

Controlling the structure of sequence-defined poly (phosphodiester)s for optimal MS/MS reading of digital information

J.-A. Amalian, A. Al Ouahabi, G. Cavallo, N. König, S. Poyer, J.-F. Lutz, L. Charles

► **To cite this version:**

J.-A. Amalian, A. Al Ouahabi, G. Cavallo, N. König, S. Poyer, et al.. Controlling the structure of sequence-defined poly (phosphodiester)s for optimal MS/MS reading of digital information. *Journal of Mass Spectrometry*, Wiley-Blackwell, 2017, 52 (11), pp.788-798. 10.1002/jms.3947 . hal-01983183

HAL Id: hal-01983183

<https://hal.archives-ouvertes.fr/hal-01983183>

Submitted on 16 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

*This is the peer reviewed version of the following article:
Jean-Arthur Amalian, Abdelaziz Al Ouahabi, Gianni Cavallo, Niklas Felix
König, Salomé Poyer, Jean-François Lutz and Laurence Charles,*

J. Mass Spectrom., 2017, 52: 788–798,

which has been published in final form at

<https://onlinelibrary.wiley.com/doi/full/10.1002/jms.3947>

*This article may be used for non-commercial purposes in accordance with Wiley
Terms and Conditions for Self-Archiving."*

Controlling the structure of sequence-defined poly(phosphodiester)s for optimal MS/MS reading of digital information

Jean-Arthur Amalian,¹ Abdelaziz Al Ouahabi,² Gianni Cavallo,² Niklas Felix König,² Salomé Poyer,¹ Jean-François Lutz^{2*} and Laurence Charles^{1*}

¹ Aix Marseille Univ, CNRS, ICR, Institut de Chimie Radicalaire, Marseille – France

² Université de Strasbourg, CNRS, Institut Charles Sadron UPR22, 23 rue du Loess, 67034 Strasbourg Cedex 2, France

* To whom correspondence should be addressed. E-mail: laurence.charles@univ-amu.fr Phone: +33 491 28 8678. Fax: +33 491 28 2897. Email: jflutz@unistra.fr. Phone: +33 388 41 4016. Fax: +33 388 41 4099.

Running title

MS/MS sequencing of coded poly(phosphodiester)s

Key words

MS/MS sequencing; digital polymers; polyphosphates; molecular design; sequence-controlled polymers; collision-induced dissociation; information storage

Abstract

Digital polymers are monodisperse chains with a controlled sequence of co-monomers, defined as letters of an alphabet, and are used to store information at the molecular level. Reading such messages is hence a sequencing task that can be efficiently achieved by tandem mass spectrometry. To improve their readability, structure of sequence-controlled synthetic polymers can be optimized, based on considerations regarding their fragmentation behavior. This strategy is described here for poly(phosphodiester)s, that were synthesized as monodisperse chains with more than 100 units but exhibited extremely complex dissociation spectra. In these polymers, two repeating units that differ by a simple H/CH₃ variation were defined as the 0- and 1-bit of the ASCII code and spaced by a phosphate moiety. They were readily ionized in the negative ion mode electrospray but dissociated *via* cleavage at all phosphate bonds upon collisional activation. Although allowing a complete sequence coverage of digital poly(phosphodiester)s, this fragmentation behavior was not efficient for macromolecules with more than 50 co-monomers and data interpretation was very tedious. The structure of these polymers was then modified by introducing alkoxyamine linkages at appropriate location throughout the chain. A first design consisted of placing these low dissociation energy bonds between each monomeric bit: while cleavage of this sole bond greatly simplified MS/MS spectra, efficient sequencing was limited to chains with up to about 50 units. In contrast, introduction of alkoxyamine bonds between each byte (*i.e.*, a set of eight monomers) was a more successful strategy. Long messages (so far, up to 8 bytes) could be read in MS³ experiments, where single-byte containing fragments released during the first activation stage were further dissociated for sequencing. The whole sequence of such byte-truncated poly(phosphodiester)s could be easily re-constructed based on a mass tagging system which permits to determine the original location of each byte in the chain.

Introduction

Information-containing macromolecules are particularly appealing for their huge storage capabilities, as best exemplified with the archetypal example of DNA storing the whole genome of any living species based on a defined sequence of four repeating units (nucleotides). Use of artificial DNA to write intentional messages on nucleic acid chains was recently demonstrated,^[1-4] after pioneering works by Church and co-workers who reported a full biology text-book written with chains of 100-150 repeating units.^[1] Alternatively, synthetic polymers can also be specifically designed to store digital information, providing that they exhibit a defined sequence of coding units.^[5-9] For this purpose, the simplest binary code was chosen, and a set of two synthetic co-monomers defined as the 0-bit and the 1-bit of the ASCII code was typically used to store digital information in co-polymeric chains. Macromolecules with such a defined co-monomeric sequence shall be monodisperse samples and can obviously not be prepared using classical chain-growth or step-growth polymerization techniques that lead to distributions of chain lengths.^[9] However, such polymers can be prepared relatively easily using multi-step growth approaches,^[9] in which coded monomers are linked one-by-one to a solid support.

Reading messages chemically encoded in chains typically relies on sequencing methods.^[10] Because the 0/1 coding system used in synthetic digital polymers was based on mass, most often with a simple H/CH₃ variation between the two co-monomeric units, tandem mass spectrometry appeared as the key sequencing method to "read" information "written" in the chains. As established for example in the field of proteomics where MS/MS is used to characterize the primary structure of peptides,^[11, 12] dissociation reactions that proceed *via* competitive cleavages of backbone bonds give rise to product ions that contain one or the other chain termination and differ in mass by a single building unit. Accordingly, as long as dissociation products retain the original co-monomer connectivity (that is, when bond cleavages occur without rearrangement), MS/MS data can be usefully employed to determine the nature and relative location of building units in any dissociating chains. However, in contrast to biopolymers such as peptides,^[13] oligonucleotides,^[14] or carbohydrates,^[15] fragmentation rules have to be established for each type of synthetic polymers, since their dissociation behavior is dictated by the types of bonds composing their skeleton.^[16] Moreover, unlike peptides which can be identified even from a partial sequence of amino acids after searching appropriate databases,^[17] total sequence coverage is mandatory for sequence defined synthetic polymers, particularly when dealing with ASCII-based coded messages. Indeed, the set of 8 bits a byte

(*i.e.*, a byte) used to code a given letter may only differ by location of one 1-bit compared to the following letter, as exemplified with **01100001** coding for "a" *versus* with **01100010** coding for "b" (Supplemental Table S-1).

In this context, a usual approach to the selection of best candidates for data storage consists of developing the synthesis protocol of any new sequence-defined polymers interactively with MS/MS, in order to check whether the fragmentation behavior of short chains (*i.e.*, oligomers with n below 10) is promising for the sequencing of longer chains (typically with $n \geq 100$). This methodology permits to rationalize synthesis efforts on the way to long sequence-controlled polymeric chains. However, it should be bear in mind that multi-step growth approaches only allow synthesis of near-monodisperse polymers if the successive reaction steps are defect-free and high yields. Thus, although the synthesis of short sequence-coded oligomers is relatively simple, the preparation of longer coded chains is a more difficult task that requires optimal coupling steps as well as capping steps that permit to eliminate truncated sequences. In that regard, several types of digital oligomers that have been recently described by our group, such as oligo(triazole amide)s,^[18-20] oligo(alkoxyamine amide)s^[21-24] and oligourethanes,^[25, 26] have not yet been prepared as long polymers. In great contrast, we have shown that synthetic poly(phosphodiester)s can easily be obtained as long digital chains. Indeed, these polymers are synthesized using convenient phosphoramidite iterative protocols that have been initially developed for DNA chemical synthesis.^[27, 28] Thus, they can be easily synthesized on an automated DNA synthesizer, which allows preparation of monodisperse chains containing more than hundred coded residues.^[29] Yet, among all reported classes of digital polymers, poly(phosphodiester)s are probably not the easiest macromolecules to sequence by MS/MS. For example, it was found that oligourethanes^[25] and oligo(alkoxyamine amide)s^[22] are particularly easy to read by MS/MS because they lead to very simple fragmentation pattern in collision-induced dissociation (CID) conditions. In comparison, the sequencing of (long) poly(phosphodiester)s chains has required more optimization and research and, in particular, a synergy between polymer synthesis and polymer characterization. In other words, based on the information given by MS/MS, the molecular structure of the polymers can be "re-engineered" to achieve the best MS/MS readability. In this context, the aim of the present feature article is to summarize our recent progress for the design of "easy-to-read" poly(phosphodiester)s. Hence, we describe here dissociation reactions experienced by these sequence-defined synthetic polymers when collisionally activated as deprotonated species, limitation associated with their MS/MS sequencing and alternative designs that were proposed to improve their MS/MS readability.

Experimental section

Materials

Sequence-coded poly(phosphodiester)s were synthesized using iterative phosphoramidite protocols as described in previous publications.^[27, 29] In this approach, two slightly different phosphoramidite monomers are used as a molecular binary code. Some of the digital poly(phosphodiester)s have been synthesized manually on crosslinked polystyrene resins^[27] and others automatically on porous glass supports.^[29] Poly(phosphodiester)s with pending groups were synthesized using two phosphoramidite monomers containing either alkyne or triisopropylsilyl-protected alkyne side groups to allow control on side-chain information by a simple post-polymerization modification method.^[30] Poly(alkoxyamine phosphodiester)s have been synthesized using an orthogonal strategy that employs successively phosphoramidite and radical-radical coupling steps.^[31] Poly(phosphodiester)s containing inter-byte alkoxyamine spacers were prepared on a DNA synthesizer using porous glass as support.^[32]

Polymer samples (a few mg) subjected to electrospray ionization (ESI) were first solubilized in methanol (SDS, Peypin, France) and then diluted (1/10 to 1/1000, v/v) in methanol supplemented with ammonium acetate (Sigma Aldrich) at a 3 mM concentration level. Poly(ethylene glycol) (PEG) and poly(methylmethacrylate) (PMMA) used as internal standards for accurate mass measurements were from Sigma Aldrich.

