Promoting carboxylate salts in the ESI source to simplify positive mode MS/MS sequencing of acid-terminated encoded polyurethanes

Salomé Poyer, Benoit Eric Petit, Sofia Telitel, Denise Karamessini, Jean-François Lutz, Laurence Charles

Promoting carboxylate salts in the ESI source to simplify positive mode MS/MS sequencing of acid-terminated encoded polyurethanes

Salomé Poyer,[1] Benoit Eric Petit,[2] Sofia Telitel,[2] Denise Karamessini,[2] 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, Strasbourg, 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.

Published in *International Journal of Mass Spectrometry*, *448*, 116271 **(2020)**
https://doi.org/10.1016/j.ijms.2019.116271

Abstract

Synthetic polymers prepared by solid-phase processes are often released from the solid support by hydrolysis, leading to an acidic α chain-end which can usefully be employed to perform mass spectrometry experiments in the negative ion mode. This is the case for sequence-defined polyurethanes, which exhibit very simple MS/MS pattern as deprotonated species in great contrast to data obtained in the positive ion mode. Indeed, after deprotonation of their acidic end-group, collision induced dissociation (CID) of these polymers proceeds *via* competitive cleavages of all carbamate bonds. This leads to a unique series of anionic fragments spaced by the mass of one of the other coding co-monomer, enabling their sequence to be readily deciphered. However, as their size increases, polyurethanes are best ionized in the positive mode but their dissociation pattern is far more complicated, with four to five fragmentation routes depending on the adducted cation. This spectral complexity could however be highly reduced when selecting precursor ions that have experienced H/Na (or any other alkali) exchange in their acidic α end-group: in these conditions, only one reaction is observed and yields pairs of complementary products. Mechanisms could be proposed to explain how, although part of an end-group, the acidic proton had a key role in the dissociation processes of polyurethane chains. Exchange of this proton was then further optimized by proper selection of the X^- anion in the NaX salt supplemented to the electrosprayed solution. By allowing simplification of MS/MS data while avoiding signal dilution over multiple ion series, CID of these $[M - H + zNa]^{(z-1)+}$ permitted to envisage reliable decoding of sequence-defined polyurethanes storing large amount of information.

1. Introduction

Tandem mass spectrometry (MS/MS) has proved to be an unrivaled technique to decipher binary information encoded in the backbone of sequence-defined polymers [1]. Such macromolecules are typically composed of co-monomers of different mass arranged in a controlled manner throughout the chain [2-4]. For example, digital polymers can be prepared with two co-monomers arbitrarily defined as the **0**- and **1**-bit of the ASCII code. However, reliable decoding of such digital messages requires MS/MS data to allow full coverage of the binary sequence regardless of the chain length. This means that the same dissociation reactions should occur in all repeating units, independently of their coding moiety. As previously demonstrated in the field of DNA storage, long messages are preferentially composed of individual chains containing about 100 coded residues rather than written in very long chains [5]. To achieve *de novo* sequencing in this size range, the number of fragmentation pathways should be limited to avoid extensive dilution of total ion signal. Since the types of backbone bonds dictate the dissociation behavior of synthetic macromolecules [6], the structure of sequence-defined polymers has to be optimized to control their dissociation extent [7].

To prevent sequence-dependent reactions, chemical groups of low MS/MS reactivity have to be selected to implement the code, such as the simple H/CH_3 variation in alkyl segments used in most digital polymers developed in our group [8-10]. Beside the nature of coding moiety, two main strategies can be envisaged to design the polymeric skeleton. One can first focus on developing synthesis protocols for production of highly monodisperse samples of long polymeric chains, which molecular structure can be further optimized to ensure their MS/MS readability [11]. For example, this approach was successfully implemented for sequence-defined poly(phosphodiester)s. These species were first synthesized as defect-free chains with degrees of polymerization (DP) above 100 using an automated synthesis protocol to implement phosphoramidite chemistry [12]. However, while readily ionized in the negative mode electrospray ionization (ESI), collisional activation of poly(phosphodiester)s led to very complex MS/MS data due to efficient cleavage of all phosphate bonds. Abundance of the resulting eight fragment series, also dispersed in terms of charge state, was too weak to allow full sequence coverage of chains with DP above 50 [11]. Specific phosphoramidite reagents were then prepared to introduce, between each group of 8 coded bits (*i.e.*, one byte), an alkoxyamine linkage which cleavage required much lower energy compared to phosphate bonds. As a result, successful decoding of long byte-truncated poly(alkoxyamine

phosphodiester)s could be achieved by performing a first activation stage to specifically induce homolysis of these C–ON bonds, followed by a second activation stage to sequence so-formed (small) single byte containing fragments [13]. The alternative bottom-up strategy consists of (i) selecting candidates based on high MS/MS readability demonstrated for short oligomers and (ii) dedicating further efforts to maintain their simple dissociation pattern while increasing their chain length. This approach is explored here for sequenced-defined polyurethanes (PUs). When subjected to collision-induced dissociation (CID), PU oligomers with DP=4-12 exhibited very simple MS/MS pattern which revealed that only one reaction occurred in all repeating units, giving rise to a single fragment series due to a fixed negative charge in the oligomer α end-group [14]. As a result, coded sequences could be literally read from peak-to-peak distance in MS/MS spectra and a simple algorithm was created for automated extraction of digital data [15]. Alternatively, the code could be graphically read using Kendrick mass defect plots [16]. Although optimal for straightforward MS/MS decoding in the negative ion mode, sequence-defined PUs exhibited decreasing ionization yield as their chain length increased. This issue could be solved by performing positive mode ESI, but CID of cationized PUs led to more complex MS/MS data, as previously reported for related species [17-19]. Because the nature of activated precursors was previously reported to highly influence CID outputs for some synthetic polymers [20-22], different cationic adducts were investigated for their influence on the MS/MS behavior of PUs.

2. Experimental section

2.1. Chemicals

HPLC-grade solvents used to prepare sequence-defined PU samples (methanol, water, tetrahydrofuran) were purchased from Carlo Erba (Peypin, France). Salts used to promote formation of PU cationic adducts were ammonium acetate, sodium chloride, lithium fluoride, potassium iodide, rubidium bromide, cesium chloride, all obtained from Sigma (Saint Louis, MO) and used without further purification. Sequence-defined PUs used in this study (Supporting Information **Table S1**) were prepared according to a synthesis protocol published elsewhere [9]. Each PU sample (a few mg) was dissolved in THF and further diluted in methanolic solutions of different salts (see text).

