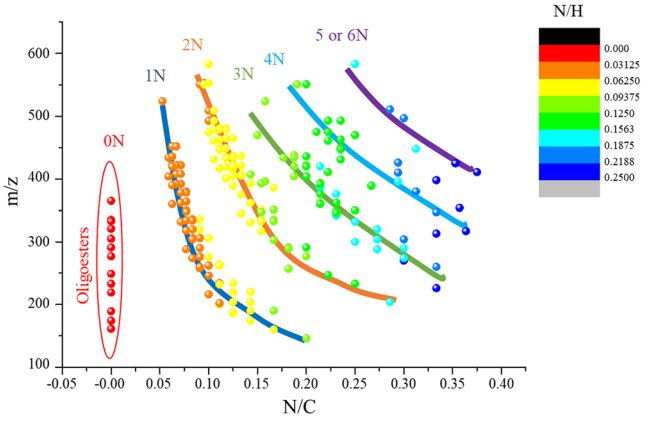

FIGURE 4 Nitrogen content of depsipeptide and polyester species detected following four dry–wet cycles of a mixture containing 3 AHAs and 11 AAs. Experiments were conducted in negative ion mode. Only species detected with the Orbitrap Fusion Lumos Tribrid and confirmed by Q-Exactive MS/MS experiments are shown [Color figure can be viewed at wileyonlinelibrary.com]

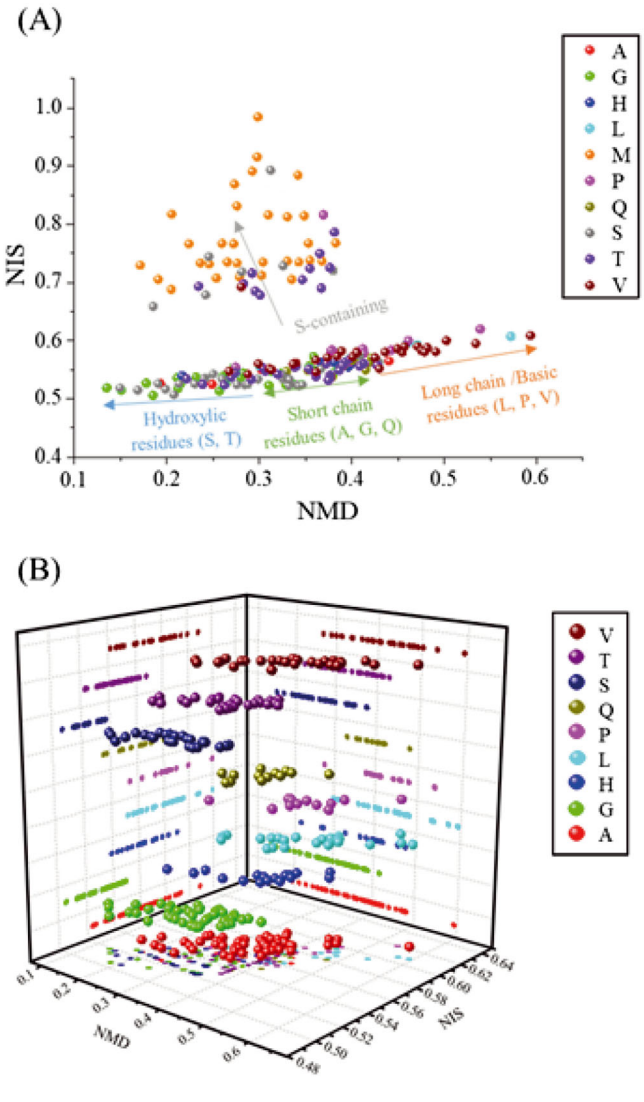

FIGURE 5 Classification of *de novo* identified depsipeptides based on their amino acid content. Normalized monoisotopic mass defect (NMD) *vs.* normalized isotopic shift (NIS) shown for A, all species detected, B, canonical separation in the Z axis of the different non-methionine-containing depsipeptide families as a function of both NMD and NIS values [Color figure can be viewed at wileyonlinelibrary.com]

To investigate trends in non-methionine-containing depsipeptides further, a "virtual" z-axis was added to the plot shown in Figure 5A using a canonical variable to describe the presence of each specific AA in the sequence (Figure 5B). This plot revealed that depsipeptides containing shorter side chains, such as alanine or glycine, grouped at higher NIS and lower NMD (<0.4), while species containing glutamine, histidine, and threonine were characterized by intermediate NMD and NIS values. Depsipeptides with aliphatic side chains, such as leucine, proline, or valine, appeared at NMD values above 0.35. Depsipeptides with hydroxyl residues (e.g. serine) were clustered at lower NMD and higher NIS values.