Mass spectrometry

Negative ion mode ESI was performed for all polymer samples, introduced in the source at a 10 $\mu\text{L min}^{-1}$ flow rate using a syringe pump. High resolution MS and MS/MS experiments performed for poly(phosphodiester)s and poly(alkoxyamine phosphodiester)s were conducted on a QStar Elite mass spectrometer (SCIEX, Concord, ON, Canada) using a -4200 V ESI capillary voltage. Air was used as nebulizing gas (10 psi) while nitrogen was used as curtain gas (20 psi) and collision gas. Ions were measured using an orthogonal acceleration time-of-flight (oa-TOF) mass analyzer operated in the reflectron mode. In CID experiments, precursor ions were selected in the quadrupole mass analyzer, injected into the collision cell, and product ions were measured in the oa-TOF. Instrument control, data acquisition and data processing were achieved using Analyst software (QS 2.0) provided by SCIEX. Pseudo-MS³ experiments implemented for the sequencing of byte-truncated poly(phosphodiester)s were also performed with a Waters Synapt G2 HDMS instrument (Manchester, UK), using the following ESI source parameters: capillary voltage: -2.27 kV; desolvation gas: N₂, heated at 35°C and flowing at 100

L h⁻¹; cone voltage: -10 V or -20 V for both MS and MS² experiments, and from -20 V to -50 V to induce in-source fragmentation in pseudo-MS³ experiments. Collision-induced dissociation (CID) was performed in the ion trap device using argon as the collision gas after selection of the primary (in MS²) or secondary (in MS³) precursor ion in the quadrupole mass analyzer of the instrument. Data analyses were conducted with the MassLynx 4.1 program provided by Waters. For both instruments, internal calibration was performed with acetate adducts of PEG or PMMA molecules in the MS mode, while precursor ions were used as references for accurate measurements of product ions in CID experiments.

Results and discussion

Digital poly(phosphodiester)s are composed of two co-monomers containing a phosphate group linked to a small alkyl segment used to implement a molecular code, that is, propyl for **0** and 2,2-dimethylpropyl for **1**. Early synthesis developments made use of manual iterative protocols, leading to polyphosphates with DP up to 24 and hydroxyl groups at both ends (**P_n** in Scheme 1). By decreasing duration of each monomer addition step from 90 to about 5 min, automation of the phosphoramidite approach on commercially available DNA synthesizer^[33, 34] greatly facilitated production of much longer chains with 104 co-monomeric units, that is, 13 bytes of information.^[29] Besides experimental conditions, the design of macromolecules was also slightly modified to accommodate a segment composed of three thymine (T) nucleotides in order to allow quantitative polymer analysis by UV spectroscopy. As shown in Scheme 1, these **P_n** polyphosphates hold OH as the α group and TTT as the ω termination.

Assessing sample monodispersity in the MS mode

A first key requirement for sequence-defined polymers to be used for information storage is their monodispersity, which could readily be assessed in the MS mode. Polyphosphates were easily produced as deprotonated molecules in the negative ion mode. However, these polyacidic chains are also prone to H/Na exchanges, leading to complicated mass spectra when using MALDI as the ionization technique due to sodium impurities commonly present in matrices.^[27] This effect could be limited to a large extent when performing ESI of polyphosphate solution supplemented with ammonium acetate.^[35] As illustrated by a set of representative examples in Fig. 1, main ESI-MS signals were indeed always observed for deprotonated chains, while ions formed after proton/cation exchanges in solution (as designated by # in Fig. 1) remained minor. Mass spectra obtained upon ESI of polyphosphates always exhibited a charge state distribution,

even for oligomers as short as **P**'s detected as singly and doubly deprotonated chains (Fig. 1a). Charge states further rapidly increased with polymer DP, and ranged from -5 to -12 when n=24 (Fig. 1b), from -11 to -23 when n=56 (Fig. 1c), and from -17 to -41 when n=104 (Fig. 1d). Overall, preferential charge states roughly corresponded to one deprotonated phosphate group out of three. In all cases, mass data obtained after charge state deconvolution of these mass spectra (insets of Fig. 1) univocally confirmed that targeted polyphosphates were produced as highly monodisperse samples.

MS/MS sequencing to read digital information

Dissociation reactions experienced by sequence-defined polyphosphates when activated by collisions with an inert gas were first established for deprotonated oligomers, prior to verify whether the same fragmentation rules still apply as the chain length increases. As shown for $[\mathbf{P}'_5 - 2\text{H}]^{2-}$ in Fig. 2a, oligo(phosphodiester)s typically dissociated *via* efficient cleavage of all phosphate bonds, leading to four sets of complementary product ions allowing coverage of the whole sequence. This is in great contrast to CID of DNA which mostly generate w-type product ions together with numerous less useful secondary fragments.^[14] While the nomenclature established for synthetic polymers states that letters to be used to designate fragments depend on the number of bonds in repeating units,^[16] we chose here to consider the number of dissociating bonds rather than the total number of bonds (Supplemental Scheme S-1). As depicted in fragmentation schemes of Fig. 2, this will allow the same nomenclature to be employed for MS/MS sequencing of any sequence-defined poly(phosphodiester)s, regardless of the length of the alkyl segment used to encode binary information. Due to the symmetric structure of **P**'s which codes for **01010**, the same m/z values were respectively measured for a_i^{z-}/z_i^{z-} , b_i^{z-}/y_i^{z-} , c_i^{z-}/x_i^{z-} , and d_i^{z-}/w_i^{z-} in the CID spectrum shown in Fig. 2a. In contrast, formation of the eight fragment series was clearly evidenced during CID of the **00101** oligomer, which exhibited the same co-monomeric composition but a different sequence compared to **P**'s (Fig. 2b; see also Supplemental Figure S-1 for detailed assignment of low abundance product ions). Interestingly, while the same dissociation reactions occurred for both isomers, their respective rate was observed to depend on their **0/1** sequence. In particular, a very different survival yield was measured for each doubly deprotonated precursor although activated in the same experimental conditions, with a much weaker relative abundance of the m/z 341.1 ion with the **00101** sequence. This might be due to steric hindrance of methyl side-groups in **1** repeating units affecting some of the proton transfers requested for phosphate bond cleavages (Supplemental Scheme S-1). However, opposite results were observed during CID of singly

deprotonated oligomers (Supplemental Figure S-2), with a much lower survival yield measured for the m/z 683.2 precursor ion generated from **01010** compared to **00101**. Since the number of deprotonated phosphate groups in poly(phosphodiester)s rapidly increased with their length (Fig. 1), influence of charge state on dissociation reactions was not further investigated. Nevertheless, for short oligo(phosphodiester)s, cleavage occurred at bonds in all co-monomers along the backbone, hence fulfilling the key requirement for full sequence coverage.

The same fragmentation pattern was also observed for oligo(phosphodiester)s holding side chains. These oligomers were synthesized using two phosphoramidite monomers containing either alkyne or triisopropylsilyl-protected alkyne side groups (respectively designated as **0** and **1** in the structure shown on top of Fig. 3a). This design allowed each type of repeating units to be modified in a stepwise manner, as depicted on top of Fig. 3b after **0** units have been reacted with an azide compound.^[30] Performing the same reaction with a different reagent after deprotection of **1** units allowed production of sequence-controlled polymers with readable side-chains (not shown) amenable to nanopore sequencing.^[36, 37] CID data obtained for the doubly deprotonated **1001** oligomer prior (Fig. 3a) and after (Fig. 3b) modification of **0** units did not exhibit new signals but the same a_i^{z-} , b_i^{z-} , c_i^{z-} and d_i^{z-} ion series as for “naked” oligo(phosphodiester)s (z_i^{z-} , y_i^{z-} , x_i^{z-} , and w_i^{z-} ion series respectively expected at the same m/z values as their α -containing congeners in these symmetric species were not annotated for the sake of clarity). These results clearly showed that studied pending groups neither had any influence on the fragmentation behavior of the polymer backbone nor underwent any dissociation reactions, probably due to specific charge-assisted mechanisms at work in these polyanions.

More importantly, fragmentation rules established for short oligomers also applied to poly(phosphodiester)s of increasing length, as illustrated by the full sequence coverage obtained for the **P24** polymer (Fig. 4a) after CID data analysis (Supplemental Figure S-3). Activation of this $[\mathbf{P24} - 8\mathbf{H}]^{8-}$ precursor ion at m/z 550.9 produced fragments in the eight expected ion series, most often detected at multiple charge states. It can however be noticed from data presented in Fig. 4a that complementary product ions were not systematically detected. For example, w_i^{7-} ions with $i=20-23$ expected to be formed complementarily to a_i^{1-} with $i=5-2$, respectively, were not observed. This strongly suggested that fragments with highest charge state and size would readily further dissociate, hence accounting for the increasing number of secondary product ions observed in MS/MS spectra of large poly(phosphodiester)s (Supplemental Figure S-3). Data shown in Fig. 4a also exhibited some “holes” in the sequence coverage. Most often, this was due to signal interferences preventing safe assignment of low

abundance fragments. As illustrated in Fig. 4b, peaks of the w_{10}^{5-} ion (annotated in purple) were detected within the isotopic cluster of a more abundant triply charged secondary fragment (annotated in black). Although these signals matched pretty well the theoretical isotopic pattern expected for w_{10}^{5-} , this would have been a very unsafe assignment in case of unknown sequence, and identification was hence not validated here.

As expected, spectral complexity further increased with the size of poly(phosphodiester)s, and the upper DP limit for full sequence coverage seemed to be around 50, as shown by the example of 56-mer coding for “Maurice” (Supplemental Figure S-4). In this example, assignment of co-monomers ranking from positions 26 to 34 in this \mathbf{P}_{56} sequence mainly relied on safe assignment of fragments at the -7 charge state from only two ion series. Of note, these were the most highly charged dissociation products detected in the MS/MS spectrum of this $[\mathbf{P}_{56} - 17\text{H}]^{17-}$, again showing occurrence of extensive secondary fragmentations. As previously mentioned in the Introduction, the structure of sequence-controlled synthetic polymers can be manipulated to achieve the most simple CID pattern. This approach was hence investigated to improve MS/MS readability of long poly(phosphodiester)s while keeping advantages offered by their automated synthesis.

