JACS Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/jacsau
Article

A Workflow Enabling the Automated Synthesis, Chain-End Degradation, and Rapid Mass Spectrometry Analysis for Molecular Information Storage in Sequence-Defined Oligourethanes

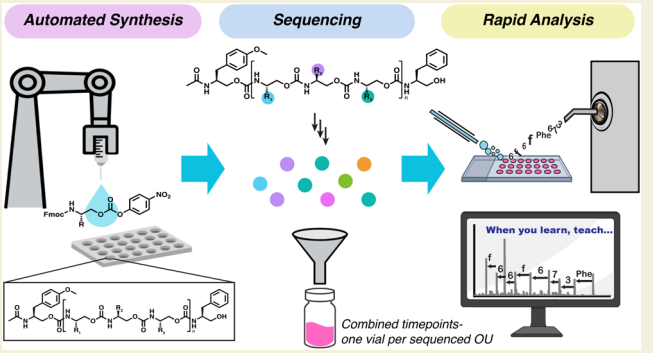
Julia R. Shuluk,[‖] Christopher D. Wight,[‖] James R. Howard, Mary E. King, Sarah R. Moor, Rachel J. DeHoog, Samuel D. Dahlhauser, Livia S. Eberlin,* and Eric V. Anslyn*

Cite This: JACS Au 2025, 5, 1232−1242

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The field of molecular information storage has recently expanded to include abiotic sequence-defined polymers. While robust methods have been developed, there is a current bottleneck in the throughput of this work as information density is increased. Herein, we introduce an automated workflow in which a commercial peptide synthesizer composed of a single XYZ liquid-handling robot was adapted to both synthesize and sequence sequence-defined oligourethanes. Our sequencing method was improved to cut down the number of samples required for each oligomer from 13 to one. Additionally, we introduce the use of desorption electrospray ionization mass spectrometry as our analysis method for sequencing, which allowed for simplified and increased speed of data acquisition. Finally, we created a Python script that is able to reconstruct the sequence information from the MS data in an automated fashion. We demonstrate this new workflow by encoding and decoding a quote from the late Maya Angelou: "When you learn, teach, when you get, give".



KEYWORDS: *polymer, sequence, urethane, information, automation*

INTRODUCTION

Abiotic sequence-defined macromolecules have attracted increasing attention over the past decade, in large part due to wide ranging applications in life and material sciences.[1,2] Sequence-defined oligomers (SDOs) and polymers (SDPs) have had a dramatic rise in the field of information storage, where they are seen as a complementary alternative to traditional silicon based storage devices.[3] Sequence-defined macromolecules have diverse chemical structure, with examples including, but not limited to, DNA,[4−6] peptides,[7,8] abiotic phosphates and phosphodiesters,[9−11] poly(alkoxyamine amide)s,[12−14] esters,[15−18] and urethanes.[19−25]

The use of sequence-defined macromolecules for information storage requires the ability to effectively "write" and "read"[26,27] the information, where the "writing" is often viewed as the synthesis, and the "reading" as sequencing. Synthetically, chemically diverse SDOs used for "writing" can be achieved in high yields by employing solid-phase synthesis (SPS). SPS methods have been optimized for numerous backbones, and often have coupling efficiencies of ≥99% per step.[28] At such coupling efficiencies, 10-mers can be synthesized in 90% or greater yields, and depending on the application, may be used sans purification. For SDOs, SPS can be applied to diverse monomers of a specific type (e.g., SPS of various amino acids) with protocols generally independent of length. In contrast, sequencing and sequence reconstruction protocols are often tailored to specific macromolecular structure.

High-throughput sequencing methods used to "read" abiotic polymers can rival those developed for biopolymers. Recently, nanopores have been shown to be capable of sequencing abiotic oligomers and polymers.[29−31] The use of nanopores for sequencing of abiotic macromolecules is still in its infancy, requiring monomers with significant differences and the addition of a large excess of noninformation encoding monomers in order to create distinct signals. Sequencing for the majority of synthetic sequence-controlled polymers is achieved using tandem mass spectrometry (MS) analysis.[8,13,14,17,22,23,32−37] Because the simplest form of reconstructing stored information postsequencing involves correlating one MS signal to one monomer or oligomer, a limitation of tandem MS analysis is the formation of adducts in a variety of charge-

Received: November 9, 2024
Revised: January 8, 2025
Accepted: January 10, 2025
Published: February 19, 2025

states. For example, in 2017 Lutz and co-workers synthesized an 8 byte containing polymer made of 64 information encoding monomers (base-2). The information encoding polymer initially ionized as adducts with multiple charge-states, and subsequent collision induced dissociation of the $[M - 24H]^{24-}$ precursor ion yielding variably charged smaller fragments that were further selected and subjected to MS. In this example, fragmentation by tandem MS analysis added additional complexity to the mass spectra that impeded automated sequence reconstruction and required manual decoding of low abundance fragments.[33] As an advance, in 2021 Lutz reported the design of a new monomer linker to mitigate undesired radical side-reactions following tandem MS fragmentation that prevented automated decoding.[32] While improvements to monomer design can help simplify spectra, this work highlights the challenges that come with sequencing by tandem MS; fragmentation in multiple locations and undesirable side-reactions following fragmentation increase spectral complexity, hinder automated decoding of information, and impart functional group limitations that decrease information storage potential. In addition, as the number of monomers increases in order to improve information density by increasing the mathematical base-coding scheme, inevitably tandem MS spectra will become inordinately complex and difficult to deconvolute. As such, in order to compliment DNA as an information storage platform and increase storage capacities beyond the current record of 18 bytes,[36] abiotic SDPs should be sequenced in an automated fashion using methodologies that allow for the maximum number of inexpensive, information rich monomers while minimizing spectral complexity to enable straightforward reconstruction.

Our group has reported the synthesis and self-sequencing of sequence-defined oligourethanes (SDOUs), with applications in information storage and steganography.[19,20,38] This platform relies on a chain-end depolymerization sequencing methodology that utilizes a thermally induced intramolecular cyclization to iteratively remove the terminal monomer. This approach allows each truncated OU to be characterized by liquid chromatography MS (LC/MS), forgoing tandem MS protocols and greatly simplifying sequence deconvolution.[38] The parent sequence can be reconstructed by identifying a delta m/z between two OUs that differ by a single monomer, where the monomer lost from sequencing at the O-terminus can be deciphered based on the m/z difference of the two strands. Most recently, we used this technique to encode a proverb in Mandarin using a base-26 framework.[39]

We previously described Mol.E-coder and Mol.E-decoder software,[20] that in short, are capable of encoding information into SDOUs based on the number of information encoding monomers used (e.g., hexadecimal encoding of bit strings encoded in base-16 with 16 unique monomers, 4 bits/monomer). The encoded hexadecimal string defines the sequence to be synthesized. Sequencing of each SDOU creates all possible chain-end depolymerized OUs, and mass differences between these sequenced strands can be fed into the decoder to retrieve the original encoded information.

The deconvolution method is simple enough that we were able to mix eight different OUs into a single sequencing vessel, and reconstruct each individual sequence based on the incorporation of unique isotope tags.[19] However, the sequencing procedure required each time point to be analyzed using an 18 min LC/MS run, which will be impractical as we move toward encoding larger amounts of data. Thus, methods that allow for expedited and higher throughput analysis are desired.

Compared to traditional MS approaches, the emergence of ambient ionization MS techniques has revolutionized molecular analysis by allowing for direct MS analysis of samples in open air conditions without the need for prior separation steps, shortening analysis times while still yielding rich molecular information.[40,41] One of the most well-established ambient ionization techniques is desorption electrospray ionization (DESI), which employs a spray of charged solvent microdroplets directed at a planar sample to desorb small molecules directly from the surface and propel the secondary microdroplets into the mass spectrometer for analysis.[42] Using DESI with a 2D moving stage allows the solvent sprayer to raster across the surface where specific areas of tissues[43] or samples can be analyzed with high spatial fidelity and different scanning speeds.[44] A study has reported the use of DESI for tandem MS analysis of SDPs.[45] In the case of screening alkenylation and azo-click reactions by the Cooks group, speeds as fast as 1 s per reaction were achieved.[46] Additionally, varying the solvent spray composition can promote the detection of specific molecules of interest, such as sequence-defined macromolecules, creating a more easily interpretable spectrum.