2.2. Mass spectrometry

MS and MS/MS experiments were carried out a hybrid Q-TOF mass spectrometer Synapt G2-HDMS instrument (Waters, Manchester, UK). The ESI source was operated at 50°C, with

a desolvation gas temperature of 100°C, using the following settings: i) negative ion mode: capillary voltage, –2.3 kV; sampling cone, –35-120 V depending on PU chain length; extraction cone, –4.0 V; ii) positive ion mode: capillary voltage, +2.8 kV; sampling cone, +30-80 V depending on the species (see text); extraction cone, +4.0 V. All ions were accurately mass measured using the orthogonal acceleration (oa) time of flight (TOF) mass analyzer, with external calibration based on sodium acetate clusters in the MS mode or using the precursor ion as an internal standard in the MS/MS mode. MS/MS experiments were performed by subjecting selected precursor ions upon collisions with argon in a collision cell. Instrument control, data acquisition and data processing of all experiments were achieved with the MassLynx 4.1 programs provided by Waters. MS^n experiments were performed with a QTrap3200 mass spectrometer (Applied Biosystems Sciex, Concord, ON, Canada) equipped with an ESI source operated in positive ion mode, using a capillary voltage of +5.5 kV and a cone voltage of +75 V. Zero-grade air was used as the nebulizing gas (at a pressure of 10 psi) whereas nitrogen was used as the curtain gas (at a pressure of 20 psi) as well as the collision gas. In this instrument, precursor ions selected by a quadrupole and activated in a collision cell led to fragments that were further selectively sampled in a linear ion trap for a second activation stage. Instrument control, data acquisition and data processing of all experiments were achieved using Analyst software (version 1.4.1) provided by Applied Biosystems.

3. Results and discussion

3.1. CID of PUs: positive *vs* negative ion mode

Sequence-defined PUs were prepared using a chemoselective multistep-growth process implemented on a hydroxyfunctionalized crosslinked polystyrene resin, and their cleavage from the solid support led to an acid moiety at their α chain-end [**9**]. This enabled their deprotonation in the negative mode ESI, which happened to be extremely useful for their MS/MS sequencing. Indeed, as detailed in a previous study [**14**], deprotonated PUs experienced one bond cleavage per unit upon CID, leading to very simple MS/MS data composed of a single series of fragments spaced by the mass of one of the other co-monomers (115.0 Da for bit-0 *vs* 129.1 Da for bit-1). These a_i^- product ions were formed upon cleavage of the O–CO bond in each carbamate group and all contained the deprotonated α termination, as exemplified in Supporting Information **Fig. S1** with the 4-mer named P1. Accordingly, the **0001** sequence of P1 could be readily re-constructed from the peak-to-peak *m/z* distance equal to the mass of one or the other coding co-monomer. Although this simple MS/MS pattern was

kept unchanged for longer PUs, deprotonation yield strongly decreased as the chain length increased. For the P2 16-mer shown in Supporting Information **Fig. S2**, acquisition time had to be increased by a factor of 50 in the MS/MS mode to ensure detectability of all fragments. This was no longer compatible with high throughput reading of low amounts of coded polymers and did not offer interesting perspectives for sequencing of much longer chains.

In contrast, high ion current was recorded in the positive ion mode for PU oligomers. For example, a nearly 20-fold enhancement was measured for the doubly charged 16-mer P2 compared to the deprotonated molecule (Supporting Information **Fig. S3**). However, CID spectra recorded in the positive mode exhibited dissociation patterns far more complex than that observed for deprotonated PUs. As shown in **Fig. 1**(a), dissociation of protonated P1 was observed to proceed by cleavage of all bonds in the carbamate moieties, conducted to four product ion series (Supporting Information **Table S2)**. Restricting the nomenclature proposed by Wesdemiotis *et al.* [6] to the sole carbamate moiety, fragments containing the left-hand side group were named α_i^+ when formed after cleavage of the CH_2–O bond and a_i^+ when resulting from the cleavage of the O–(CO) bond, while x_j^+ and y_j^+ ions both containing the ω termination arose from cleavage at the O–(CO) and the (CO)–NH bonds, respectively (**Fig. 1**, top scheme). Although coverage of the **0001** sequence of P1 was achieved with each ion series observed in **Fig.1**(a), this made these MS/MS data highly redundant. Moreover, product ions were also dispersed in terms of charge states when formed from multiply protonated species generated upon ESI of long PUs. In addition, at the energy levels requested to produce all fragments for full sequence coverage of increasing length PUs, extensive secondary dissociation occurred and further increased CID data complexity (Supporting Information **Fig. S4** and **Table S3**). Alkali adducts could also be easily formed in the positive mode ESI since PUs contain carbonyl and ether oxygen atoms. However, their activation led to even more complex MS/MS spectra compared to protonated species, with an additional series of w_j^+ ions (containing the ω end-group and formed after cleavage of the CH_2–O bond) as illustrated in **Fig. 1**(b) for $[P1 + Na]^+$. Lithiated P1 exhibited the same dissociation behavior whereas CID of $[P1 + K]^+$, $[P1 + Rb]^+$ or $[P1 + Cs]^+$ conducted to the release of the naked alkali (Supporting Information **Fig. S5**), as commonly reported for polyethers where the strength of alkali adduction was shown to decrease for increasing size cations [20, 21]. In great contrast, highly simplified CID data were recorded for $[P1 – H + 2Na]^+$ ions, *i.e.*, sodium adducts of P1 in which H/Na exchange has occurred in the carboxylic acid group of the α termination (**Fig. 1**(c)). A major complete distribution of a_i^+ fragments was observed, together with a few (less abundant) members of the x_j^+ series. Surprisingly, the same MS/MS pattern was obtained for

precursors containing alternative alkali, regardless of their size (Supporting Information **Fig. S6**). These last results suggested that the adducted cation was more tightly bound to the alkali carboxylate form of P1 compared to its acidic form. Interestingly, the same simple dissociation behavior still applied as the chain size was increased (*vide infra*), permitting to anticipate high benefits for full sequence coverage of much larger PUs.