Design of sequence-defined poly(alkoxyamine phosphodiester)s

As learnt from CID of oligo(alkoxyamine amide)s, C–ON bonds have very low bond dissociation energies particularly when nitroxide moieties are connected to highly substituted carbon atoms.^[21, 22, 38] Such alkoxyamine linkages hence appeared as valuable options to introduce weak points in the backbone of poly(phosphodiester)s and make coded messages they contain more easy-to-read by MS/MS.

A first strategy consisted of placing alkoxyamine bonds between each phosphodiester moieties, by implementing two chemoselective reactions in a successive manner, namely phosphoramidite coupling and nitroxide radical coupling.^[31] As recently described in details,^[39] structure of these poly(alkoxyamine phosphodiester)s had to be optimized to allow sufficient distance between consecutive phosphate groups for all of them to be deprotonated simultaneously. Indeed, this feature was experimentally found to be a key condition to obtain extremely simple MS/MS pattern (*vide infra*). This was achieved with the architecture shown in Scheme 2, which contains a TEMPO-hexanamide nitroxide connected to a tertiary carbon, and n-butyl (R=H) or 2-methyl-butyl (R=CH₃) groups as segments coding for **0** or **1**, respectively. For a first proof of concept, these sequence-controlled macromolecules were

synthesized manually, and hold a hydroxypropyl moiety as the α termination and a hydroxyl group as ω .

As expected, low activation of deprotonated poly(alkoxyamine phosphodiester)s mainly induced competitive homolysis of all alkoxyamine bonds, leading to complementary radical anions named c_i^{z-} when containing α and y_i^{z-} when containing ω . However, the peculiar MS/MS pattern systematically observed for fully deprotonated poly(alkoxyamine phosphodiester)s arose from their specific structural design.^[39] As illustrated in Fig. 5 with CID data obtained for the **01001110** oligomer, all c_i^{z-} ions (in brown) were detected at the left-hand side of the precursor ion peak and their m/z increased with their size, while the whole y_i^{z-} ion series (in purple) was observed on the other side with m/z decreasing as their size increased. Because each repeating unit had a deprotonated phosphate in fully deprotonated precursor, the charge state z of any fragment equaled the number i of monomers they contained. As a result, the $\Delta m/z$ difference measured between two consecutive members in each fragment series, $F_{i+1}^{(i+1)-}$ and F_i^{i-} , could be mathematically defined as

$$\Delta m/z = m/z(F_{i+1}^{(i+1)-}) - m/z(F_i^{i-}) = \frac{m_{\text{monomer}}}{i+1} - \frac{m/z(F_i^{i-})}{i+1} \quad (\text{Eq. 1})$$

where m_{monomer} is the mass of the deprotonated monomer at the $(i+1)^{\text{th}}$ position, that is, 504.2 Da for **0** or 518.2 Da for **1**. In addition, the particular location of C–ON linkages in oligo(alkoxyamine phosphodiester)s had two main consequences: as can be seen in the structure of the **01001110** oligomer shown on top of Fig. 5, the mass of deprotonated monomers is i) larger than the mass of the molecular segment at the left-hand side of the first cleavable C–ON bond (*i.e.*, 281.1 Da) but ii) lower than the mass of the segment at the right-hand side of the last cleavable C–ON bond (*i.e.*, 549.2 Da). As a result, values obtained for $\Delta m/z$ according to Eq. 1 were i) always positive for c_i^{z-} ions while ii) always negative for y_i^{z-} ions. Consequently, m/z values of c_i^{z-} fragments increased with their size while remaining lower than that of the precursor ion, whereas m/z values of y_i^{z-} fragments decreased as their size increased while remaining higher than that of the precursor ion. Overall, these spectral features allowed product ions to be readily assigned, and hence the **0/1** sequence of any poly(alkoxyamine phosphodiester)s to be characterized in a straightforward manner by simple application of Eq. 1. For example, the first unit in the chain was identified from the m/z value measured for c_1^{1-} , which can only be either 295.1 or 309.1 for sequence beginning as α -**0** or α -**1**, respectively. Then, $\Delta m/z = 111.6$ measured between c_2^{2-} (m/z 406.7) and c_1^{1-} indicated a mass of 518.3 Da for the second monomer, hence identified as **1**. Applying the same operation throughout the

c_i^{z-} series allowed the digitally-coded **01001110** message to be read from the left- to the right-hand side, while a similar analysis of y_i^{i-} product ions permitted to re-construct the sequence from the ω to the α chain-end.

Such a particular MS/MS pattern, with $\Delta m/z$ rapidly decreasing as the size of product ions increases, will however lead to spectral congestion and require high resolving capabilities for safe assignment of highly charged fragments expected to be formed from very long chains. Based on isotopic simulations and owing to the resolving power offered by the QqTOF instrument used in this study, structures alternating phosphate groups with nitroxide moieties would remain advantageous for MS/MS sequencing of chains with DP up to about 50.^[39] Although messages digitally encoded in these poly(alkoxyamine phosphodiester)s were always much easier to decipher than those written in poly(phosphodiester)s of similar DP, such a molecular design would not allow targeted polymers with more than 100 repeating units to be efficiently sequenced.

Byte-truncated poly(phosphodiester)s

Still taking advantage of the very low dissociation energy of alkoxyamine linkages which made any other chemical bonds MS/MS-silent, an alternative design was developed to further improve sequencing of longer polyphosphate chains. As depicted in Scheme 3a, these weak bonds were placed between each byte of information instead of each bit, in order to perform i) a first activation stage to release intact individual bytes upon cleavage of all inter-byte alkoxyamine bonds, and ii) a second activation stage to sequence these primary fragments according to polyphosphate dissociation rules, which should be an easy task for these oligo(phosphodiester) 8-mers.^[32]

In order to retrieve original location of bytes released during the first activation step, specific tags were used to label each byte. These tags were carefully chosen to create an unequivocal mass signature based on the following two main criteria: i) neither the mass of a byte-tag nor the mass difference between two byte-tags shall be a multiple of 28, which is the mass difference between **1** (166.0 Da) and **0** (138.0 Da) synthons used here, and ii) the mass difference between two byte-tags shall not be smaller than 3 Da since triply charged species were studied in MS³ (*vide infra*). As shown in Scheme 3c, selected byte-labels were nucleotides which offered the additional key advantage of being incorporated at specific locations in the chain with no modification of the automated synthesis protocol. Obviously, the byte-tag sequence defined by the “writer” has to be known by the “reader”.

This concept was successfully applied to MS³ sequencing of digital polymers containing up to 8 bytes of information (that is, 64 co-monomeric units),^[32] and is illustrated in Fig. 6 for a 4-byte polymer that contained the ASCII-encoded word “Byte”, using the following byte-tag sequence: “no tag”, A, C, T. Although the most stable charge state for this polymer was $z = 11$ as revealed by ESI-MS (Fig. 6a), the $[M - 12H]^{12-}$ ion was selected for fragmentation in order to favor production of byte fragments at a preferential -3 charge state from this 4-byte precursor. Since cleavage of alkoxyamine bonds in such precursor ions was expected to occur in a competitive manner, collision energy in this first activation stage was optimized to promote consecutive dissociation, in order to generate all targeted single-byte fragments, while preventing fragmentation of phosphate bonds within bytes. Accordingly, MS/MS spectra mainly exhibited product ions arising from C–ON bond homolysis (Fig. 6b) and found, based on their charge state, to contain 1 ($z=3$), 2 ($z=6$) or 3 ($z=9$) bytes. Thanks to the implemented mass tag system, original location of molecular segments detected as triply deprotonated species was readily identified, with byte 1 in m/z 433.7, byte 2 in m/z 644.1, byte 3 in m/z 626.8, and byte 4 in m/z 563.1 (Supplemental Table S-3).

In order to take advantage of the resolving capabilities of oaTOF analyzer for safe assignment of fragments in MS³ sequencing, the second activation stage was also conducted in the QqTOF instrument used here to perform MS and MS/MS experiments. Typically, raising the cone voltage allowed deprotonated precursors to undergo in-source CID (first activation step), and so-released triply charged byte-fragments were further selected for activation in the collision cell (second activation step). Analysis of the four pseudo-MS³ spectra recorded for the 4-byte polymer shown in Fig. 6 led to the sequence coverage charts presented in Fig. 7. The **0/1** sequence of all bytes was accurately characterized, with at least three ion series allowing full coverage in each case. Interestingly, very partial series were systematically inventoried for fragments w_i^{z-} , x_i^{z-} , y_i^{z-} , and z_i^{z-} when dissociating ions hold a carbon-centered radical ω termination, that is, for all bytes but the last one. This result suggested that alternative radical-driven reactions efficiently competed with phosphate bond cleavages that led to ω -containing fragments, as supported by abundant new product ions specifically observed in pseudo-MS³ spectra recorded for these bytes (Supplemental Figure S-5). Future efforts will focus on a new design of this molecular moiety in order to limit these alternative dissociation pathways. In contrast, high stability of the TEMPO nitroxide found at the α end of all bytes but the first one did not exhibit any reactivity upon activation, allowing (nearly) complete a_i^{z-} , b_i^{z-} , c_i^{z-} , and d_i^{z-} series to be inventoried for useful byte sequencing.