Here we describe a new method that addresses the main barriers in the use of OUs for high-throughput information storage by automating as well as increasing the speed of synthesis, sequencing, and sequence reconstruction. We describe the development of separate automated synthesis and sequencing platforms by adapting a commercially available synthesizer to perform the two different chemistries. We also increase the throughput of sequencing by combining all time points into one vessel, reducing the number of data points from 13 per OU down to only one per OU, and we introduce a new end-cap which can be added during automated synthesis removing the need for postsynthetic modification steps. Further, we describe the development and implementation of a desorption electrospray ionization mass spectrometry (DESI-MS)-based data acquisition method with a new automated sequence reconstruction program. Finally, we show a proof-of-concept trial of the entire process by encoding and decoding a short quote by Maya Angelou, "*When you learn, teach, when you get, give.*" Thus, this report describes the automated "writing" and "reading" of data in the form of SDOUs.

RESULTS AND DISCUSSION

Workflow Overview

Analogous to the storage of information in other SDPs, we posit that the synthesis, sequencing, and combination of the truncated SDOUs can be viewed as the "writing" of information, while the DESI-MS analysis and subsequent sequence reconstruction, can be viewed as the "reading" of said information. With this approach, sequenced SDOUs can be stored long-term and aliquots taken out when needed to allow for multiple "readings" of the information.

Central to our goal of automating the "writing" process was creating a workflow that utilized a robotic liquid handler for synthesis. To streamline this, previously reported procedures for SPS of urethanes[19,20,38] were adapted and optimized using a commercially available peptide synthesizer composed of an XYZ robot with a liquid handler. Optimization of synthesis protocols resulted in the fully automated synthesis of SDOUs that required no postsynthetic modifications and could be

https://doi.org/10.1021/jacsau.4c01070
JACS Au 2025, 5, 1232−1242

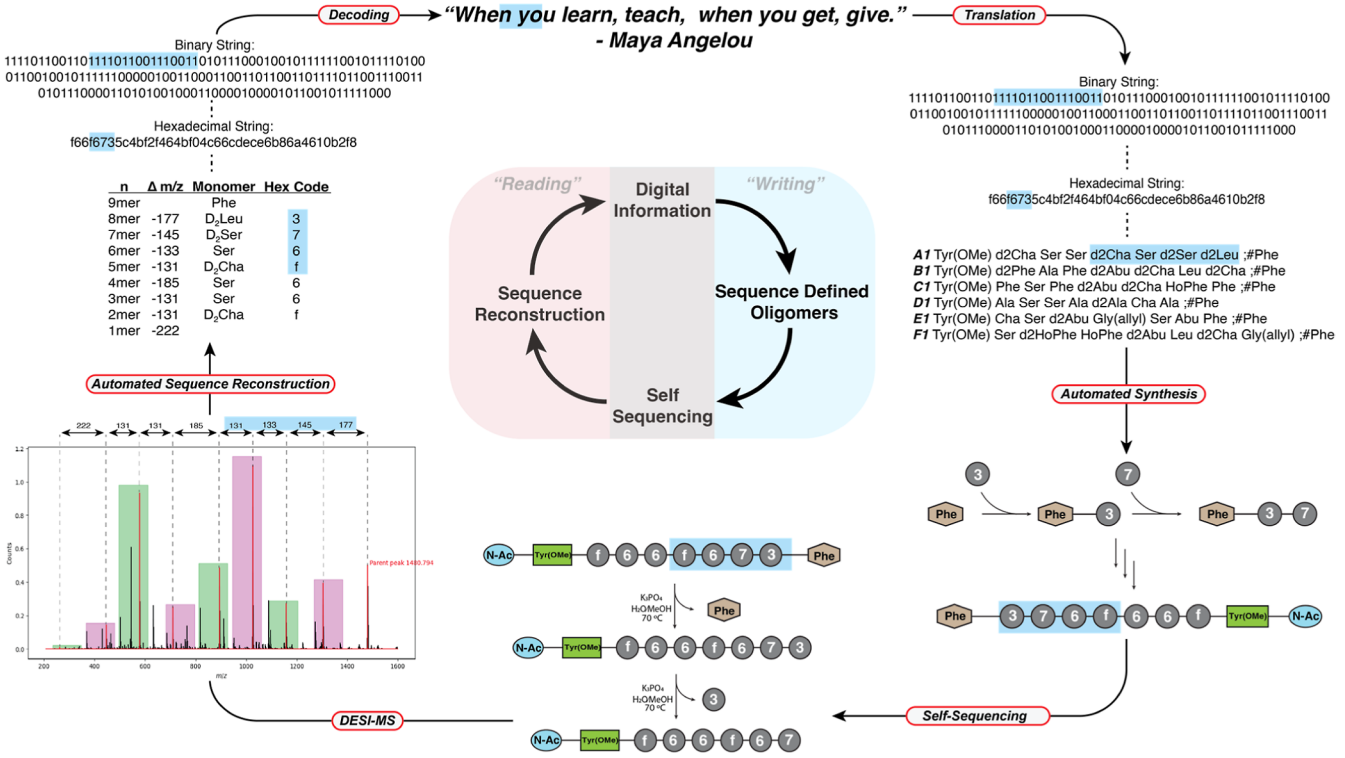


Figure 1. Workflow overview for the "reading" and "writing" of information. For the "writing" of information: digital information to be stored is shown in English at the top, and is followed by translation of the quote into binary. The binary string is then converted to hexadecimal (base-16) and each individual hex code is assigned to a unique monomer. Because the hexadecimal string for this quote is 42 characters long and we aim to have seven information encoding monomers per oligomer, we ended up with six SDOUs to synthesize (A1-F1). The blue highlight in the top right section of the figure tracks an example of the pieces of the binary string, hexadecimal string, and oligomer string that correspond to each other. Once the desired SDOUs are planned, the information is entered into the automated synthesizer and all six SDOUs are synthesized simultaneously. After automated synthesis the oligomers are sequenced via addition of base in the presence of heat. The combined aliquots of sequenced oligomer is where the "writing" process ends, as they contain all of the information needed in a "ready-to-read" format. For the "reading" of information: the six sequenced oligomer samples are spotted on a surface and subjected to DESI-MS analysis. The data from DESI-MS is then input into our in-house Python reconstruction program which automatically identifies the differences in mass between all sequenced oligomers. This process is described in detail in the "DESI-MS" and "Sequence Reconstruction" sections of the paper. Finally, we use Mol.E-decoder to correlate the mass differences to the correct monomers and the monomers to their unique hex codes. The rebuilt hexadecimal string is converted back into binary, and the binary back into English, thereby retrieving the original stored information.

sequenced sans purification, in part due to high-yielding individual coupling steps (≥98%). For example, although a 10-mer would result in ∼82% purity, the remaining 18% that is impurities would be split between each of the "incomplete" or truncated sequences. Each individual impurity is thereby negligible, contributing only ∼2% of the total impurities if distributed equally. Additionally, the need for purification was negated by incorporation of truncated caps, as is discussed later. We were then able to develop an automated sequencing protocol that employed the same robotic liquid handler, simply by manipulating the programming of the software. The "reading" process that followed sequencing involved optimizing DESI-MS experimental parameters to enhance the signal of the desired species (parent OU and all chain-end depolymerized OU strands).

Automation Workflow

We chose to combine and test our independently optimized synthesis and reconstruction programs to produce a continuous workflow for the automated encoding and decoding of information (Figure 1). To test this workflow, we selected a quote by the writer Maya Angelou: *"When you learn, teach, when you get, give."* To write this quote in base-16, we begin by converting the information from English into binary using the Huffman coding algorithm, and then convert the binary string into hexadecimal via an ASCII table. Each hex symbol, e.g. character of the hexadecimal string, corresponds to a unique monomer and from this monomer pool six 9-mers were synthesized and sequenced (Figure 2d). Six spots (one per sequenced oligomer), each 4 mm in diameter, were deposited onto a glass slide. The slide was then directly analyzed by DESI-MS at a stage rate of 1500 μm/s, allowing completion of the entire analysis in under 30 s. Note that Amalian et al. recently established the use of DESI in MS2 mode for reading molecularly encoded SDPs.[45] In their work, eight spots (each 3 mm in diameter) of five oligomers were needed to encode the four-character word "styx" in binary, which took approximately 5.5 min for the analysis to be completed using a stage rate of 500 μm/s, which is 5 min longer than our proposed method using MS1 only. Our acquired MS data were then fed into our Python reconstruction program. For each of the six 9-mers, the delta m/z values between truncated species were accurately matched to the associated monomer at that position. The data were processed with Mol.E-decoder to reconstruct the original quote.