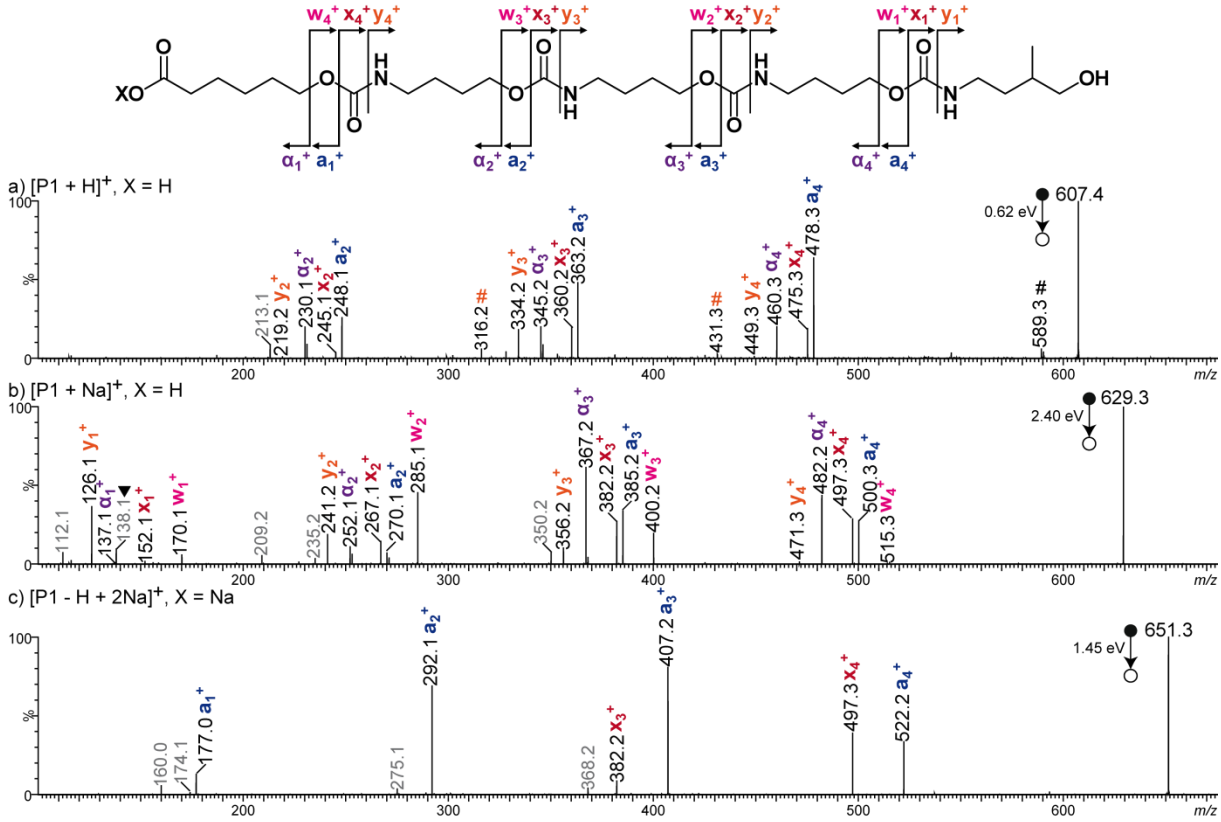


Fig. 1. ESI-MS/MS spectra of (a) $[P1 + H]^+$ at *m/z* 607.4 (X=H), (b) $[P1 + Na]^+$ at *m/z* 629.3 (X=H) and (c) $[P1 – H + 2Na]^+$ at *m/z* 651.3 (X=Na). Top: dissociation scheme with X = H or Na. Peaks annotated in grey correspond to secondary fragments. # designates dehydration products while the peak annotated with a black triangle in part b) correspond to a sodiated monomer (see text). For the sake of clarity, most minor peaks were not annotated but a complete list of fragments is reported in Supporting Information **Table S2** for $[P1 + H]^+$, **Table S4** for $[P1 + Na]^+$ and **Table 1** for $[P1 – H + 2Na]^+$. These experiments were performed with the Synapt G2-HDMS mass spectrometer. Activation energies are given in the center-of-mass frame.

	before HDX			*after HDX*		
	composition	m/z_{th}	m/z_{exp}	composition	m/z_{th}	m/z_{exp}
a_1^+	$C_6H_{11}O_3Na_2^+$	177.0498	177.0501	$C_6H_{10}DO_3Na_2^+$	178.0561	178.0566
a_2^+	$C_{11}H_{20}NO_5Na_2^+$	292.1131	292.1134	$C_{11}H_{18}D_2NO_5Na_2^+$	294.1257	294.1260
a_3^+	$C_{16}H_{29}N_2O_7Na_2^+$	407.1765	407.1758	$C_{16}H_{26}D_3N_2O_7Na_2^+$	410.1953	410.1953
a_4^+	$C_{21}H_{38}N_3O_9Na_2^+$	522.2398	522.2401	$C_{21}H_{34}D_4N_3O_9Na_2^+$	526.2649	526.2659
x_1^+	$C_6H_{11}NO_2Na^+$	152.0682	*n.d.*	$C_6H_{10}DNO_2Na^+$	153.0745	*n.d.*
x_2^+	$C_{11}H_{20}N_2O_4Na^+$	267.1315	267.1306	$C_{11}H_{18}D_2N_2O_4Na^+$	269.1441	*n.d.*
x_3^+	$C_{16}H_{29}N_3O_6Na^+$	382.1949	382.1944	$C_{16}H_{26}D_3N_3O_6Na^+$	385.2137	385.2139
x_4^+	$C_{21}H_{38}N_4O_8Na^+$	497.2582	497.2585	$C_{21}H_{34}D_4N_4O_8Na^+$	501.2833	501.2831

Table 1. Accurate mass measurement of fragments formed in CID of sodiated P1-Na (**Fig. 1**(c)) using the precursor ion ($C_{27}H_{49}N_4O_{11}Na_2^+$, *m/z* 651.3188 before H/D exchanges, left; $C_{27}H_{44}D_5N_4O_{11}Na_2^+$, *m/z* 656.3502 after H/D exchanges, right) as an internal standard. *n.d.*: not detected