Conclusion

Synthetic digital polymers have recently emerged as a promising new class of macromolecules to store information at the molecular level. Besides binary codes making use of two comonomers and, for example, arbitrarily defined as the 0- and 1-bits of the ASCII alphabet, a variety of more sophisticated coding systems can be envisaged to increase storage capacity of polymeric chains. Because information is encrypted in the primary structure of digital polymers, efficiency of reading methodologies mainly relies on their ability to accurately decipher comonomer connectivity in the coded chains. Sequencing capabilities of MS/MS have long been proved effective but complexity of MS/MS pattern highly depends on the nature of backbone bonds. However, in contrast to biopolymer structure dictating their biologic functions, digital polymers do not have to exhibit any specific properties but to deliver messages. This offers a great degree of freedom as their structure can be manipulated so as to achieve the best MS/MS readability. As demonstrated for poly(phosphodiester)s, distribution of specific chemical moieties at appropriate locations in the chain allowed dissociation pathways to be modulated on purpose. As a result, using the great variety of experiments offered by mass spectrometry, information reading could be greatly simplified. However, throughput of this crucial step is still to be improved. On the one hand, although not requiring a long preparation protocol, digital polymers have to be solubilized in the appropriate salt solution and transferred in a syringe for being introduced in the ESI source. Use of ambient ionization techniques such as DESI would definitively be a much faster approach. In addition, imaging capabilities of such desorption techniques would allow fast reading of much longer messages composed of different sequence-controlled polymers distributed onto a surface. On the other hand, data reading has to be further accelerated, using dedicated software currently developed in our group for automated MS/MS spectra interpretation. Alternative activation methods that induce more specific dissociation reactions compared to CID should also be explored. In particular, photo-fragmentation processes are particularly appealing as they should allow controlled cleavages of bonds in sequence-controlled polymers decorated with specifically designed chromophores. Finally, while efforts still need to be focused on the preparation of long MS/MS readable chains for data storage, a promising application of small sequence-defined oligomers is their use as molecular tags in anticounterfeiting technologies, as recently demonstrated for high value materials such as intraocular implants tagged with oligourethanes.^[26]

Acknowledgments

The authors thank the French National Research Agency (ANR project 00111001) for financial support. L.C. also acknowledges support from Spectropole, the Analytical Facility of Aix-Marseille University, by allowing a special access to the instruments purchased with European Funding (FEDER OBJ2142-3341). J.F.L. also acknowledges the H2020 program of the European Union (project Euro-Sequences, H2020-MSCA-ITN-2014, grant agreement n°642083). The PhD positions of G.C. and N.K. are supported by the ITN Euro-Sequences.

References

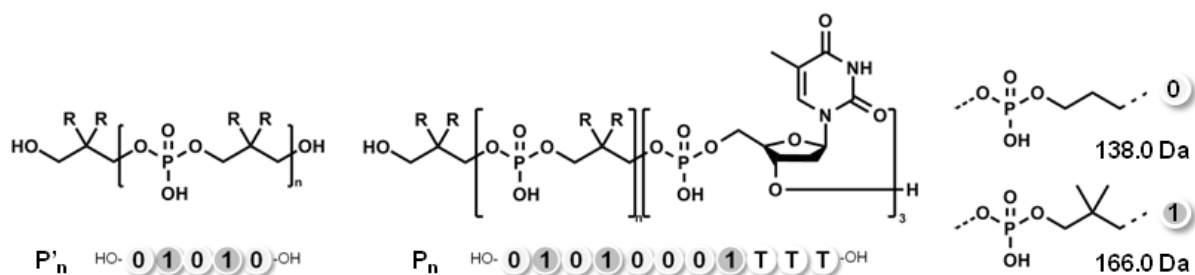
- [1] G. M. Church, Y. Gao, S. Kosuri, Next-Generation Digital Information Storage in DNA. *Science* **2012**, *337*, 1628-1628.
- [2] N. Goldman, P. Bertone, S. Y. Chen, C. Dessimoz, E. M. LeProust, B. Sipos, E. Birney, Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature* **2013**, *494*, 77-80.
- [3] C. Mayer, G. R. McInroy, P. Murat, P. Van Delft, S. Balasubramanian, An Epigenetics-Inspired DNA-Based Data Storage System. *Angew. Chem. Int. Ed.* **2016**, *55*, 11144-11148.
- [4] Y. Erlich, D. Zielinski, DNA Fountain enables a robust and efficient storage architecture. *Science* **2017**, *355*, 950-953.
- [5] J.-F. Lutz, M. Ouchi, D. R. Liu, M. Sawamoto, Sequence-Controlled Polymers. *Science* **2013**, *341*, 1238149.
- [6] J.-F. Lutz, Writing on Polymer Chains. *Acc. Chem. Res.* **2013**, *46*, 2696-2705.
- [7] H. Colquhoun, J.-F. Lutz, Information-containing macromolecules. *Nat. Chem.* **2014**, *6*, 455-456.
- [8] J.-F. Lutz, Coding Macromolecules: Inputting Information in Polymers Using Monomer-Based Alphabets. *Macromolecules* **2015**, *48*, 4759-4767.
- [9] J.-F. Lutz, J.-M. Lehn, E. W. Meijer, K. Matyjaszewski, From precision polymers to complex materials and systems. *Nat. Rev. Mater.* **2016**, *1* (5) 16024.
- [10] H. Mutlu, J.-F. Lutz, Reading Polymers: Sequencing of Natural and Synthetic Macromolecules. *Angew. Chem. Int. Ed.* **2014**, *53*, 13010-13019.
- [11] E. Phizicky, P. I. H. Bastiaens, H. Zhu, M. Snyder, S. Fields, Protein analysis on a proteomic scale. *Nature* **2003**, *422*, 208-215.
- [12] M. Tyers, M. Mann, From genomics to proteomics. *Nature* **2003**, *422*, 193-197.

- [13] K. Biemann, Laying the groundwork for proteomics - Mass spectrometry from 1958 to 1988. *Int. J. Mass Spectrom.* **2007**, *259*, 1-7.
- [14] S. A. McLuckey, G. J. Vanberkel, G. L. Glish, Tandem mass-spectrometry of small, multiply charged oligonucleotides. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60-70.
- [15] B. Domon, C. E. Costello, A systematic nomenclature for carboglycane fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate J.* **1988**, *5*, 397-409.
- [16] C. Wesdemiotis, N. Solak, M. J. Polce, D. E. Dabney, K. Chaicharoen, B. C. Katzenmeyer, Fragmentation pathways of polymer ions. *Mass Spectrom. Rev.* **2011**, *30*, 523-559.
- [17] H. Steen, M. Mann, The ABC's (and XYZ's) of peptide sequencing. *Nat. Rev. Molec. Cell Biol.* **2004**, *5*, 699-711.
- [18] T. T. Trinh, C. Laure, J.-F. Lutz, Synthesis of Monodisperse Sequence-Defined Polymers Using Protecting-Group-Free Iterative Strategies. *Macromol. Chem. Phys.* **2015**, *216*, 1498-1506.
- [19] T. T. Trinh, L. Oswald, D. Chan-Seng, L. Charles, J.-F. Lutz, Preparation of Information-Containing Macromolecules by Ligation of Dyad-Encoded Oligomers. *Chem. Eur. J.* **2015**, *21*, 11961-11965.
- [20] J.-A. Amalian, T. T. Trinh, J.-F. Lutz, L. Charles, MS/MS Digital Readout: Analysis of Binary Information Encoded in the Monomer Sequences of Poly(triazole amide)s. *Anal. Chem.* **2016**, *88*, 3715-3722.
- [21] R. K. Roy, A. Meszynska, C. Laure, L. Charles, C. Verchin, J.-F. Lutz, Design and synthesis of digitally encoded polymers that can be decoded and erased. *Nat. Commun.* **2015**, *6*, 7237.
- [22] L. Charles, C. Laure, J.-F. Lutz, R. K. Roy, MS/MS Sequencing of Digitally Encoded Poly(alkoxyamine amide)s. *Macromolecules* **2015**, *48*, 4319-4328.
- [23] R. K. Roy, C. Laure, D. Fischer-Krauser, L. Charles, J.-F. Lutz, Convergent synthesis of digitally-encoded poly(alkoxyamine amide)s. *Chem. Commun.* **2015**, *51*, 15677-15680.
- [24] C. Laure, D. Karamessini, O. Milenkovic, L. Charles, J.-F. Lutz, Coding in 2D: Using Intentional Dispersity to Enhance the Information Capacity of Sequence-Coded Polymer Barcodes. *Angew. Chem. Int. Ed.* **2016**, *55*, 10722-10725.
- [25] U. S. Gunay, B. E. Petit, D. Karamessini, A. Al Ouahabi, J.-A. Amalian, C. Chendo, M. Bouquey, D. Gigmes, L. Charles, J.-F. Lutz, Chemoselective synthesis of uniform sequence-coded polyurethanes and their use as molecular tags. *Chem* **2016**, *1*, 114-126.