https://doi.org/10.1021/jacsau.4c01070
JACS Au 2025, 5, 1232−1242

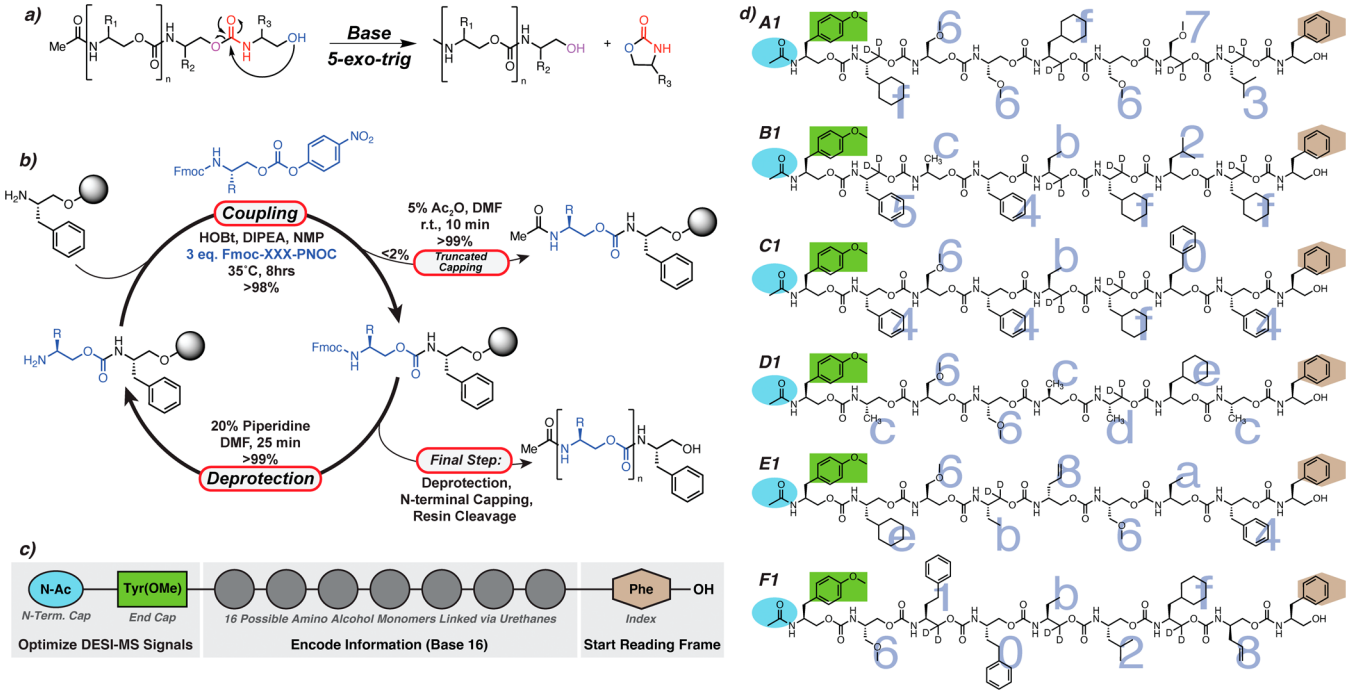


Figure 2. (a) Self-sequencing of β-amino alcohol-derived OUs in base proceeds O → N via a 5-*exo*-trig mechanism, successively removing the O-terminal monomer. (b) SPS of SDOUs. (c) General structure of SDOUs used in this report. (d) Specific structure of the six SDOUs required to encode the Maya Angelou quote, with hex codes shown in gray overlaying the corresponding monomers.

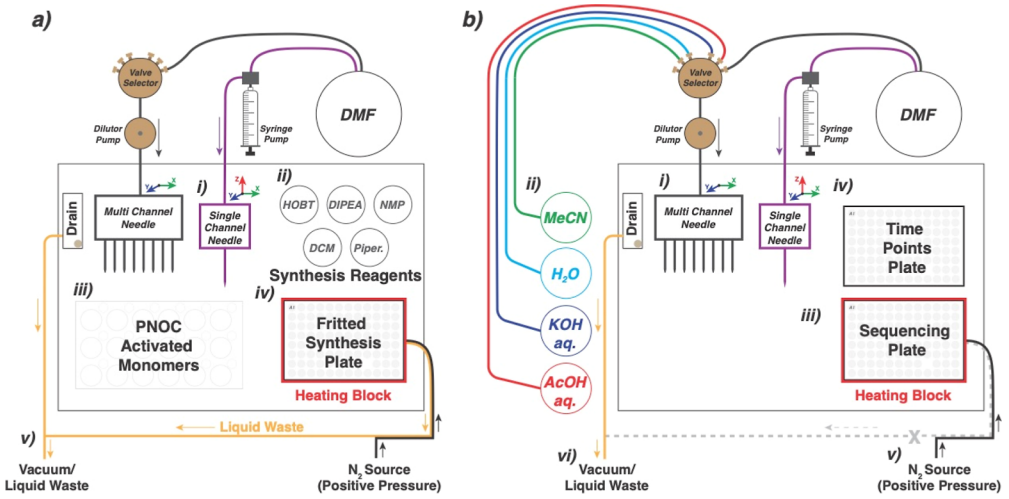


Figure 3. (a) Schematic of robotic liquid handler used for automated synthesis. (ai) Single channel needle that can move in the *XYZ* directions and functions to aspirate and dispense synthesis reagents and monomers into the fritted synthesis plate. (aii) Zone where synthesis reagents are placed. (aiii) Zone where PNOC activated monomers are placed along with tubes for mixing all reagents before the final dispense into the synthesis plate. (aiv) Fritted synthesis 96-well plate underneath which sits a heating block. Resin for all oligomers is deposited here in separate wells before synthesis begins. 96-well plate can be switched out for a 48-slot syringe column module that has a shaker underneath instead of the heating block. (av) Vacuum line, connected to the drain for rinsing the single channel needle in between uses and also connected to the fritted synthesis module for draining reagents postcoupling and washing steps. (b) Schematic of robotic liquid handler used for sequencing of SDOUs. (bi) Multichannel needle that can move in the *XY* directions and can only dispense reagents connected via the automated valve selector. For sequencing it is responsible for dispensing all reagents: solvent, base, and acid quench solution. (bii) Sequencing reagents prepped in bulk solutions and placed on the outer portion of the MP instrument, connected to the multichannel needle through the valve selector at the top of the instrument. (biii) Fritted sequencing 96-well plate fitted with a heating block underneath for sequencing, where the SDOUs are placed in separate wells after cleavage from the resin. (biv) Time points plate (not fritted), where aliquots taken from the sequencing plate by the single channel needle throughout the experiment are deposited. All aliquots taken from a particular well (e.g., A1) in the sequencing plate are deposited in the corresponding well (A1) of the time points plate, combining the data from the sequencing experiments and allowing for only one subsequent analysis per OU. (bv) N_2 manifold adapted to this instrument as a source of positive pressure to keep reagents from draining through the fritted apparatus' during synthesis and sequencing. (bvi) Vacuum that is still connected to the drain for rinsing of the single channel needle, but is disconnected from the sequencing plates during sequencing to avoid any loss of reagent. The materials in the sequencing plate need to remain in solution at all times.