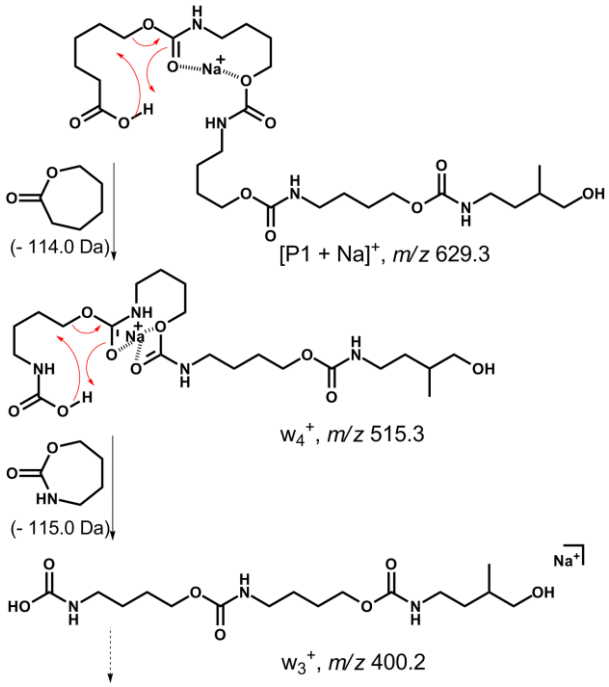
Overall, MS/MS data recorded for the salt form of P1 (**Fig. 1**(c)) strongly suggested that the acidic proton of the α end-group played a key role in the dissociation of $[P1+H]^+$ (**Fig. 1**(a)) and of $[P1+Na]^+$ (**Fig. 1**(b)). This is similar to reported works on unsaturated fatty acids, where formation of $[M – H + 2X]^+$ ions (with X = alkali) by Hsu and Turk [23] or of $[M – H + Ba]^+$ ions by Volmer et al. [24] was shown to inhibit fragmentation mechanisms driven by the carboxylate while promoting charge-remote fragmentation along the hydrocarbon backbone to yield useful information about double bond location. These reactions were further investigated in order to understand the influence of the carboxylate α termination on CID of PUs, using the example of the small P1 4-mer for the sake of clarity.

3.2. Mechanistic study of PU dissociation

In order to understand how the nature of the dissociating precursor influenced its dissociation pathways, complementary experiments were performed to rationalize MS/MS data of **Fig. 1**. Hydrogen/deuterium exchanges (HDX) were promoted by preparing PU solutions in deuterated methanol. In such conditions, P1 had six protons exchanged to deuterium, one in the α carboxylic acid function, one in the ω alcohol group and one in each carbamate moiety. CID spectra were hence recorded for $[P1\text{-}d6 + D]^+$ at *m/z* 614.4, $[P1\text{-}d6 + Na]^+$ at *m/z* 635.4 and $[P1\text{-}d6 – D + 2Na]^+$ at *m/z* 656.4. All CID fragments detected before or

after HDX were accurately mass measured. MS^3 experiments were also conducted to check whether product ions were primary or secondary fragments or to distinguish consecutive from competitive pathways.

Production of w_j^+ fragments was only observed upon activation of $[P1 + Na]^+$ (**Fig. 1**(b)), indicating that their formation required both the acidic proton of the α end-group and an adducted sodium. Moreover, MS^3 experiments showed that they were formed according to a consecutive process (Supporting Information **Fig. S7**). Accordingly, the mechanism depicted in **Scheme 1** was proposed, where solvation of the adducted sodium by oxygen atoms of the first units of P1 would favor the transfer of the α-end acidic proton onto the first carbonyl oxygen, leading to the formation of w_4^+ after elimination of the 114.0 Da caprolactone.

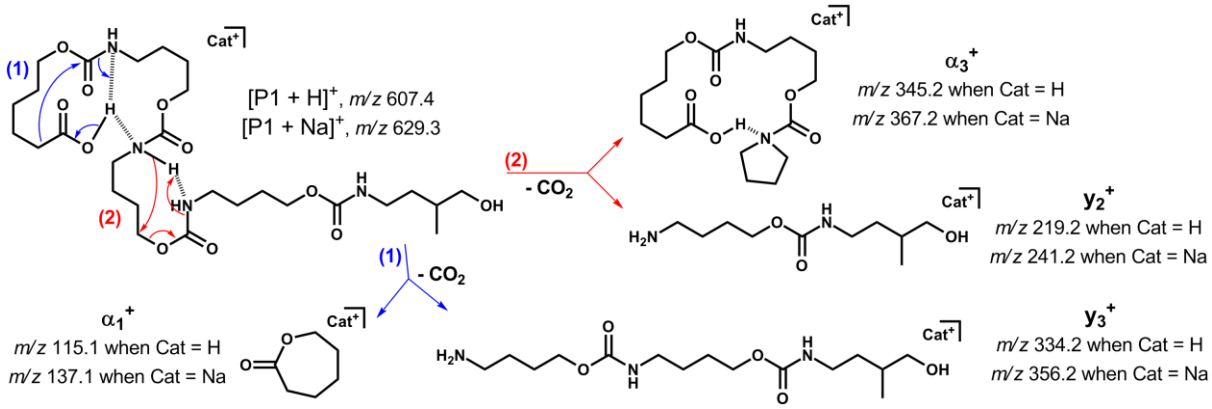


Scheme 1. Mechanism proposed to account for the formation of w_j^+ fragments upon CID of $[P1 + Na]^+$ at *m/z* 629.3.

Then, similar sodium solvation in w_4^+ would allow the first coding unit to be released as a cyclic 115.0 neutral to yield w_3^+. Iteration of the latter reaction would conduct to w_2^+ and w_1^+. Detection of the cyclic monomer at *m/z* 138.1 when adducted to sodium (designated by a black triangle in **Fig. 1**(b)) would further support such a multiple step process. Of note, the sodiated caprolactone expected as a complementary *m/z* 137.1 fragment in the first step of **Scheme 1** could not be distinguished from the α_1^+ fragment of the same $C_6H_{10}O_2Na^+$ composition. Mechanisms proposed in **Scheme 1** were also consistent with HDX

experiments, showing that each w_j^+ ion was formed upon D-transfer (Supporting Information **Table S4**).

The lack of α_i^+ and y_j^+ fragments in the CID spectrum of $[P1 - H + 2Na]^+$ (**Fig. 1**(c)) indicated that the α acidic proton also had a key role in the formation of these product ions. Indeed, when generated from $[P1 + Na]^+$ (**Fig. 1**(b)), y_j^+ ions were detected as sodiated species, strongly suggesting that the dissociating *m/z* 629.3 precursor was a sodium adduct of P1 rather that the protonated form of P1 having a sodium carboxylate as the α termination. HDX experiments indicated that α_i^+ species lacked one exchangeable proton while y_j^+ ions had gained one (Supporting Information **Table S2 and Table S4**). Accordingly, proposed mechanisms involved a structure with intra-molecular hydrogen bonds which would favor i) transfer of the α acidic proton to a carbamate NH group (**Scheme 2**, pathway 1 in blue) or ii) abstraction of the carbamate proton of one unit by the NH moiety of the next one (**Scheme 2**, pathway 2 in red). Both reactions would eliminate CO_2 and produce either α_i^+ or y_j^+ depending on the location of the adducted cation.