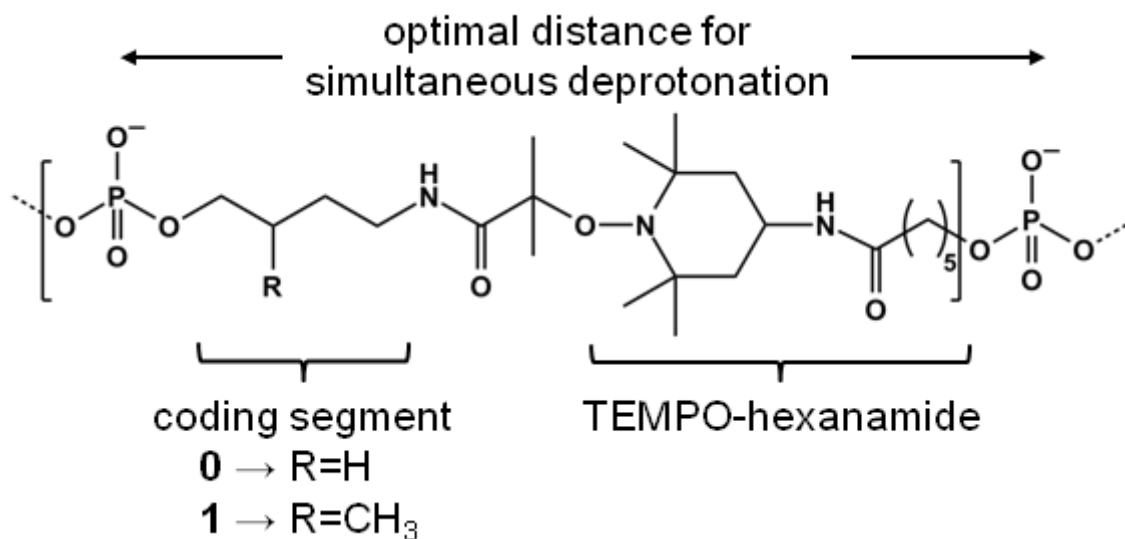
- [26] D. Karamessini, B. E. Petit, M. Bouquey, L. Charles, J.-F. Lutz, Identification-tagging of methacrylate-based intraocular implants using sequence defined polyurethane barcodes. *Adv. Funct. Mater.* **2017**, *27*, 1604595.
- [27] A. Al Ouahabi, L. Charles, J.-F. Lutz, Synthesis of Non-Natural Sequence-Encoded Polymers Using Phosphoramidite Chemistry. *J. Am. Chem. Soc.* **2015**, *137*, 5629-5635.
- [28] L. Oswald, A. Al Ouahabi, L. Charles, J.-F. Lutz, Optimal ATRP-Made Soluble Polymer Supports for Phosphoramidite Chemistry. *Chem. Eur. J.* **2016**, *22*, 3462-3469.
- [29] A. Al Ouahabi, M. Kotera, L. Charles, J.-F. Lutz, Synthesis of Monodisperse Sequence-Coded Polymers with Chain Lengths above DP100. *ACS Macro Lett.* **2015**, *4*, 1077-1080.
- [30] N. F. König, A. Al Ouahabi, S. Poyer, L. Charles, J.-F. Lutz, A simple post-polymerization modification method for controlling side-chain information in digital polymers. *Angew. Chem. Int. Ed.* **2017**, *accepted*. DOI: 1002/anie.201702384R1.
- [31] C. Cavallo, A. Al Ouahabi, L. Oswald, L. Charles, J.-F. Lutz, Orthogonal Synthesis of "Easy-to-Read" Information-Containing Polymers Using Phosphoramidite and Radical Coupling Steps. *J. Am. Chem. Soc.* **2016**, *138*, 9417-9420.
- [32] A. Al Ouahabi, J.-A. Amalian, L. Charles, J.-F. Lutz, Mass spectrometry sequencing of long digital polymers facilitated by programmed inter-byte fragmentation. *Nat. Commun.* **2017**, *submitted*.
- [33] R. Noir, M. Kotera, B. Pons, J. S. Remy, J. P. Behr, Oligonucleotide-oligospermine conjugates (zip nucleic acids): A convenient means of finely tuning hybridization temperatures. *J. Am. Chem. Soc.* **2008**, *130*, 13500-13505.
- [34] T. G. W. Edwardson, K. M. M. Carneiro, C. J. Serpell, H. F. Sleiman, An Efficient and Modular Route to Sequence-Defined Polymers Appended to DNA. *Angew. Chem. Int. Ed.* **2014**, *53*, 4567-4571.
- [35] J. T. Stults, J. C. Marsters, Improved electrospray ionization of synthetic oligodeoxynucleotides. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 359-363.
- [36] D. Branton, D. W. Deamer, A. Marziali, H. Bayley, S. A. Benner, T. Butler, M. Di Ventra, S. Garaj, A. Hibbs, X. H. Huang, S. B. Jovanovich, P. S. Krstic, S. Lindsay, X. S. S. Ling, C. H. Mastrangelo, A. Meller, J. S. Oliver, Y. V. Pershin, J. M. Ramsey, R. Riehn, G. V. Soni, V. Tabard-Cossa, M. Wanunu, M. Wiggin, J. A. Schloss, The potential and challenges of nanopore sequencing. *Nat. Biotechnol.* **2008**, *26*, 1146-1153.
- [37] Y. L. Ying, J. J. Zhang, R. Gao, Y. T. Long, Nanopore-Based Sequencing and Detection of Nucleic Acids. *Angew. Chem. Int. Ed.* **2013**, *52*, 13154-13161.

- [38] L. Charles, C. Laure, J.-F. Lutz, R. K. Roy, 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*, 22-28.
- [39] L. Charles, G. Cavallo, V. Monnier, L. Oswald, R. Szweda, J.-F. Lutz, MS/MS-assisted design of sequence-controlled synthetic polymers for improved reading of encoded information. *J. Am. Soc. Mass Spectrom.* **2017**, *in press*. DOI: 10.1007/s13361-016-1543-5.

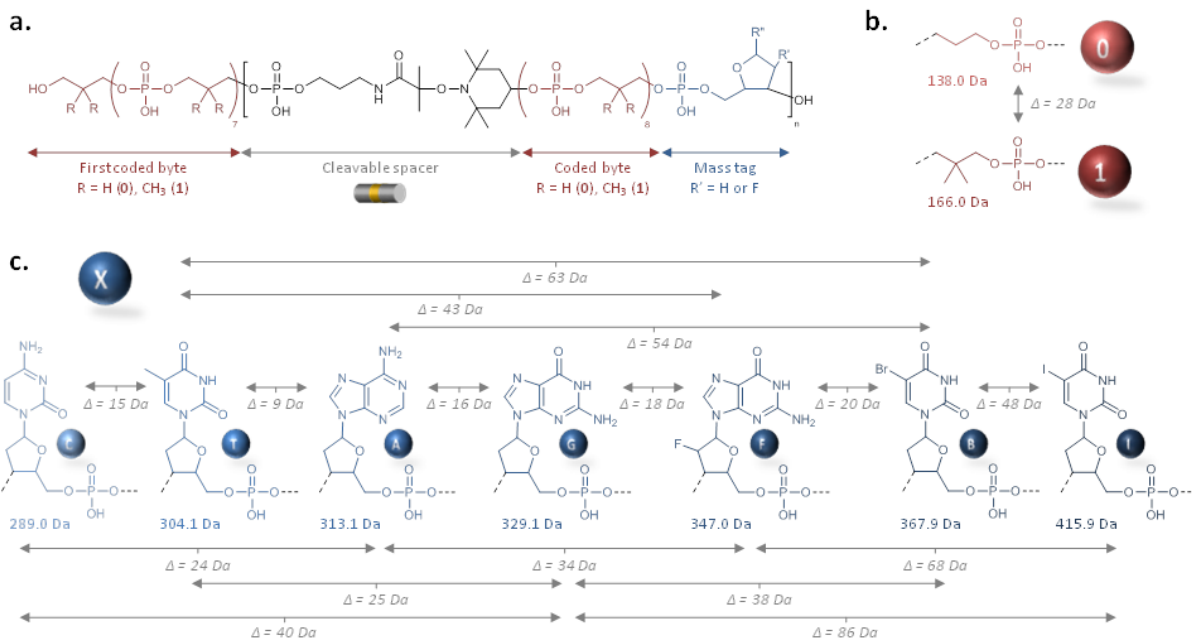
Captions to Figures & Schemes



Scheme 1. Structure of poly(phosphodiester)s prepared manually, with OH at both chain ends (P'_n , left), and obtained with an automated synthesizer, with OH as α and a segment composed of three thymine (T) nucleotides as ω (P_n , middle), both composed of a defined sequence of **0** and **1** repeating units (right).



Scheme 2. Specific design of monomeric units for oligo(alkoxyamine phosphodiester)s to exhibit high gas phase stability when fully deprotonated.



Scheme 3. a) Structural design of byte-truncated poly(phosphodiester)s, with a cleavable spacer (in black) containing an alkoxyamine bond placed between sets of eight coding units (in red, structure of **0** and **1** shown in b) and a mass tag (in blue, selected structures shown in c) to label each coded byte.