https://doi.org/10.1021/jacsau.4c01070
JACS Au 2025, 5, 1232−1242

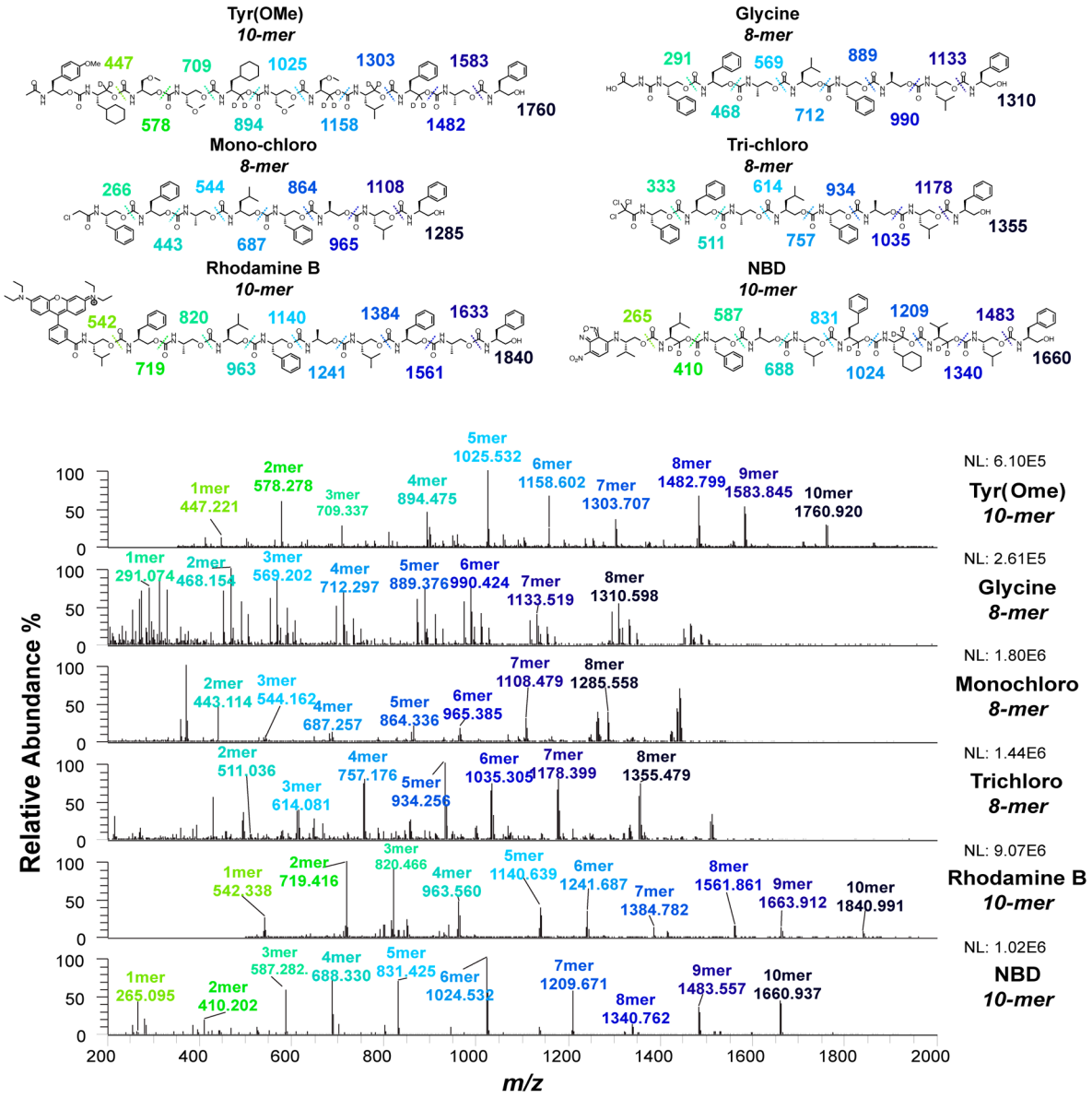


Figure 4. Comparison of DESI-MS data (bottom) obtained from test OUs appended with the six different end-caps (top). The first five end-caps were analyzed in positive mode with MeCN as the solvent. NBD was analyzed in negative mode with 98:2 MeOH/H_2O and 0.2% ammonium acetate as the solvent. Ultimately Tyr(OMe) was selected as the end-cap moving forward due to ease of synthesis and bias toward $[M + K]^+$ adducts.

Oligourethane Sequence Design

In order to optimize the automation workflow, automated synthesis of SDOUs needed to (1) achieve near quantitative coupling steps, (2) prevent deletion products from continuing in synthesis, (3) bias sequenced products to form one type of adduct (e.g., $[M + K]^+$) in DESI-MS, and (4) suppress ion formation of undesired byproducts.

Several changes to monomers were made compared to our previous reports to improve synthesis and sequence reconstruction. Monomers with protected side chains were omitted in order to simplify and standardize cleavage from resin, and monomers were chosen in order to keep the largest monomer mass less than twice the mass of the smallest monomer. While a subtle point, this change improved our ability to decipher the last information encoding monomer, as we will discuss later.

Automated Synthesis

SDOUs were synthesized by adapting an automated parallel peptide synthesizer (CEM MultiPep 2 Parallel Peptide Synthesizer, "MP") for the coupling of PNOC-activated amino alcohol monomers to an amino alcohol loaded resin (Figure 2b). A schematic showing the setup of the MP for synthesis is shown in Figure 3. This modification required programming designated zones within the software to accommodate the OU coupling reagents as well as calibrating the XYZ coordinates of the liquid-handling arm to be able to access each reagent or vessel. This robotic arm has attached two sets of needles, a single channel needle that can move in the *x*, *y*, and *z* directions and functions to aspirate and dispense all reagents, and a multichannel needle that can move in the *x* and *y* direction and dispenses solvent only. Our original method used fritted syringes as reaction vessels to allow for easy drainage with a vacuum pump between coupling steps.

https://doi.org/10.1021/jacsau.4c01070
JACS Au 2025, 5, 1232−1242

However, a source of positive pressure was required to keep the reagents from draining out over the reaction wait times during synthesis, so a nitrogen manifold was adapted to the instrument. The synthesis block is fitted with a shaking apparatus underneath the fritted syringe rack, as well as multiple syringe module sizes that hold 12−48 syringes. Additionally, we wanted to adapt this system for even higher throughput, so a second method was developed using fritted 96-well plates and a heating block to replace the shaker. Our previously reported synthesis procedures were scaled to the correct parameters for the instrument and input into the instrument "Method" file while monomer information was stored in the "Derivatives" file.

Optimization of OU synthesis was aided by the development of a sequence deletion and MS ion adduct calculator. This calculator requires the input of monomers being used along with their corresponding m/z, the designated sequence being synthesized, possible functional groups placed on the N- and O-terminus, and any possible ion adducts with corresponding expected observed m/z. The program produces a list of combinations for both positive and negative ion modes that include all possibilities of monomer deletions with different N- and O-terminus functionalities and different adducts. This calculator helped to deconvolute mass spectra during the design process to determine if deletions were occurring and the identity of the deleted monomer. Further details for both the automated synthesis and the MS adduct calculator can be found in the Supporting Information.

OU End-Caps

An "end-cap" is a single monomer added to the N-terminus of the target OU, thereby being present on the parent OU and all subsequent products formed during the sequencing reaction. End-caps have two major functions: (1) increase the relative signal intensity of the parent and sequenced OUs from background ions and undesired byproducts and (2) favor the formation of a single adduct type in DESI-MS (e.g., $[M + K]^+$). By biasing sequencing products to form a single adduct, the sequence reconstruction process is greatly simplified as the m/z difference between all observed ions simply matches the mass of the monomer lost. Six different end-caps were evaluated, including 4-fluoro-7-nitrobenzofurazan (NBD), chloroacetyl (Mono-Cl), trichloroacetyl (Tri-Cl), Glycine (Gly), RhodamineB (RhoB), and O-methyl Tyrosine [Tyr(OMe)] (Figure 4). NBD was the end-cap of choice in our previous work mainly due to its strong UV absorbance. With the switch from LC−MS to DESI-MS, a UV absorbance trace was no longer necessary. Additionally, NBD-F is an expensive reagent ($240/100 mg), requires an extra synthetic step postautomation, and hydrolyzes off over time during sequencing due to the aqueous reaction conditions leading to successively lower abundance of traceable product during the latter part of the experiment. Next, Mono-Cl and Tri-Cl were tested aiming to exploit their distinct isotopic patterns and easily select for the sequenced species of interest in the midst of spectral noise. Unfortunately, with both of these examples we observed the lower m/z species (e.g., 2-mer and 1-mer) had either low measured relative abundance or were not detected. Similar abundance issues were found with the Gly end-cap, as was the appearance of multiple adducts ($[M + H]^+$, $[M + Na]^+$, $[M + K^+]$). The fifth end-cap, RhoB, yielded promising results as it had high relative abundance and the heavier m/z of the end-cap itself shifted the entire mass spectrum toward a higher m/z range bringing all of the OU signals away from the m/z 100−300 range which often contains the most noise and background ions. However, under basic sequencing conditions RhoB undergoes a spirocyclization to form a lactam resulting in a loss of color due to a loss of conjugation. This process is easily reversible with the addition of acid or metal cations, but would require extra steps for ease of use during analysis. Additionally, RhoB requires two steps to be appended to the oligomers postautomated synthesis.