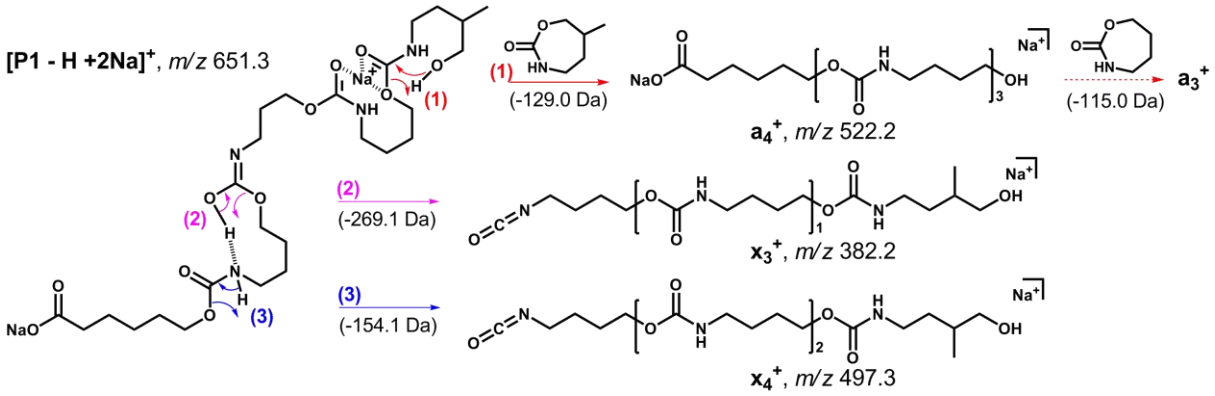


Scheme 2. Mechanism proposed to explain the formation of α_i^+ or y_j^+ fragments upon CID of precursor ions holding the acidic α termination such as $[P1 + H]^+$ at *m/z* 607.4 and $[P1 + Na]^+$ at *m/z* 629.3.

The same mechanism could apply when starting from either protonated or sodiated precursors. However, relative intensity of y_j^+ ions was very different in each case. When generated from $[P1 + H]^+$, y_j^+ ions exhibited a Gaussian-like distribution (**Fig. 1**(a)), and MS^3 experiments showed that they were formed according to a consecutive process (Supporting Information **Fig. S8**(a)). In contrast, when formed from $[P1 + Na]^+$, abundance of y_j^+ fragments strongly increased as their size decreased (**Fig. 1**(b)). As shown in MS^3, monomer elimination was indeed a minor dissociation pathway for y_j^+ formed from $[P1+Na]^+$

(Supporting Information **Fig. S8**(b)). Alternatively, these y_j^+ ions were readily generated as secondary fragments after loss of CO_2 from w_j^+ ions, as supported by MS^3 data (Supporting Information **Fig. S7**).

Finally, CID data of **Fig. 1** showed that a_i^+/x_j^+ product ions were detected regardless of the type of precursor ion. However, relative ion abundance measured in these two ion series slightly varied as a function of the dissociating species. As previously shown in the case of deprotonated PUs [14], multiple reactions could be envisaged for the formation of these fragments. Exploring all these processes is beyond the scope of this study which, instead, focused on the case of $[P1 - H + 2Na]^+$. Because a_i^+ and x_j^+ ions obtained upon cleavage of the same O–CO bond in this *m/z* 651.3 precursor were measured with very different abundances (see a_1^+ *vs* x_4^+, a_2^+ *vs* x_3^+, a_3^+ *vs* x_2^+, a_4^+ *vs* x_1^+ in **Fig. 1**(c)), they were not envisaged as complementary fragments formed during a charge-remote reaction but, instead, as ions produced *via* two distinct mechanisms. Since all a_i^+ ions contained two Na (**Table 1**), they were proposed to be formed upon successive elimination of cyclic monomers after transfer of the terminal hydroxy proton onto the ether oxygen (activated by the adducted sodium) of the preceding carbamate group (**Scheme 3**, pathway 1 in red).



Scheme 3. Proposed mechanisms for the dissociation of $[P1 – H + 2Na]^+$ species: pathway 1 (in red) could rationalize consecutive production of a_i^+ ions, whereas pathway 2 (in pink) and pathway 3 (in blue) would account for the competitive formation of x_j^+ ions.

In contrast, to account for the decreasing abundance of x_j^+ ions as their size decreased, competitive reactions involving a 1,3-proton transfer were considered, as proposed by Gies and Hercules for related species [19]. Consistent with HDX experiments (**Table 1**), a proton would be transferred either from the NH group onto the ether oxygen of the same carbamate moiety (**Scheme 3**, pathway 3 in blue). Alternatively, one could also consider transfer of the

hydroxyl proton in the enol form of a carbamate group, as illustrated by pathway 2 (in pink, **Scheme 3**). The latter reaction was also proposed since the enol form of carbamate moieties was evidenced in the formation of internal fragments at *m/z* 160.0, *m/z* 174.1 and *m/z* 275.1 (annotated in grey in **Fig. 1**(c)), as described in Supporting Information **Scheme S1**. None of the reactions depicted in **Scheme 3** produced the complementary fragments expected at *m/z* 152.1 (sodiated unit **1**) and *m/z* 138.1 (sodiated unit **0**) in pathway 1, *m/z* 292.1 in pathway 2 and *m/z* 177.1 in pathway 3. This result is consistent with the proposed strong solvation of the adducted cation by multiple oxygen atoms of P1 (**Scheme 3**), which would also prevent release of the naked alkali, regardless of its size (Supporting Information **Fig. S6**). In contrast, presence of the acidic α termination in $[P1 + Na]^+$ would favor the formation of a hydrogen bond network (as proposed in **Scheme 2**), preventing strong interaction with the adducted alkali and its facile release as its size increased, as observed when activating P1 cationized with K^+, Rb^+ or Cs^+ (Supporting Information **Fig. S5**).