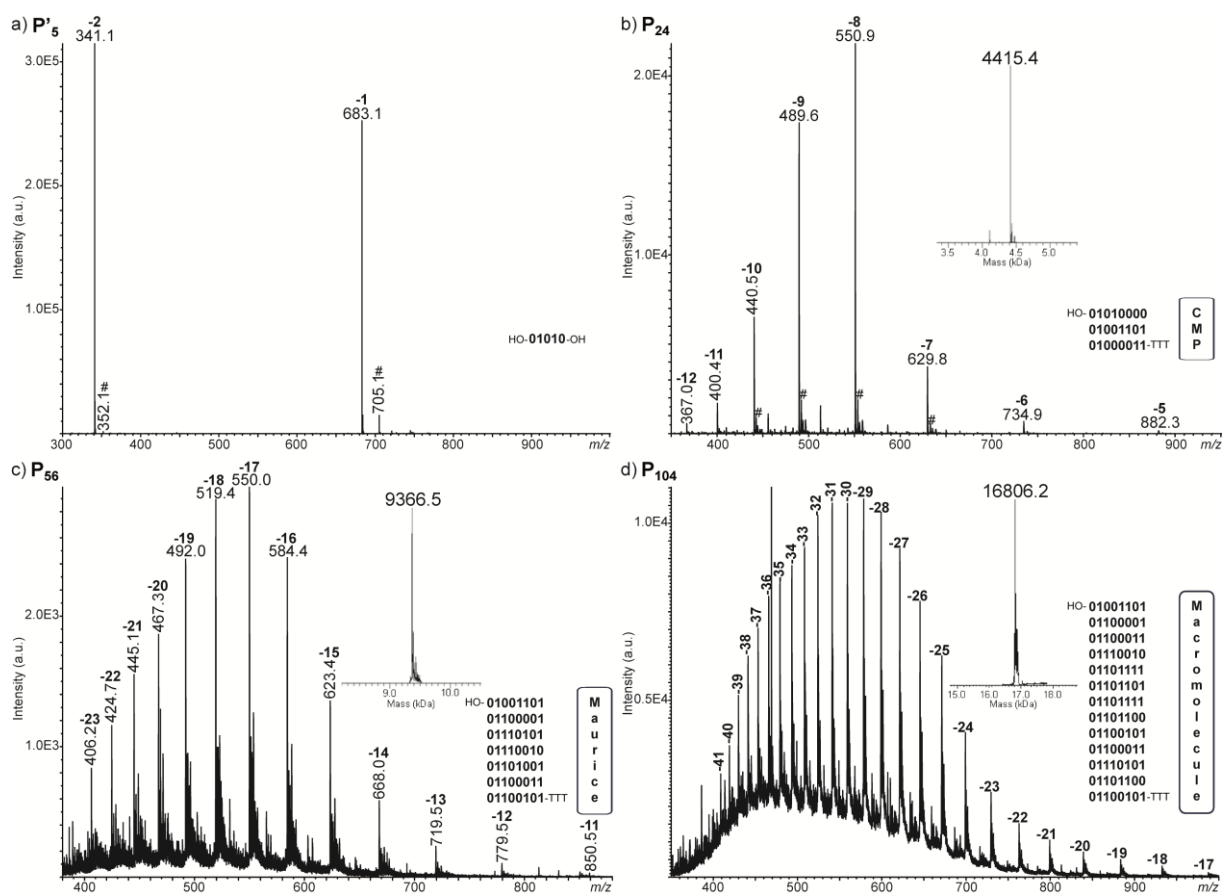


Figure 1. Negative mode ESI mass spectra of sequence-defined phosphodiester oligomer a) **P'5** (684.1478 Da, $C_{19}H_{44}O_{18}P_4$), and polymers b) **P24** (4415.6641 Da at isotopic maximum, $C_{120}H_{244}N_6O_{115}P_{26}$), c) **P56** (9366.5273 Da at isotopic maximum, $C_{254}H_{544}N_6O_{243}P_{58}$), and d) **P104** (16806.8379 Da at isotopic maximum, $C_{456}H_{996}N_6O_{435}P_{106}$). For **P104**, m/z values are listed in Supplemental Table S-2. Deconvoluted mass spectra, **0/1** co-monomeric sequence and associated encoded words are shown in inset. The manually synthesized **P'5** oligomer holds OH as both terminations, whereas **P24**, **P56** and **P104** polymers prepared with an automated DNA synthesizer have their TTT 3'-end of the sequence on the right-hand side, following DNA conventions. # designates signals due to H/Na and H/K exchanges.

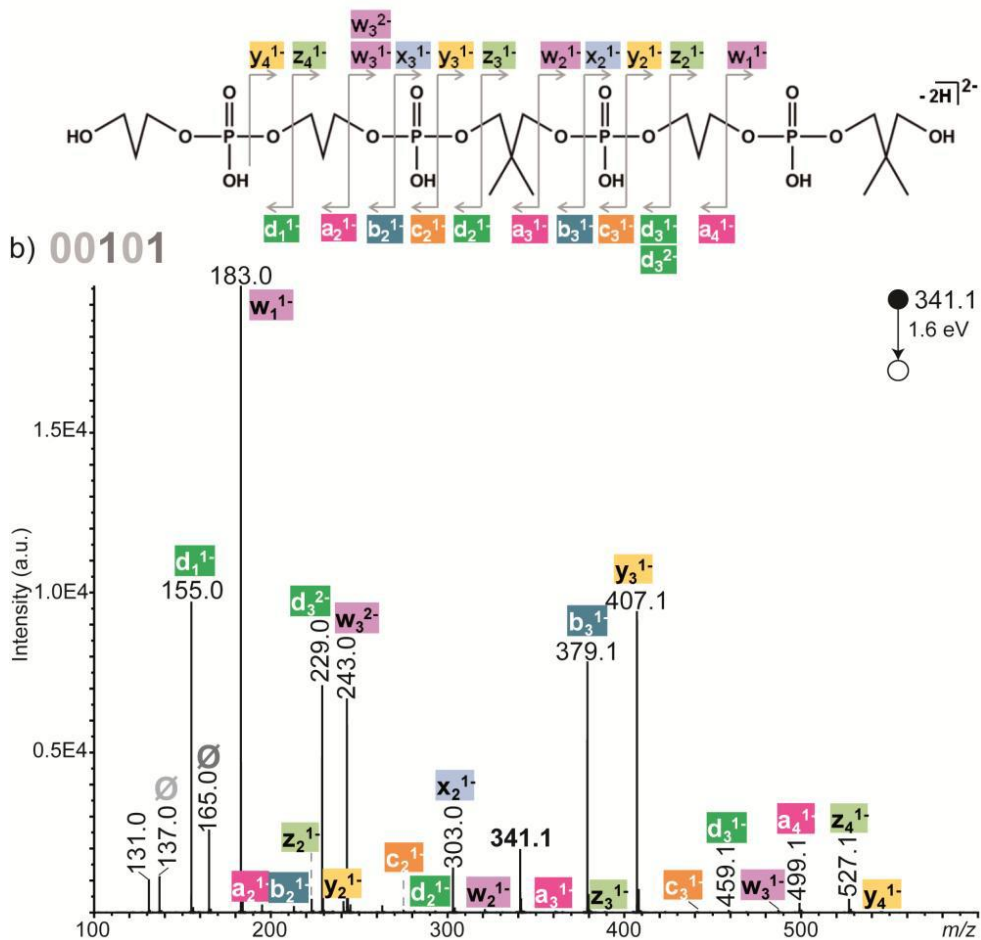
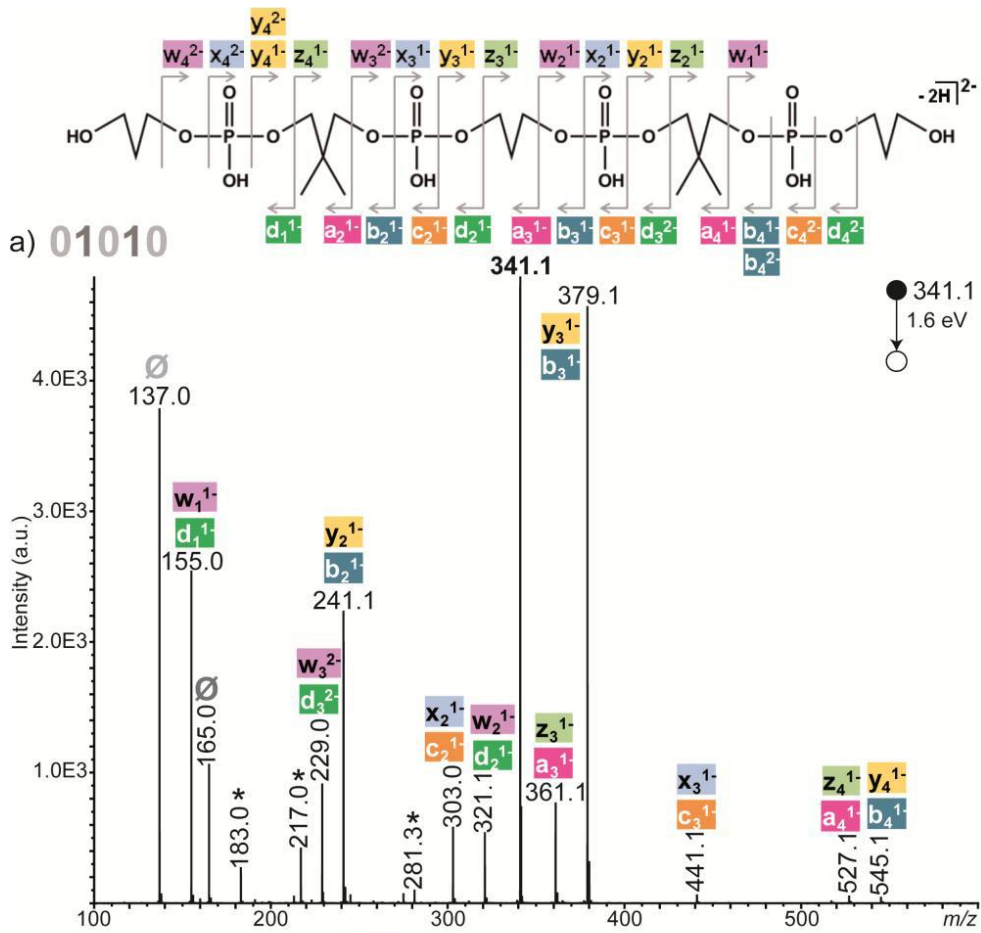


Figure 2. CID spectra of the isomeric oligophosphates a) **01010** and b) **00101**, both produced as doubly deprotonated molecules at m/z 341.1 (in bold), with product ion assignment according to fragmentation schemes shown on top. \emptyset designates deprotonated monomers **0** (m/z 137.0, pale grey) and **1** (m/z 165.0, dark grey) released from primary product ions, while other secondary fragments are annotated with an asterisk.