In the end, Tyr(OMe) was chosen as the ideal end-cap as (1) Fmoc-Tyr(OMe)-PNOC could be coupled to the growing strand using identical conditions as the information encoding monomers, eliminating the need for any synthetic modifications after removal from the MultiPep, (2) the monomer could be synthesized and recrystallized on a multigram scale, and (3) it was observed to exclusively form singly $[M + K]^+$ adducts using DESI-MS. The exclusive formation of this adduct in acetonitrile (MeCN) is attributed to the strong intermolecular cation-pi interaction,[47,48] and the presence of potassium cations (e.g., $K_3PO_4{\cdot}H_2O$) used during chain-end depolymerization sequencing.

OU Truncated Caps

"Truncated caps" were employed to block unreacted amines on the growing OU chain following an incomplete coupling step, as well as cap the N-terminus following completion of synthesis. Quantitative capping using a 5% acetic anhydride solution prevented deletion products from proceeding in synthesis, simplifying data analysis and aiding sequence reconstruction. As described above, this method largely contributed to our ability to skip any purification steps postsynthesis. Further, the acetylation of free amines prevented the formation of cationic ammonium salts, which were observed to have large ionization intensities when present.

Sequencing

Our group has previously reported methods for sequencing SDOUs, which include adding aqueous base and heating the samples while taking consecutive time points. Each of these time points was aliquoted in its own vial for later MS analysis. For the Maya Angelou quote, the six oligomers were sequenced using the original procedure, however we realized that the multiple time points taken could be gathered in a single well, further simplifying data analysis. Later, taking into account the rest of the automation workflow and our newly adaptable robotic synthesizer, we envisioned using this instrument to automate the sequencing program (Figure 3b). While the XYZ robot and liquid handler used for the project was commercially produced for the purpose of synthesis, we recognized the generality of the system and were able to apply it to another parallel experimental platform. Ensuring the MP could accommodate the sequencing of SDOUs required many adjustments to the existing procedure. Some of which included adapting the temperature and solvent system to allow heating while open to air, introducing acid after time point sampling to halt sequencing, as well as altering the concentration of base and time between sampling to allow for uniform data collection of up to 96 wells.

In short, sequencing in the MP involved the automated addition of sequencing reagents (e.g., base) to a 96-well plate that contained the SDOUs after cleavage from the resin. Mild heat (38 °C) was applied to the fritted sequencing plate to allow the reaction to proceed without agitation, and a constant flow of N_2 was present as a source of positive pressure to keep the reagents from leaking through the frits. Approximately

https://doi.org/10.1021/jacsau.4c01070
JACS Au 2025, 5, 1232−1242

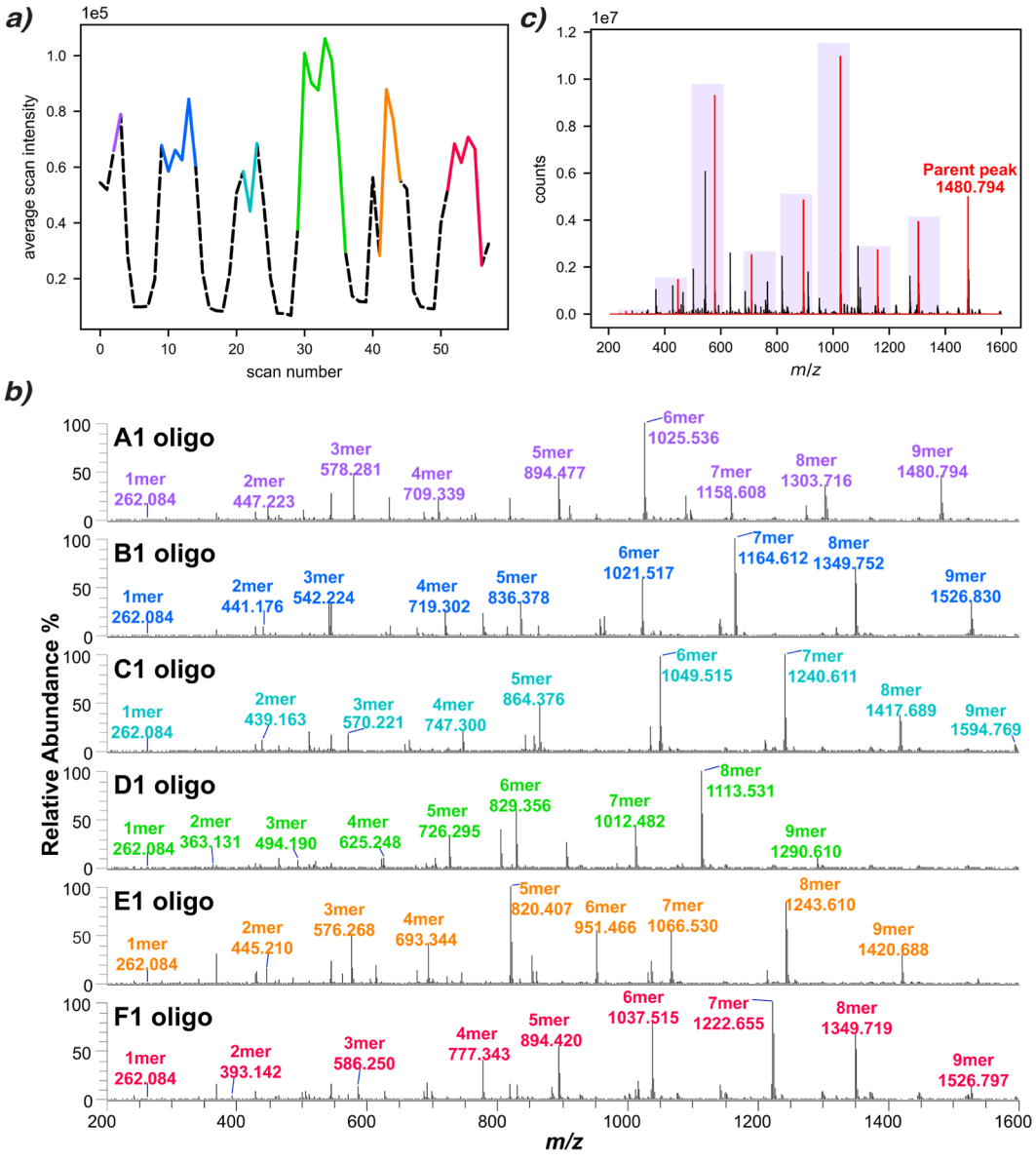


Figure 5. (a) Overall scan intensity for the DESI-MS raster of a plate with 6 samples spotted on the plate. The raw scan intensities are colored in black with the selected scans for analysis colored differently to correspond to the sample being analyzed. (b) Stacked spectra of six Tyr(OMe) capped OUs colored to correspond with the scans in part (a). (c) DESI-MS sequencing data for 9-mer A1. The selected signals associated with an $[M + K]^+$ OU are highlighted in red. The m/z windows in which monomer losses are identified are emphasized in purple for consecutive $n - 1$ oligomers. The DesiSequencer program takes into account both peak abundance and proximity to an idealized m/z of a truncation, meaning the largest peaks within a window are not necessarily selected for if a smaller peak better matches one of the $n - 1$ oligomers.

every 60 min, the single channel needle aspirates an aliquot from a sequencing well and dispenses into the corresponding well of a "time points" 96-well plate. Each OU is designated one well in the "time points" plate, thus, at the end of sequencing, multiple time points are still combined in a single well. For a full comparison of sequencing procedures and further details on the automation of the *5-exo-trig* sequencing see the Oligourethane Sequencing section of the Supporting Information.

Desorption Electrospray Ionization Mass Spectrometry

Initial work was done to develop and optimize a DESI-MS method for screening end-caps of OUs, with a goal to find the combination of parameters that resulted in all of the sequenced species being present and in high abundance. We evaluated the ionization of OUs capped with five of the six different groups described previously using DESI-MS in the positive ion mode with different solvent systems. NBD-capped OUs were evaluated in the negative ion mode. Initial results indicated that all sequenced species of test OUs were easily observed using MeCN as the spray solvent, with the exception of NBD which required a 98:2 MeOH solution with 0.2% ammonium acetate (Figure 4).

Additionally, the velocity of the 2D moving stage was optimized to 1500 μm/s to accelerate the analysis without compromising data quality. By tuning these DESI-MS parameters, analysis of one OU with all its iterative truncations was achieved in approximately 3 s using the optimized stage rate. Further details on DESI-MS optimization are found in the Supporting Information.

https://doi.org/10.1021/jacsau.4c01070
JACS Au 2025, 5, 1232−1242

The raw data from the DESI-MS analysis is then converted to a plaintext file containing m/z and count values for each scan. The peak centers are selected based on the Savitzky–Golay filtered raster data (Figure 5a). From these peak centers, the individual scans behind and ahead of the peak center are combined to generate an overall mass spectrum for that sample (Figure 5b).