3.3. Promoting formation of PU salts

Although $[P1 – H + 2Na]^+$ ions exhibited a highly promising CID behavior for MS/MS sequencing of larger PUs, it should be acknowledged that abundance of these species remained minor (about 30%) compared to $[P1 + Na]^+$, when employing a 1 mM NaCl solution. Due to the low volatility of sodium salts, the amount of NaCl could only be increased up to 3 mM but with no significant improvement of the absolute (or relative) intensity of the $[P1 – H + 2Na]^+$ signal. Optimizing the declustering potential (DP) that governed ion sampling from the ESI atmospheric pressure into the vacuum side of the mass spectrometer allowed an overall signal enhancement although it was not specific to the sodium salt form of PUs: as observed for the P3 8-mer of sequence **01000010**, increasing DP from 20 to 60 V allowed ion abundance to be roughly increased by a factor of about 3 for both $[P3 + 2Na]^{2+}$ at *m/z* 563.3 and $[P3 – H +3Na]^{2+}$ at *m/z* 574.3 (Supporting Information **Fig. S9**). Different NaX salts were then tested in order to optimize production of $[P – H +zNa]^{(z–1)+}$ species, using the P4 8-mer of sequence **01000100** as a model. The best salt to use was found to be NaOH because it is both basic and highly dissociated in polar solvents, and hence allowed the $[P4 – H + 3Na]^{2+}$ ion at *m/z* 574.3 to be obtained with maximal absolute abundance (**Fig. 2**(a)). Addition of 1 mM NaOH in the electrosprayed solution did not induce major loss of sensitivity, as shown in **Fig. 2**(b) by similar abundances measured for $[P4 + 3Na]^{2+}$ when using NaI and $[P4 – H + 3Na]^{2+}$ when using NaOH.

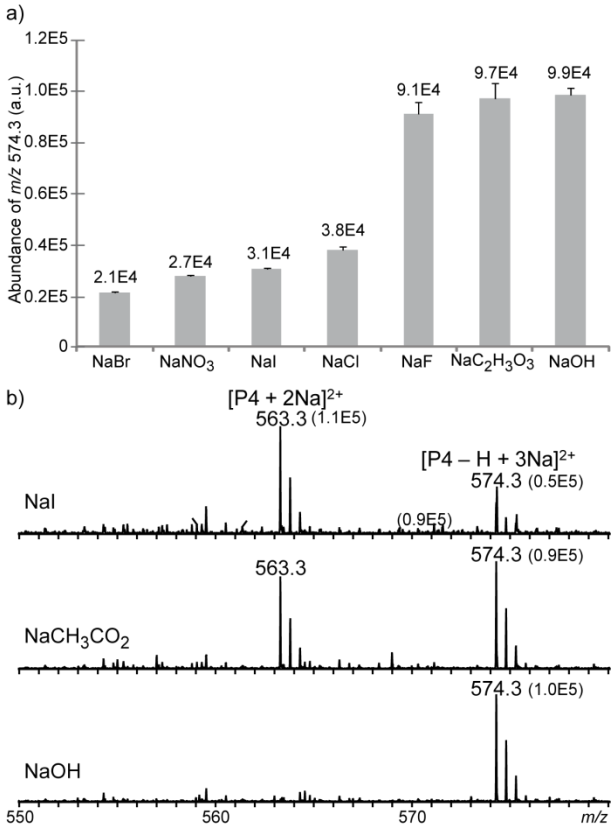


Fig. 2. (a) Absolute abundance of the $[P4 – H + 3Na]^{2+}$ ion at *m/z* 574.3 (grey bars, with average values on top) as a function of the NaX salt supplemented to the electrosprayed solution. Error bars as standard deviation of 3 replicates. (b) ESI(+) mass spectra (*m/z* 550-580) showing relative abundance of $[P4 + 2Na]^{2+}$ (*m/z* 563.3) and $[P4 – H + 3Na]^{2+}$ (*m/z* 574.3) when using NaI, $NaCH_3COO$ and NaOH as the cationization agent (each at 1 mM). Absolute abundance of these two species is indicated into parentheses. These experiments were performed with the Synapt G2-HDMS mass spectrometer.

3.4. Sequencing of long PU oligomers

These optimal conditions were then used for positive mode sequencing of the P2 16-mer. When prepared in a methanolic solution of NaOH (1 mM), this oligomer was readily electrosprayed as the doubly charged $[P2 – H + 3Na]^{2+}$ species and was observed to dissociate into the two main a_i^{z+} and x_j^{z+} ion series (**Fig. 3**(a) and Supporting Information **Table S5**). Using low activation energy level (*e.g.*, 40 eV laboratory frame for data shown in **Fig. 3**(a)) permitted to minimize secondary fragments (annotated in grey) while achieving full coverage of P2 sequence with the complete series of a_i^{z+} ions (annotated in blue), detected with z = 1 from i = 1-12 and z = 2 from i = 10-16. The **1101001011011000** code written in the backbone of P2 was best read after charge state deconvolution of these MS/MS data (**Fig. 3**(b)). Starting

from a_1 always expected with a 154.0 Da mass since it does not contain any coding moiety (**Fig. 1**, top scheme), the sequencing task simply consisted of searching for the next a_{i+1} signal by adding either 115.1 Da (mass of the **0**-bit) or 129.1 Da (mass of the **1**-bit). As a result, following blue arrows in **Fig. 3**(b) allowed the binary message to be readily deciphered. The partial sequence reconstructed when using x_j species had, instead, to be read from the right- to the left-hand side (as indicated by red arrows below the x-axis of **Fig. 3**(b)) and was used as a validation tool.