Mass defect distributions coded by AHA content showed interesting trends also. As shown in Figure S9 (supporting information), malic acid containing sequences clustered at lower NMD values. Acidic AA residues, such as glutamic and aspartic acid, are

known for producing this type of shifts in mass defect plots.[23] As malic acid is a structural analog of aspartic acid, this trend was not surprising. The other two AHAs used, lactic acid and 2-HIBA, that have short side chains, led to NMD values above 0.3. Overall, it was observed that glycine, methionine, and serine tended to co-polymerize with malic acid, and with each other. Other AAs (A, H, L, P, Q) tended to co-polymerize with lactic acid and 2-HIBA.

4 | CONCLUSIONS

We have reported on the performance of a TENG nanoelectrospray ionization source coupled to UHR-MS for the analysis of complex depsipeptide libraries (11 AAs and 3 AHAs). Peptidomics strategies were adapted to take advantage of the appealing aspects of direct infusion analysis, including short analysis times and no need for mobile phase additives or buffers. This approach led to the successful sequencing of more than 500 depsipeptides and/or oligoesters prepared in a single simulated prebiotic mixture. Mass-mapping plots simplified the visualization of depsipeptide libraries composition into different subgroups that included the number of nitrogen atoms in the elemental composition and the relative AA and AHA composition. The libraries characterized in the present work represent a relatively simple system in terms of complexity when compared with a true prebiotic soup, but they do serve as a testing ground to understand the analytical peak capacity needed to examine libraries containing upwards of 20 AAs and their corresponding AHAs. Future applications of TENG UHR-MS involve the coupling to higher resolving power mass spectrometers, such as Fourier Transform Ion Cyclotron Resonance (FTICR) instruments, and the incorporation of ion mobility (IM) gas-phase separations to further decrease spectral congestion. While addition of a chromatographic step should also improve sequence coverage, ester hydrolysis during the liquid phase separation adds unnecessary complications to an already difficult analytical task.

ACKNOWLEDGEMENTS

This work was jointly supported by NSF and the NASA Astrobiology Program, under the NSF Center for Chemical Evolution, CHE-1504217. The authors thank Seema Sharma, Eloy Wouters, Jesse D. Canterbury, and Donna Earley for access to and assistance with Orbitrap Fusion Lumos Tribrid instrumentation.

ORCID

Marcos Bouza https://orcid.org/0000-0002-2601-9776
Jay G. Forsythe https://orcid.org/0000-0003-1041-1113
Facundo M. Fernández https://orcid.org/0000-0002-0302-2534

REFERENCES

1. Lazcano A, Miller SL. The origin and early evolution of life: Prebiotic chemistry, the pre-RNA world, and time. *Cell*. 1996;85(6):793-798.
2. Miller SL. A production of amino acids under possible primitive earth conditions. *Science*. 1953;117(3046):528-529.
3. Fox SW, Harada K. Thermal copolymerization of amino acids to a product resembling protein. *Science*. 1958;128(3333):1214.
4. Oró J. Mechanism of synthesis of adenine from hydrogen cyanide under possible primitive earth conditions. *Nature*. 1961;191(4794):1193-1194.
5. Weber AL. Prebiotic sugar synthesis: Hexose and hydroxy acid synthesis from glyceraldehyde catalyzed by iron(III) hydroxide oxide. *J Mol Evol*. 1992;35(1):1-6.
6. Deamer DW, Georgiou CD. Hydrothermal conditions and the origin of cellular life. *Astrobiology*. 2015;15(12):1091-1095.
7. Alberts B, Johnson A, Lewis J, et al. *Molecular Biology of the Cell*. 4th ed. New York: Garland Science; 2002.
8. Peltzer ET, Bada JL, Schlesinger G, Miller SL. The chemical conditions on the parent body of the Murchison meteorite: Some conclusions based on amino, hydroxy and dicarboxylic acids. *Adv Space Res*. 1984;4(12):69-74.
9. Danger G, Plasson R, Pascal R. Pathways for the formation and evolution of peptides in prebiotic environments. *Chem Soc Rev*. 2012;41(16):5416-5429.
10. Jakubke H-D, Kuhl P, Könnecke A. Basic principles of protease-catalyzed peptide bond formation. *Angew Chem Int Ed*. 1985;24(2):85-93.
11. Parker Eric T, Zhou M, Burton Aaron S, et al. A plausible simultaneous synthesis of amino acids and simple peptides on the primordial earth. *Angew Chem Int Ed*. 2014;53(31):8132-8136.
12. Montalbetti CAGN, Falque V. Amide bond formation and peptide coupling. *Tetrahedron*. 2005;61(46):10827-10852.
13. Rode BM. Peptides and the origin of life. *Peptides*. 1999;20(6):773-786.
14. Rodriguez-Garcia M, Surman AJ, Cooper GJT, et al. Formation of oligopeptides in high yield under simple programmable conditions. *Nat Commun*. 2015;6(1):8385.
15. Le Son H, Suwannachot Y, Bujdak J, Rode BM. Salt-induced peptide formation from amino acids in the presence of clays and related catalysts. *Inorg Chim Acta*. 1998;272(1):89-94.
16. Forsythe JG, Yu S-S, Mamajanov I, et al. Ester-mediated amide bond formation driven by wet–dry cycles: A possible path to polypeptides on the prebiotic earth. *Angew Chem Int Ed*. 2015;54(34):9871-9875.
17. Bada JL. New insights into prebiotic chemistry from Stanley Miller's spark discharge experiments. *Chem Soc Rev*. 2013;42(5):2186-2196.
18. Yu S-S, Krishnamurthy R, Fernandez FM, et al. Kinetics of prebiotic depsipeptide formation from the ester-amide exchange reaction. *PCCP*. 2016;18(41):28441-28450.
19. Forsythe JG, Petrov AS, Millar WC, et al. Surveying the sequence diversity of model prebiotic peptides by mass spectrometry. *Proc Natl Acad Sci U S A*. 2017;114(37):E7652-E7659.
20. Sleighter RL, Hatcher PG. The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter. *J Mass Spectrom*. 2007;42(5):559-574.
21. McKenna AM, Nelson RK, Reddy CM, et al. Expansion of the analytical window for oil spill characterization by ultrahigh resolution mass spectrometry: Beyond gas chromatography. *Environ Sci Technol*. 2013;47(13):7530-7539.
22. D'Andrilli J, Cooper WT, Foreman CM, Marshall AG. An ultrahigh-resolution mass spectrometry index to estimate natural organic matter lability. *Rapid Commun Mass Spectrom*. 2015;29(24):2385-2401.
23. Artemenko KA, Zubarev AR, Samgina TY, Lebedev AT, Savitski MM, Zubarev RA. Two dimensional mass mapping as a general method of