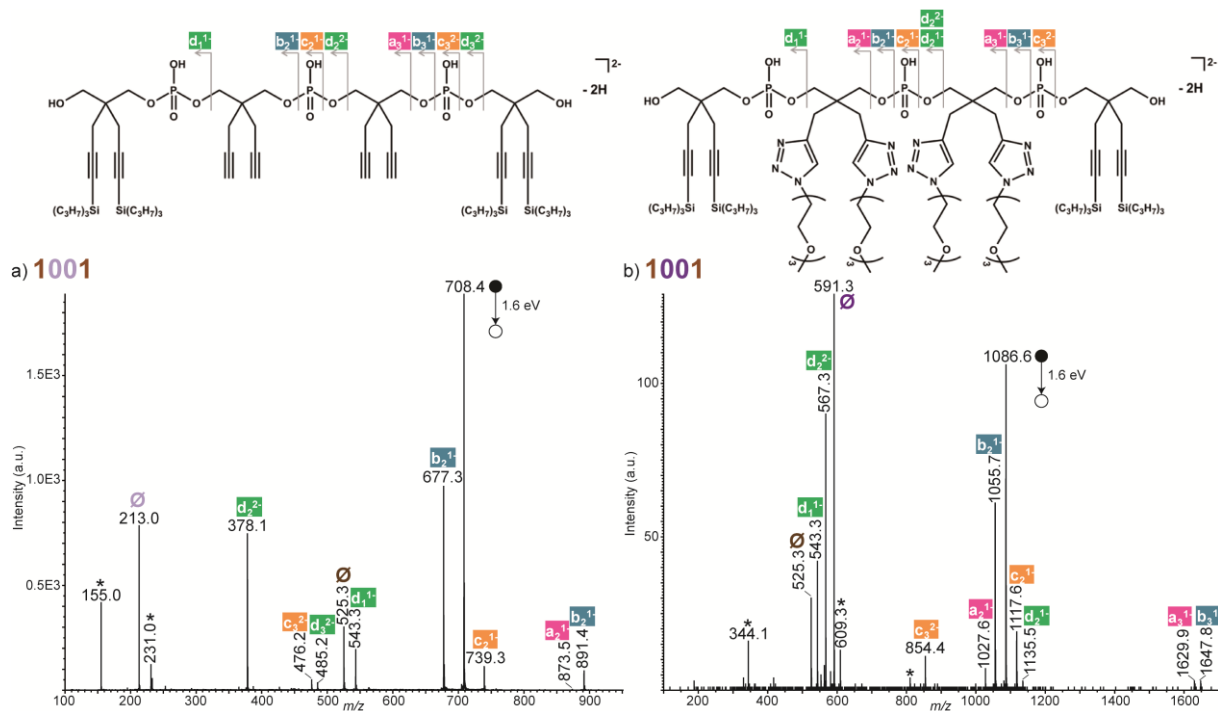


Figure 3. CID spectra of the oligophosphate **1001** a) prior and b) after modification of the **0** units. Both species were produced as doubly deprotonated molecules, at m/z 708.4 for the oligomer with unmodified **0** units (in pale purple) of mass $m_0 = 214.0$ Da and at m/z 1086.6 with modified **0** units (in dark purple) of mass $m_0 = 592.3.0$ Da (the mass of 1 units being $m_1 = 526.3$ Da). Product ion assignment is described in fragmentation schemes shown on top. \emptyset designates deprotonated monomers (unmodified **0**, pale purple: m/z 213.0; modified **0**, dark purple: m/z 591.3; **1**: m/z 525.3) evaporated from primary product ions, while other secondary fragments are annotated with an asterisk.

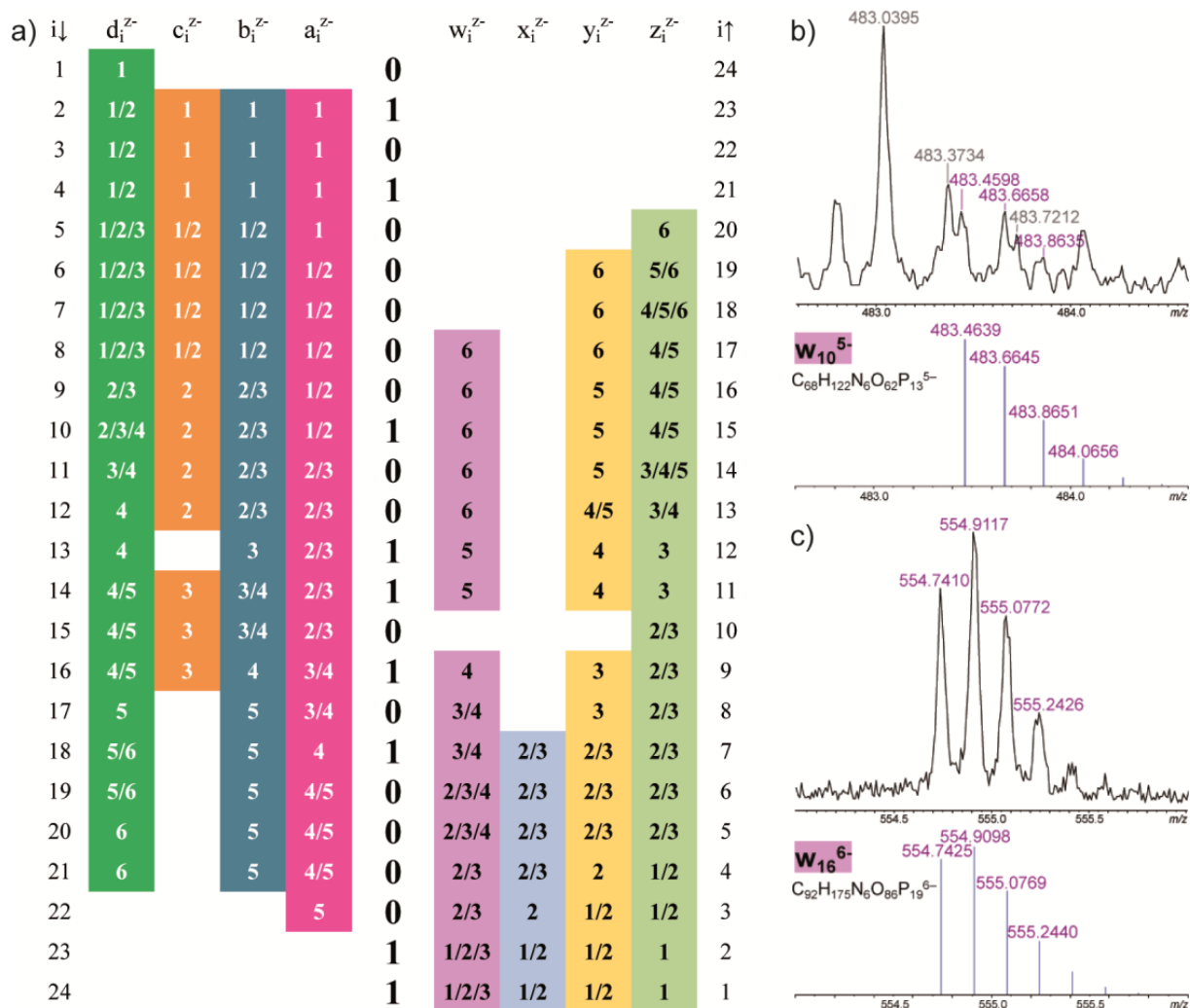


Figure 4. a) Sequence coverage chart obtained for P_{24} after MS/MS of the -8 deprotonated molecule. The full 0/1 sequence, written vertically from top to bottom, was successfully reconstructed by combining fragments containing α (a_i^{z-} , b_i^{z-} , c_i^{z-} , and d_i^{z-} , to be read from top to bottom in left panel) and fragments containing ω (w_i^{z-} , x_i^{z-} , y_i^{z-} , and z_i^{z-} , to be read from bottom to top in right panel), detected with charge states ranging from 1 to 6 as indicated for each product ion (values spaced by "/" in case of multiple charge states). "Holes" in this sequence coverage was most often explained by isotopic interferences that prevented clear ion identification, as exemplified for b) the w_{10}^{5-} which assignment would have been highly doubtful in case of unknown sequence despite matching quite well theoretical isotopic pattern shown on bottom, whereas c) safe identification of fragments at even higher charge state was achieved, as illustrated with the w_{16}^{6-} .

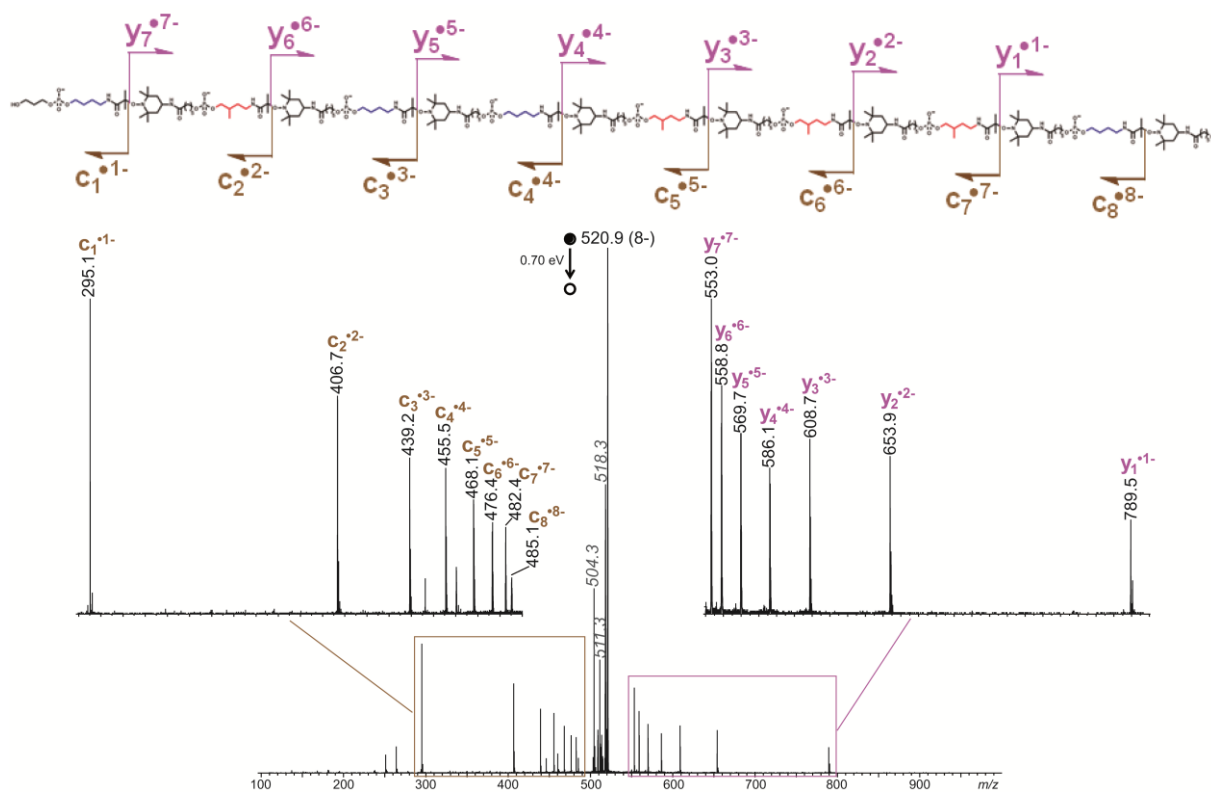


Figure 5. ESI-MS/MS spectrum of the fully deprotonated **01001110** poly(alkoxyamine phosphodiester) at m/z 520.9 ($z = 8$). Top: dissociation scheme of the polyanionic precursor, with c_i^{i-} and y_i^{i-} fragments designated in brown and in purple, respectively. In this structure, coding segments were highlighted in blue for **0** and in red for **1**. Peaks annotated with grey italicized m/z values correspond to internal fragments identified as either cyclic or biradical species such as $[0_i1_j - (i+j)H]^{(i+j)-}$. Collision energy indicated in the center-of-mass frame.