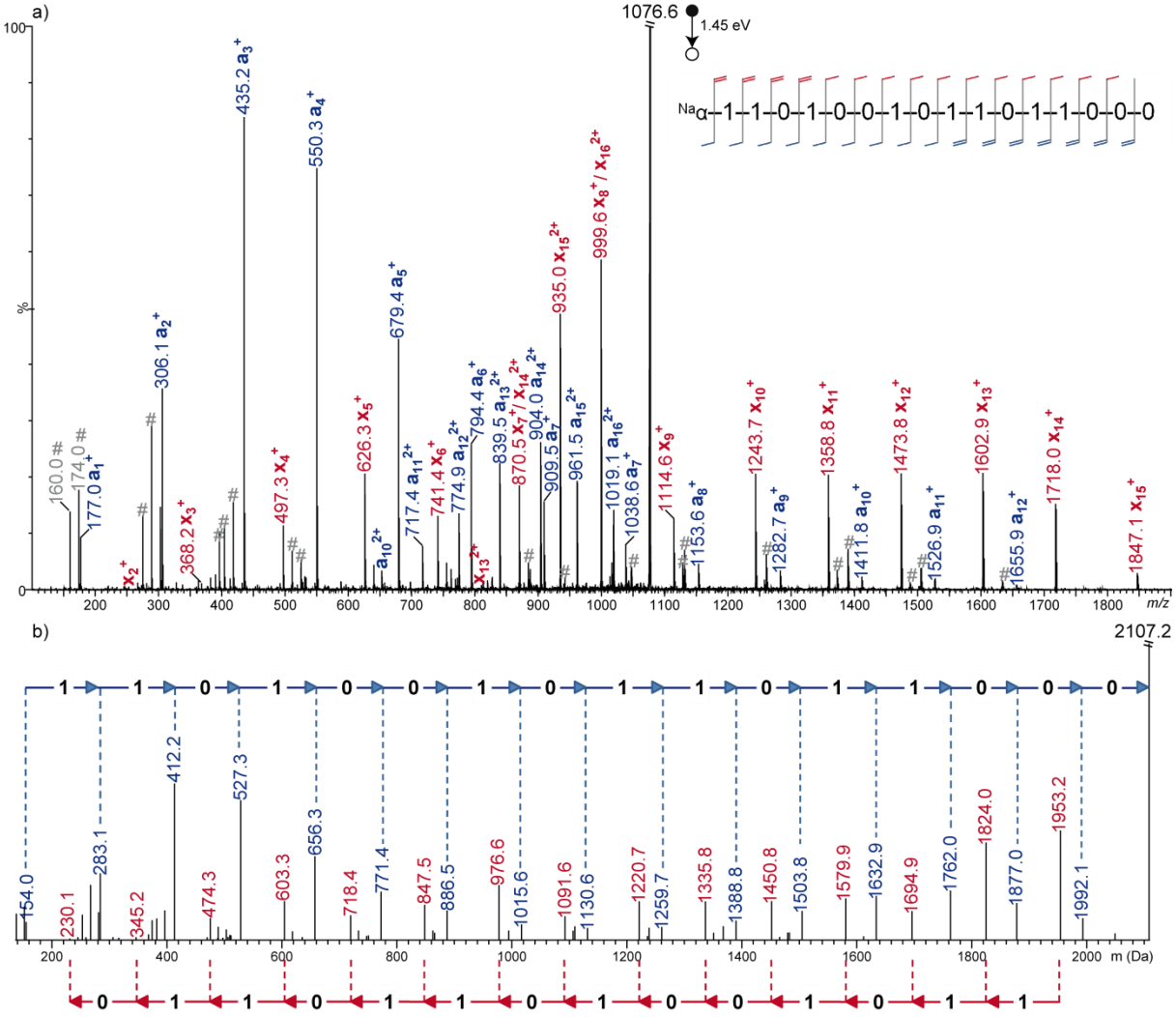


Fig. 3. (a) ESI(+)-MS/MS of $[P2 - H + 3Na]^{2+}$ at *m/z* 1076.6, using a 1.45 eV activation energy (center-of-mass frame). In the associated dissociation scheme (inset), singly and doubled charged fragments are respectively designated with single and double pattern, in blue for a_i^{z+} and in red for x_j^{z+}. Grey # symbols designate secondary fragments, including the two internal *m/z* 160.0 and *m/z* 174.0 ions. (b) Charge-deconvoluted MS/MS spectrum with masses of a_i and x_j species annotated in blue and red, respectively. Unlabelled minor peaks

correspond to secondary fragments. In both spectra, the y-axis has been expanded by a factor of 3 for the sake of clarity. These experiments were performed with the Synapt G2-HDMS mass spectrometer.

4. Conclusion

The dissociation pattern of sequence-defined PUs in the positive ion mode could be strongly modified when changing their acidic α termination into a carboxylate moiety. This H/Na (or other alkali) exchange was easily performed in solution and permitted to efficiently suppress all reactions involving the carboxylic acid α group of these polymers. Collisional activation of carboxylate-terminated PUs mainly induced cleavage of the O–(CO) bond of carbamate group present in each repeating unit, which led to the formation of two fragment series only and hence fulfilled conditions requested for full sequence coverage by MS/MS. This is another example of the "MS/MS design" approach which aims at optimizing the structure of encoded polymers to control their dissociation behavior and hence improve their sequencing. However, in contrast to previously reported cases where synthesis works were requested, changing a carboxylic acid into its carboxylate form was readily achieved in the electrosprayed solution supplemented with the appropriate alkali salts. Thank to multiple charging during ESI, PUs of increasing length were efficiently ionized in the positive ion mode and exhibited the same simple MS/MS pattern regardless of their charge state. With this new methodology in hand, production of sequence-defined polymers to store multiple bytes of information can now be envisaged.

Acknowledgments

L.C. and J.-F.L. thank the French National Research Agency (ANR project 00111001, grant numbers ANR-16-CE29-0004-01 and ANR-16-CE29-0004-02) for financial support. L.C. also acknowledges support from Spectropole, the Analytical Facility of Aix-Marseille University, by granting a special access to the instruments purchased with European Funding (FEDER OBJ2142-3341).

References

[1] L. Charles. Tandem mass spectrometry sequencing of sequence-controlled and sequence-defined synthetic polymers. In: Sequence controlled polymers, J.-F. Lutz, Ed. Wiley: 2018; pp 479-504.

[2] H. Colquhoun, J.-F. Lutz. Information-containing macromolecules. Nat. Chem. 6 (2014) 455-456.

[3] J.-F. Lutz. Coding Macromolecules: Inputting Information in Polymers Using Monomer-Based Alphabets. Macromolecules 48 (2015) 4759-4767.

[4] M. Rutten, R.W. Vaandrager, J. Elemans, R.J.M. Nolte. Encoding information into polymers. Nat. Rev. Chem. 2 (2018) 365-381.

[5] G.M. Church, Y. Gao, S. Kosuri. Next-generation digital information storage in DNA. Science 337 (2012) 1628-1628.

[6] C. Wesdemiotis, N. Solak, M.J. Polce, D.E. Dabney, K. Chaicharoen, B.C. Katzenmeyer, Fragmentation pathways of polymer ions. Mass Spectrom. Rev. 30 (2011) 523-559.

[7] 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. 28 (2017) 1149-1159.

[8] 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. 6 (2015) 7237.

[9] 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 1 (2016) 114-126.

[10] G. 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. 138 (2016) 9417-9420.

[11] J.-A. Amalian, A. Al Ouahabi, G. Cavallo, N.F. Konig, S. Poyer, J.-F. Lutz, L. Charles. Controlling the structure of sequence-defined poly(phosphodiester)s for optimal MS/MS reading of digital information. J. Mass Spectrom. 52 (2017) 788-798.

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

[13] 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. 8 (2017) 967.