Sequence Reconstruction

With an efficient sequencing method in hand, we then developed a general protocol for extracting sequence information from the MS data. Using a Python script, we first automatically identify the mass of the parent oligomer. The program then searches the range of m/z values which would encompass the result of losing the heaviest monomer to the lightest monomer for the $n - 1$ oligomer. The signal of the $n - 1$ oligomer is then selected based on a scoring function that focuses primarily on the proximity of a potential $n - 1$ oligomer signal to an ideal $n - 1$ oligomer mass. Because there are often signals that are close to the $n - 1$ oligomer mass, the selection of the $n - 1$ oligomer signal is biased toward high intensity signals to avoid selecting noise. The difference between the parent signal and the identified signal of the sequenced product (i.e., the $n - 1$ oligomer) is then calculated and compared to a dictionary of possible monomers to assign the mass loss to a monomer identity. To accomplish this, the differences between the mass lost and all possible monomer masses is calculated and the lowest difference is assigned as the monomer. The process is then repeated using the $n - 1$ oligomer as the new parent oligomer to identify subsequent mass losses ($n - 2$, $n - 3$, etc.). This occurs until the end-cap monomer mass is identified (in the present case Tyr(OMe)) and the resulting sequence data is obtained (Figure 5b). This strategy is robust and was found to extract the sequence information without error for the 9-mers. For further details, see the Automated Sequence Reconstruction using DesiSequencer section of the Supporting Information.

The final step required to translate the DESI-MS data back into the original quote is to convert the oligomer sequences into their Huffman-coded representations using Mol.E-decoder. This process is achieved by converting the monomer identities into hexadecimal and then into Huffman-encoded half-byte representations (e.g., binary). The binary can be further translated back into English using the Huffman compression results, allowing us to read the original quote without error.[20]

General Considerations

As described above, the automation of: (1) the synthesis, (2) the 5-*exo*-trig degradation, (3) the combination of degradation time points, (4) the DESI spotting and analysis, and finally (5) the sequence reconstruction took considerable time, effort, and computer program development. These aspects of the work would represent a bottleneck for anyone wanting to immediately implement this methodology. In fact, our original one-by-one oligourethane synthesis followed by LC–MS was simpler, albeit it lacked the ability for increasing throughput. But importantly, given that organic chemistry endeavors are progressively turning to automation,[49−53] the obstacles for high throughput experiments will continue to diminish, and this work is a step in that direction.

Additionally, we find that automation blurs the line between writing and reading. To understand this, let us first remove the human from the process, because we clearly decide what is to be written and we do the final reading with our eyes. Thus, with a focus solely on the steps in this paper, we defined above the conversion of English to binary, then to hexadecimal, the oligourethane synthesis, and the *5-exo-trig* degradation including the combination of truncated oligomers into single samples, as the steps involved in writing. By contrast the steps of spotting the DESI plate, followed by the DESI-MS analysis and running the Mol.E-decoder software to convert monomer m/z to hexadecimal, then to binary, and back into English all count as reading. But this was an arbitrary separation. In our opinion, when considering the use of SDPs for information storage, the point at which the synthetic material is placed into storage ends the writing process. In our case, this could be the synthetic oligomers themselves, or their combined truncated oligomers, or the spotted DESI plates. Reading would commence from whatever step ends the writing. The emphasis is to have a tangible physical object in some form to end the writing as an analogy to a book or a computer screen.

When humans read, we typically recognize symbols as representing words, either individually or in combination, in a process of decoding. But at that point we have not parsed the words, linking them to mental representations that have meaning to us. In a sense, our use of chain-end degradation and the combination of the truncated strands into a single sample is analogous to conversion of symbols on a page to words, i.e., decoding. The DESI-MS data is akin to having a written document, while the Mol.E-decoder finally generates meaningful information, i.e., the parsing. The combination of organic chemistry with information science, as well as linguistics, is ripe for investigation of how these fields blend, complement, and enhance one another.

CONCLUSION

Herein, we report an automated workflow for the storage and retrieval of information with sequence-defined oligourethanes using DESI-MS. A commercially available peptide synthesizer was adapted to accommodate oligourethane synthesis, therefore speeding up the process and allowing up to 96 oligomers to be synthesized in parallel in an automated fashion. Independently, the same instrument was used to then automate sequencing of oligomers. We chose DESI-MS as our data acquisition method, and the end-caps of our oligomers were tailored to complement and deconvolute the data. A Python script was developed that allowed for automated data analysis of all six oligomers simultaneously, and was shown to correctly track and identify each monomer lost during sequencing. Our previously developed Mol.E-coder and Mol.E-decoder programs were used for encoding and decoding a proof-of-concept quote that we used to string the entire process together. The high-throughput automation of this workflow has implications for the future of molecular information storage, paving the way for a greater density of information to be stored as it cuts down on time and direct attention required from the researcher. Such advances also spark discussions of how chemistry, human cognition, and information/linguistics will merge and create new interdisciplinary ventures.

ASSOCIATED CONTENT

Data Availability Statement

The sequence deletion calculator, an interactive Jupyter notebook demonstrating automated sequence reconstruction,

https://doi.org/10.1021/jacsau.4c01070
JACS Au 2025, 5, 1232−1242

and algorithms for encoding and decoding information can be found on github at (1) https://github.com/PhysicalOrganic/SequenceDeletionCalculator/tree/main, (2) https://github.com/PhysicalOrganic/DesiSequencer, (3) https://github.com/PhysicalOrganic/Mol.E-coder.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacsau.4c01070.

Detailed experimental procedures, supplementary data, and spectral data for all compounds (PDF)

AUTHOR INFORMATION

Corresponding Authors

Livia S. Eberlin − *Department of Surgery, Baylor College of Medicine, Houston, Texas 77030, United States;* orcid.org/0000-0002-3885-3215; Email: livia.eberlin@bcm.edu

Eric V. Anslyn − *Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, United States;* orcid.org/0000-0002-5137-8797; Email: anslyn@autin.utexas.edu

Authors

Julia R. Shuluk − *Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, United States;* orcid.org/0009-0001-1851-884X

Christopher D. Wight − *Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, United States; Department of Chemistry, Texas State University, San Marcos, Texas 78666, United States;* orcid.org/0000-0003-3389-1762

James R. Howard − *Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, United States;* orcid.org/0000-0002-9184-6954

Mary E. King − *Department of Surgery, Baylor College of Medicine, Houston, Texas 77030, United States*

Sarah R. Moor − *Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, United States;* orcid.org/0000-0001-5579-407X

Rachel J. DeHoog − *Department of Surgery, Baylor College of Medicine, Houston, Texas 77030, United States*

Samuel D. Dahlhauser − *Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, United States;* orcid.org/0000-0002-4034-5172

Complete contact information is available at: https://pubs.acs.org/10.1021/jacsau.4c01070

Author Contributions

[‖]J.R.S and C.D.W contributed equally. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare the following competing financial interest(s): J.R.S, C.D.W., J.R.H., M.E.K., S.R.M., S.D.D., L.S.E., and E.V.A. are co-inventors on the following patent: COMPOSITIONS AND METHODS FOR ENCRYPTING, STORING, AND DECRYPTING INFORMATION IN OLIGOMERS. PCT/US2024/026494. 10046-505WO1. 7846 ANS.

ACKNOWLEDGMENTS

We gratefully acknowledge financial support for this work from the Keck Foundation (UTA20-000926), the NSF (2203354), and the Welch Reagents Chair to E.V.A. (F-0046). We again thank the Welch Foundation for support to L.S.E. (Q-1895-20220331). We would like to acknowledge the UT Mass Spectrometry Facility for their instrumental help and the UT NMR facilities for Bruker AVANCE III 500: NIH Grant Number 1 S10 OD021508-01. We also acknowledge our linguistics collaborators Drs. Danny Law and Todd Krause for helpful discussions.