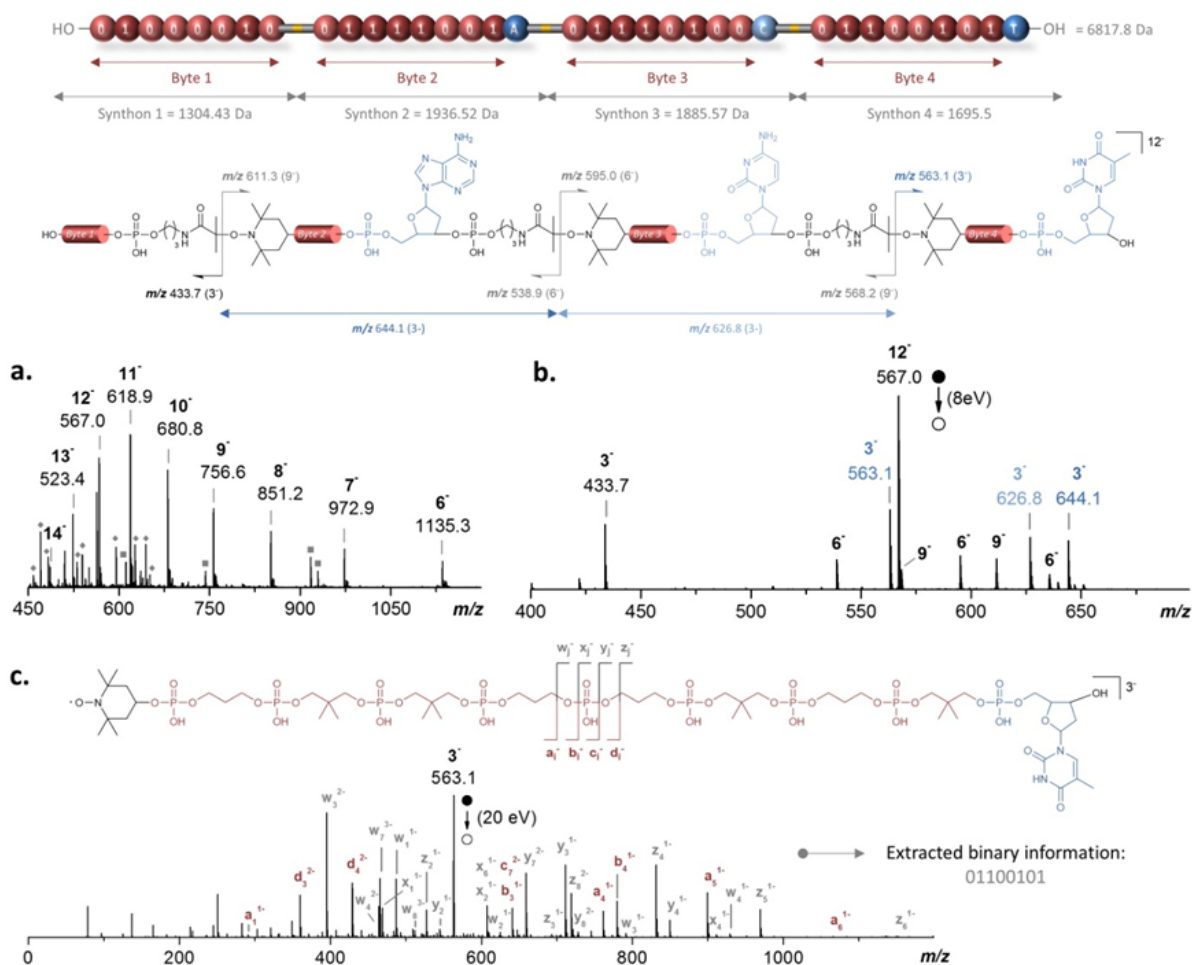
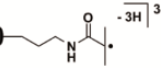


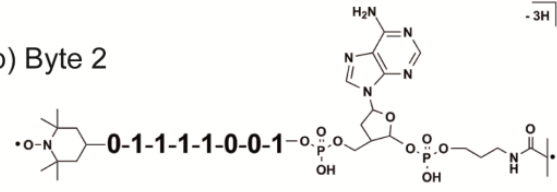
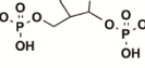
Figure 6. a) Negative mode ESI-MS of the byte-truncated poly(phosphodiester) containing the ASCII-encoded word “Byte”. b) CID spectrum of the m/z 567.0 ion, showing that activation of the $[M - 12H]^{12-}$ precursor mainly leads to homolysis of all alkoxyamine bonds placed between each byte (see top structure), and hence production of fragment of charge state proportional to the number of bytes they contain ($z=3$ for 1 byte, $z=6$ for 2 bytes, and $z=9$ for 3 bytes).

a) Byte 1

HO-0-1-0-0-0-0-1-0 

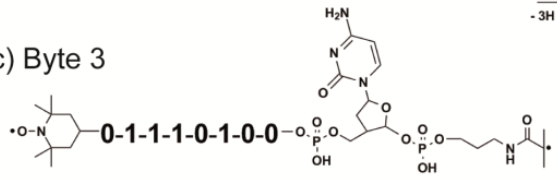
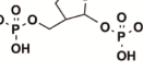
i↓	d _i ^{z-}	c _i ^{z-}	b _i ^{z-}	a _i ^{z-}	0	w _i ^{z-}	x _i ^{z-}	y _i ^{z-}	z _i ^{z-}	i↑
1	1	1			0					8
2	1	1	1		1				2	7
3	1/2	1	1	1	0	2/3				6
4	1/2	1	1	1	0		1/2	2	1/2	5
5	1/2	1/2	1/2	1	0		1	1	1/2	4
6	1/2	1/2	1/2	1/2	0	1	1	1	1	3
7	2/3	2	1/2		1	1		1		2
8	2/3	3	2	2	0				1	1

b) Byte 2

 0-1-1-1-1-0-0-1 

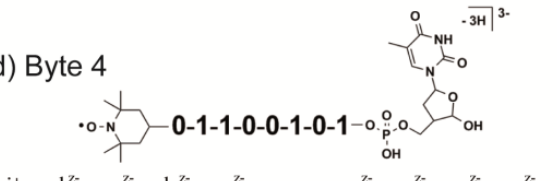
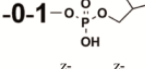
i↓	d _i ^{z-}	c _i ^{z-}	b _i ^{z-}	a _i ^{z-}	0	w _i ^{z-}	x _i ^{z-}	y _i ^{z-}	z _i ^{z-}	i↑
1	1	1	1	1	0	3		2	2	8
2		1	1	1	1	3		2		7
3	1/2	1	1	1	1	2		2	2	6
4	1/2	1	1	1	1			2	1	5
5	1/2	1/2	1/2	1	1	2			1	4
6	1/2	1/2	1/2	1/2	0	2			1	3
7	2/3	1/2	1/2	1/2	0	1/2		1		2
8	2/3	2/3	1/2	1/2	1	1/2		1	1	1

c) Byte 3

 0-1-1-1-0-1-0-0 

i↓	d _i ^{z-}	c _i ^{z-}	b _i ^{z-}	a _i ^{z-}	0	w _i ^{z-}	x _i ^{z-}	y _i ^{z-}	z _i ^{z-}	i↑
1	1	1	1	1	0	3	3		2	8
2	1	1	1	1	1	3	3	2	2	7
3	1/2	1	1		1	2/3				6
4	1/2	1/2	1	1	1					5
5	1/2	1/2	1/2	1/2	0	2				4
6	1/2/3	1	1/2	1/2	1	2			1	3
7	2/3	1/2/3	1/2	1/2	0				1	2
8	2/3	2/3	1/2	1/2	0					1

d) Byte 4

 0-1-1-0-0-1-0-1 

i↓	d _i ^{z-}	c _i ^{z-}	b _i ^{z-}	a _i ^{z-}	0	w _i ^{z-}	x _i ^{z-}	y _i ^{z-}	z _i ^{z-}	i↑
1	1	1	1	1	0	3	3	2	2/3	8
2	1	1	1	1	1	2/3	2/3	2	2/3	7
3	1/2	1	1	1	1	2/3	2	1/2	1	6
4	1/2	1/2	1	1	0	2	1/2	1/2	1/2	5
5	1/2	1/2	1/2	1/2	0	1/2	1/2	1	1/2	4
6	2/3	2	1/2	1/2	1	1/2	2	1	1	3
7	2/3	2	2	1/2	0	1/2	1	1	1	2
8		3	2	2	1	1	1	1	1	1

Figure 7. Sequence coverage charts obtained for a) byte 1 holding no tag, b) byte 2 tagged with A, c) byte 3 tagged with C, and d) byte 4 tagged with T, upon pseudo-MS³ of the byte-truncated poly(phosphodiester) that contains the ASCII-encoded word “Byte”. Top: structure of triply charged species activated in the second dissociation stage.