data representation in comprehensive analysis of complex molecular mixtures. *Anal Chem*. 2009;81(10):3738-3745.

24. Schmitt-Kopplin P, Gabelica Z, Gougeon RD, et al. High molecular diversity of extraterrestrial organic matter in Murchison meteorite revealed 40 years after its fall. *Proc Natl Acad Sci U S A*. 2010;107(7):2763-2768.
25. Schmidt EM, Pudenzi MA, Santos JM, et al. Petroleomics via Orbitrap mass spectrometry with resolving power above 1 000 000 at *m/z* 200. *RSC Adv*. 2018;8(11):6183-6191.
26. Li A, Zi Y, Guo H, Wang ZL, Fernández FM. Triboelectric nanogenerators for sensitive nano-coulomb molecular mass spectrometry. *Nat Nanotechnol*. 2017;12(5):481-487.
27. Wang ZL. Triboelectric nanogenerators as new energy technology and self-powered sensors - principles, problems and perspectives. *Faraday Discuss*. 2014;176(0):447-458.
28. Zaia DAM, Zaia CTBV, De Santana H. Which amino acids should be used in prebiotic chemistry studies? *Orig Life Evol Biosph*. 2008;38(6):469-488.
29. Pizzarello S, Wang Y, Chaban GM. A comparative study of the hydroxy acids from the Murchison, GRA 95229 and LAP 02342 meteorites. *Geochim Cosmochim Acta*. 2010;74(21):6206-6217.
30. Coombes KR, Koomen JM, Baggerly KA, Morris JS, Kobayashi R. Understanding the characteristics of mass spectrometry data through the use of simulation. *Cancer Informatics*. 2007;1(1):41-52.
31. Yu S-S, Solano MD, Blanchard MK, et al. Elongation of model prebiotic proto-peptides by continuous monomer feeding. *Macromolecules*. 2017;50(23):9286-9294.
32. Keller M, Kuhn KK, Einsiedel J, et al. Mimicking of arginine by functionalized Nω-carbamoylated arginine as a new broadly applicable approach to labeled bioactive peptides: High affinity angiotensin, neuropeptide Y, neuropeptide FF, and neurotensin receptor ligands as examples. *J Med Chem*. 2016;59(5):1925-1945.
33. Kew W, Blackburn JWT, Clarke DJ, Uhrín D. Interactive van Krevelen diagrams – Advanced visualisation of mass spectrometry data of complex mixtures. *Rapid Commun Mass Spectrom*. 2017;31(7):658-662.
34. Hughey CA, Hendrickson CL, Rodgers RP, Marshall AG, Qian K. Kendrick mass defect spectrum: A compact visual analysis for ultrahigh-resolution broadband mass spectra. *Anal Chem*. 2001;73(19):4676-4681.
35. Sleno L. The use of mass defect in modern mass spectrometry. *J Mass Spectrom*. 2012;47(2):226-236.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Bouza M, Li A, Forsythe JG, Petrov A, Wang ZL, Fernández FM. Compositional characterization of complex protopeptide libraries *via* triboelectric nanogenerator Orbitrap mass spectrometry. *Rapid Commun Mass Spectrom*. 2019;33:1293-1300. https://doi.org/10.1002/rcm.8469