ABBREVIATIONS

SDP, sequence-defined polymer; SDO, sequence-defined oligomer; SPS, solid-phase synthesis; MS, mass-spectrometry; LC/MS, liquid-chromatography mass spectrometry; DESI, desorption electrospray ionization; NBD, 4-fluoro-7-nitrobenzofurazan; Mono-Cl, monochloroacetyl; Tri-Cl, trichloroacetyl; Gly, glycine; RhoB, RhodamineB; Tyr(OMe), *O*-methyl tyrosine; MeCN, acetonitrile; OU, oligourethane; SDOU, sequence-defined oligourethane; MP, MultiPep 2

REFERENCES

(1) Aksakal, R.; Mertens, C.; Soete, M.; Badi, N.; Du Prez, F. Applications of Discrete Synthetic Macromolecules in Life and Materials Science: Recent and Future Trends. *Advanced Science* **2021**, *8* (6), 2004038.

(2) Deng, Z.; Gillies, E. R. Emerging Trends in the Chemistry of End-to-End Depolymerization. *JACS Au* **2023**, *3* (9), 2436−2450.

(3) Zhirnov, V.; Zadegan, R. M.; Sandhu, G. S.; Church, G. M.; Hughes, W. L. Nucleic Acid Memory. *Nat. Mater.* **2016**, *15* (4), 366−370.

(4) Church, G. M.; Gao, Y.; Kosuri, S. Next-Generation Digital Information Storage in DNA. *Science* **2012**, *337* (6102), 1628.

(5) Pan, C.; Tabatabaei, S. K.; Tabatabaei Yazdi, S. M. H.; Hernandez, A. G.; Schroeder, C. M.; Milenkovic, O. Rewritable Two-Dimensional DNA-Based Data Storage with Machine Learning Reconstruction. *Nat. Commun.* **2022**, *13*, 2984.

(6) Zhang, C.; Wu, R.; Sun, F.; Lin, Y.; Liang, Y.; Teng, J.; Liu, N.; Ouyang, Q.; Qian, L.; Yan, H. Parallel Molecular Data Storage by Printing Epigenetic Bits on DNA. *Nature* **2024**, *634* (8035), 824−832.

(7) Cafferty, B. J.; Ten, A. S.; Fink, M. J.; Morey, S.; Preston, D. J.; Mrksich, M.; Whitesides, G. M. Storage of Information Using Small Organic Molecules. *ACS Cent. Sci.* **2019**, *5* (5), 911−916.

(8) Rössler, S. L.; Grob, N. M.; Buchwald, S. L.; Pentelute, B. L. Abiotic Peptides as Carriers of Information for the Encoding of Small-Molecule Library Synthesis. *Science* **2023**, *379* (6635), 939−945.

(9) Al Ouahabi, A.; Charles, L.; Lutz, J.-F. Synthesis of Non-Natural Sequence-Encoded Polymers Using Phosphoramidite Chemistry. *J. Am. Chem. Soc.* **2015**, *137* (16), 5629−5635.

(10) Al Ouahabi, A.; Kotera, M.; Charles, L.; Lutz, J.-F. Synthesis of Monodisperse Sequence-Coded Polymers with Chain Lengths above DP100. *ACS Macro Lett.* **2015**, *4* (10), 1077−1080.

(11) König, N. F.; Al Ouahabi, A.; Poyer, S.; Charles, L.; Lutz, J.-F. A Simple Post-Polymerization Modification Method for Controlling Side-Chain Information in Digital Polymers. *Angew. Chem., Int. Ed. Engl.* **2017**, *56* (25), 7297−7301.

(12) Charles, L.; Laure, C.; Lutz, J.-F.; Roy, R. K. Tandem Mass Spectrometry Sequencing in the Negative Ion Mode to Read Binary Information Encoded in Sequence-Defined Poly(Alkoxyamine Amide)s. *Rapid Commun. Mass Spectrom.* **2016**, *30* (1), 22−28.

(13) Roy, R. K.; Meszynska, A.; Laure, C.; Charles, L.; Verchin, C.; Lutz, J.-F. Design and Synthesis of Digitally Encoded Polymers That Can Be Decoded and Erased. *Nat. Commun.* **2015**, *6* (1), 7237.

(14) Charles, L.; Laure, C.; Lutz, J.-F.; Roy, R. K. MS/MS Sequencing of Digitally Encoded Poly(Alkoxyamine Amide)s. *Macromolecules* **2015**, *48* (13), 4319−4328.
(15) Lee, J. M.; Kwon, J.; Lee, S. J.; Jang, H.; Kim, D.; Song, J.; Kim, K. T. Semiautomated Synthesis of Sequence-Defined Polymers for Information Storage. *Sci. Adv.* **2022**, *8* (10), No. eabl8614.
(16) Lee, J. M.; Jang, H.; Lee, S. W.; Kim, K. T. Nondestructive Sequencing of Enantiopure Oligoesters by Nuclear Magnetic Resonance Spectroscopy. *JACS Au* **2022**, *2* (9), 2108−2118.
(17) Lee, J. M.; Koo, M. B.; Lee, S. W.; Lee, H.; Kwon, J.; Shim, Y. H.; Kim, S. Y.; Kim, K. T. High-Density Information Storage in an Absolutely Defined Aperiodic Sequence of Monodisperse Copolyester. *Nat. Commun.* **2020**, *11* (1), 56.
(18) Soete, M.; Du Prez, F. E. Sequencing of Uniform Multifunctional Oligoesters via Random Chain Cleavages. *Angew. Chem., Int. Ed.* **2022**, *61* (24), No. e202202819.
(19) Dahlhauser, S. D.; Wight, C. D.; Moor, S. R.; Scanga, R. A.; Ngo, P.; York, J. T.; Vera, M. S.; Blake, K. J.; Riddington, I. M.; Reuther, J. F.; Anslyn, E. V. Molecular Encryption and Steganography Using Mixtures of Simultaneously Sequenced, Sequence-Defined Oligourethanes. *ACS Cent. Sci.* **2022**, *8* (8), 1125−1133.
(20) Dahlhauser, S. D.; Moor, S. R.; Vera, M. S.; York, J. T.; Ngo, P.; Boley, A. J.; Coronado, J. N.; Simpson, Z. B.; Anslyn, E. V. Efficient Molecular Encoding in Multifunctional Self-Immolative Urethanes. *Cell Rep. Phys. Sci.* **2021**, *2* (4), 100393.
(21) Soete, M.; Van Hoorde, J.; Du Prez, F. Discrete, Self-Immolative N-Substituted Oligourethanes and Their Use as Molecular Tags. *Polym. Chem.* **2022**, *13* (28), 4178−4185.
(22) Gunay, U. S.; Petit, B. E.; Karamessini, D.; Al Ouahabi, A.; Amalian, J.-A.; Chendo, C.; Bouquey, M.; Gigmes, D.; Charles, L.; Lutz, J.-F. Chemoselective Synthesis of Uniform Sequence-Coded Polyurethanes and Their Use as Molecular Tags. *Chem* **2016**, *1* (1), 114−126.
(23) Martens, S.; Landuyt, A.; Espeel, P.; Devreese, B.; Dawyndt, P.; Du Prez, F. Multifunctional Sequence-Defined Macromolecules for Chemical Data Storage. *Nat. Commun.* **2018**, *9* (1), 4451.
(24) Mondal, T.; Charles, L.; Lutz, J.-F. Damage and Repair in Informational Poly(N-Substituted Urethane)s. *Angew. Chem., Int. Ed.* **2020**, *59* (46), 20390−20393.
(25) Zwillinger, M.; Fischer, L.; Sályi, G.; Szabó, S.; Csékei, M.; Huc, I.; Kotschy, A. Isotope Ratio Encoding of Sequence-Defined Oligomers. *J. Am. Chem. Soc.* **2022**, *144* (41), 19078−19088.
(26) Soete, M.; Mertens, C.; Badi, N.; Du Prez, F. E. Reading Information Stored in Synthetic Macromolecules. *J. Am. Chem. Soc.* **2022**, *144* (49), 22378−22390.
(27) Holloway, J. O.; Van Lijsebetten, F.; Badi, N.; Houck, H. A.; Du Prez, F. E. From Sequence-Defined Macromolecules to Macromolecular Pin Codes. *Advanced Science* **2020**, *7* (8), 1903698.
(28) Cho, C. Y.; Moran, E. J.; Cherry, S. R.; Stephans, J. C.; Fodor, S. P. A.; Adams, C. L.; Sundaram, A.; Jacobs, J. W.; Schultz, P. G. An Unnatural Biopolymer. *Science* **1993**, *261* (5126), 1303−1305.
(29) Cao, C.; Krapp, L. F.; Al Ouahabi, A.; König, N. F.; Cirauqui, N.; Radenovic, A.; Lutz, J.-F.; Peraro, M. D. Aerolysin Nanopores Decode Digital Information Stored in Tailored Macromolecular Analytes. *Sci. Adv.* **2020**, *6* (50), No. eabc2661.
(30) Tabatabaei, S. K.; Pham, B.; Pan, C.; Liu, J.; Chandak, S.; Shorkey, S. A.; Hernandez, A. G.; Aksimentiev, A.; Chen, M.; Schroeder, C. M.; Milenkovic, O. Expanding the Molecular Alphabet of DNA-Based Data Storage Systems with Neural Network Nanopore Readout Processing. *Nano Lett.* **2022**, *22* (5), 1905−1914.
(31) Yan, S.; Wang, L.; Wang, Y.; Cao, Z.; Zhang, S.; Du, X.; Fan, P.; Zhang, P.; Chen, H.-Y.; Huang, S. Non-Binary Encoded Nucleic Acid Barcodes Directly Readable by a Nanopore. *Angew. Chem., Int. Ed.* **2022**, *61* (20), No. e202116482.
(32) Launay, K.; Amalian, J.-A.; Laurent, E.; Oswald, L.; Al Ouahabi, A.; Burel, A.; Dufour, F.; Carapito, C.; Clément, J.-L.; Lutz, J.-F.; Charles, L.; Gigmes, D. Precise Alkoxyamine Design to Enable Automated Tandem Mass Spectrometry Sequencing of Digital Poly(Phosphodiester)s. *Angew. Chem., Int. Ed.* **2021**, *60* (2), 917−926.
(33) Al Ouahabi, A.; Amalian, J.-A.; Charles, L.; Lutz, J.-F. Mass Spectrometry Sequencing of Long Digital Polymers Facilitated by Programmed Inter-Byte Fragmentation. *Nat. Commun.* **2017**, *8* (1), 967.
(34) Cavallo, G.; Al Ouahabi, A.; Oswald, L.; Charles, L.; Lutz, J.-F. Orthogonal Synthesis of "Easy-to-Read" Information-Containing Polymers Using Phosphoramidite and Radical Coupling Steps. *J. Am. Chem. Soc.* **2016**, *138* (30), 9417−9420.
(35) Liu, B.; Shi, Q.; Hu, L.; Huang, Z.; Zhu, X.; Zhang, Z. Engineering Digital Polymer Based on Thiol−Maleimide Michael Coupling toward Effective Writing and Reading. *Polym. Chem.* **2020**, *11* (10), 1702−1707.
(36) Laurent, E.; Amalian, J.-A.; Parmentier, M.; Oswald, L.; Al Ouahabi, A.; Dufour, F.; Launay, K.; Clément, J.-L.; Gigmes, D.; Delsuc, M.-A.; Charles, L.; Lutz, J.-F. High-Capacity Digital Polymers: Storing Images in Single Molecules. *Macromolecules* **2020**, *53* (10), 4022−4029.
(37) Soete, M.; De Bruycker, K.; Du Prez, F. Rewritable Macromolecular Data Storage with Automated Read-Out. *Angew. Chem., Int. Ed.* **2022**, *61* (13), No. e202116718.
(38) Dahlhauser, S. D.; Escamilla, P. R.; VandeWalle, A. N.; York, J. T.; Rapagnani, R. M.; Shei, J. S.; Glass, S. A.; Coronado, J. N.; Moor, S. R.; Saunders, D. P.; Anslyn, E. V. Sequencing of Sequence-Defined Oligourethanes via Controlled Self-Immolation. *J. Am. Chem. Soc.* **2020**, *142* (6), 2744−2749.
(39) Zhang, L. B.; Krause, T.; Deol, H.; Pandey, B.; Xiao, Q.; Park, H. M.; Iverson, B. L.; Law, D.; Anslyn, E. V. Chemical and Linguistic Considerations for Encoding Chinese Characters: An Embodiment Using Chain-End Degradable Sequence-Defined Oligourethanes Created by Consecutive Solid Phase Click Chemistry. *Chem. Sci.* **2024**, *15* (14), 5284−5293.
(40) Takáts, Z.; Wiseman, J. M.; Gologan, B.; Cooks, R. G. Mass Spectrometry Sampling Under Ambient Conditions with Desorption Electrospray Ionization. *Science* **2004**, *306* (5695), 471−473.
(41) Cooks, R. G.; Ouyang, Z.; Takats, Z.; Wiseman, J. M. Ambient Mass Spectrometry. *Science* **2006**, *311* (5767), 1566−1570.
(42) Takáts, Z.; Wiseman, J. M.; Cooks, R. G. Ambient Mass Spectrometry Using Desorption Electrospray Ionization (DESI): Instrumentation, Mechanisms and Applications in Forensics, Chemistry, and Biology. *J. Mass Spectrom.* **2005**, *40* (10), 1261−1275.
(43) Sans, M.; Krieger, A.; Wygant, B. R.; Garza, K. Y.; Mullins, C. B.; Eberlin, L. S. Spatially Controlled Molecular Analysis of Biological Samples Using Nanodroplet Arrays and Direct Droplet Aspiration. *J. Am. Soc. Mass Spectrom.* **2020**, *31* (2), 418−428.
(44) Wleklinski, M.; Loren, B. P.; Ferreira, C. R.; Jaman, Z.; Avramova, L.; Sobreira, T. J. P.; Thompson, D. H.; Cooks, R. G. High Throughput Reaction Screening Using Desorption Electrospray Ionization Mass Spectrometry. *Chem. Sci.* **2018**, *9* (6), 1647−1653.
(45) Amalian, J.-A.; Mondal, T.; Konishcheva, E.; Cavallo, G.; Petit, B. E.; Lutz, J.-F.; Charles, L. Desorption Electrospray Ionization (DESI) of Digital Polymers: Direct Tandem Mass Spectrometry Decoding and Imaging from Materials Surfaces. *Adv. Mater. Technol.* **2021**, *6* (4), 2001088.
(46) Huang, K.-H.; Ghosh, J.; Xu, S.; Cooks, R. G. Late-Stage Functionalization and Characterization of Drugs by High-Throughput Desorption Electrospray Ionization Mass Spectrometry. *ChemPlusChem* **2022**, *87* (1), No. e202100449.
(47) Dougherty, D. A. The Cation−π Interaction. *Acc. Chem. Res.* **2013**, *46* (4), 885−893.
(48) Mecozzi, S.; West, A. P.; Dougherty, D. A. Cation-Pi Interactions in Aromatics of Biological and Medicinal Interest: Electrostatic Potential Surfaces as a Useful Qualitative Guide. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93* (20), 10566−10571.
(49) Leonov, A. I.; Hammer, A. J. S.; Lach, S.; Mehr, S. H. M.; Caramelli, D.; Angelone, D.; Khan, A.; O'Sullivan, S.; Craven, M.; Wilbraham, L.; Cronin, L. An Integrated Self-Optimizing Program-

https://doi.org/10.1021/jacsau.4c01070
JACS Au 2025, 5, 1232−1242

mable Chemical Synthesis and Reaction Engine. *Nat. Commun.* **2024**, *15* (1), 1240.
(50) Wang, W.; Angello, N. H.; Blair, D. J.; Tyrikos-Ergas, T.; Krueger, W. H.; Medine, K. N. S.; LaPorte, A. J.; Berger, J. M.; Burke, M. D. Rapid Automated Iterative Small-Molecule Synthesis. *Nat. Synth.* **2024**, *3* (8), 1031−1038.
(51) Bubliauskas, A.; Blair, D. J.; Powell-Davies, H.; Kitson, P. J.; Burke, M. D.; Cronin, L. Digitizing Chemical Synthesis in 3D Printed Reactionware. *Angew. Chem., Int. Ed.* **2022**, *61* (24), No. e202116108.
(52) Rauschen, R.; Guy, M.; Hein, J. E.; Cronin, L. Universal Chemical Programming Language for Robotic Synthesis Repeatability. *Nat. Synth.* **2024**, *3* (4), 488−496.
(53) El-khawaldeh, R.; Mandal, A.; Yoshikawa, N.; Zhang, W.; Corkery, R.; Prieto, P.; Aspuru-Guzik, A.; Darvish, K.; Hein, J. E. From Eyes to Cameras: Computer Vision for High-Throughput Liquid-Liquid Separation. *Device* **2024**, *2* (7), 100404.

https://doi.org/10.1021/jacsau.4c01070
JACS Au 2025, 5, 1232−1242