[14] J.-A. Amalian, S. Poyer, B.E. Petit, S. Telitel, V. Monnier, D. Karamessini, D. Gigmes, J.-F. Lutz, L. Charles. Negative mode MS/MS to read digital information encoded in sequence-defined oligo(urethane)s: A mechanistic study. Int. J. Mass Spectrom. 421 (2017) 271-278.

[15] A. Burel, C. Carapito, J.-F. Lutz, L. Charles. MS-DECODER: Milliseconds Sequencing of Coded Polymers. Macromolecules 50 (2017) 8290-8296.

[16] S. Poyer, T. Fouquet, H. Sato, J.-F. Lutz, L. Charles. Convenient graphical visualization of messages encoded in sequence-controlled synthetic polymers using Kendrick mass defect analysis of their MS/MS data. Macromol. Chem. Phys. 219 (2018) 1800173.

[17] V. Mass, W. Schrepp, B. Von Vacano, H. Pasch. Sequence analysis of an isocyanate oligomer by MALDI-TOF mass spectrometry using collision induced dissociation. Macromol. Chem. Phys. 210 (2009) 1957-1965.

[18] A.P. Gies, W.H. Heath, R.J. Keaton, J.J. Jimenez, J.J. Zupanic. MALDI-TOF/TOF CID study of polycarbodiimide branching reactions. Macromolecules 46 (2013) 7616-7637.

[19] A.P. Gies, D.M. Hercules. Collision induced dissociation study of ester-based polyurethane fragmentation reactions. Anal. Chim. Acta 808 (2014) 199-219.

[20] R.P. Lattimer. Tandem mass spectrometry of lithium-attachment ions from polyglycols. J. Am. Soc. Mass Spectrom. 3 (1992) 225-234.

[21] R.P. Lattimer. Tandem mass spectrometry of poly(ethylene glycol) lithium-attachment ions. J. Am. Soc. Mass Spectrom. 5 (1994) 1072-1080.

[22] A. Tintaru, C. Chendo, Q. Wang, S. Viel, G. Quelever, L. Peng, P. Posocco, S. Pricl, L. Charles. Conformational sensitivity of conjugated poly(ethylene oxide)-poly(amidoamine) molecules to cations adducted upon electropsray ionization. A mass spectrometry, ion mobility and molecular modelling study. Anal. Chim. Acta 808 (2014) 163-174.

[23] F.F. Hsu, J. Turk. Distinction among isomeric unsaturated fatty acids as lithiated adducts by electrospray ionization mass spectrometry using low energy collisionally activated dissociation on a triple stage quadrupole instrument. J. Am. Soc. Mass Spectrom. 10 (1999) 600-612.

[24] N. Zehethofer, D.M. Pinto, D.A. Volmer. Plasma free fatty acid profiling in a fish oil human intervention study using ultra-performance liquid chromatography/electrospray

ionization tandem mass spectrometry. Rapid Commun. Mass Spectrom. 22 (2008) 2125-2133.

Supporting information

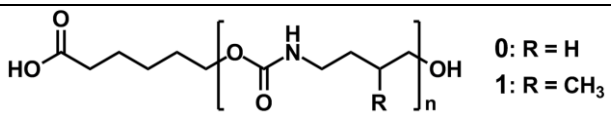
Promoting carboxylate salts in the ESI source to simplify positive mode MS/MS sequencing of acid-terminated encoded polyurethanes

Salomé Poyer,[1] Benoit Eric Petit,[2] Sofia Telitel,[2] Denise Karamessini,[2] 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, Strasbourg, France

Content



	n	sequence	composition	mass (Da)
P1	4	α-0001-ω	$C_{27}H_{50}N_4O_{11}$	606.3476
P2	16	α-1101001011011000-ω	$C_{94}H_{172}N_{16}O_{35}$	2085.2171
P3	8	α-01000010-ω	$C_{48}H_{88}N_8O_{19}$	1080.6166
P4	8	α-01000100-ω	$C_{48}H_{88}N_8O_{19}$	1080.6166

Table S1. Structural description of PU oligomers used in this study.

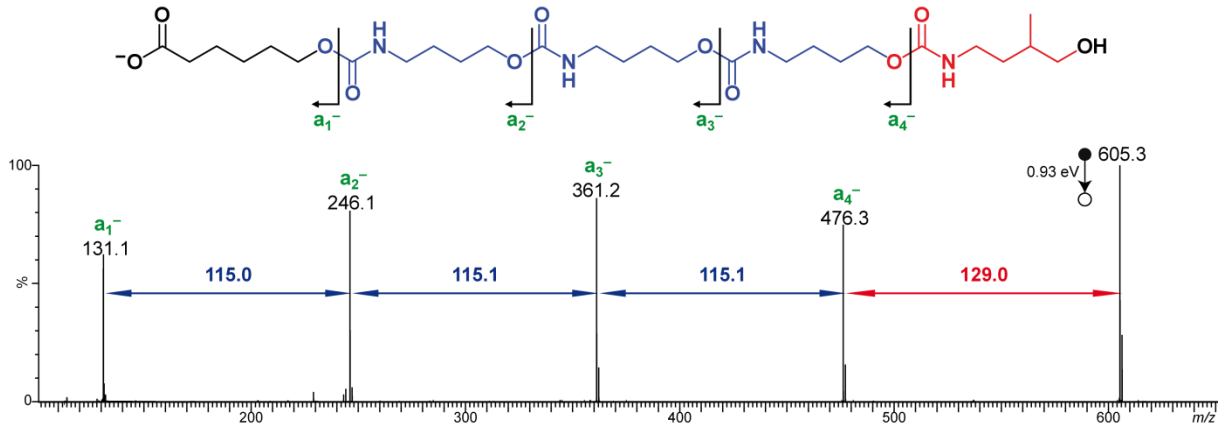


Fig. S1. ESI(–)-MS/MS of the deprotonated P1 oligomer of sequence **0001** (top structure) detected at *m/z* 605.3. Activation energy is given in the center-of-mass frame. Starting from the a_1^- ion always detected at *m/z* 131.1 since this ion does not contain any coding moiety, the **0001** sequence of P1 is re-constructed from the peak-to-peak distance equal to the mass of the 0-bit (115.0 Da) or of the 1-bit (129.0 Da): Δ*m/z* (a_2^-/a_1^-) = 115.0, Δ*m/z* (a_3^-/a_2^-) = 115.1, Δ*m/z* (a_4^-/a_3^-) = 115.1 and Δ*m/z* ([PU-1 – H]$^-/a_4^-$) = 129.0. This experiment was performed with the Synapt G2-HDMS mass spectrometer.

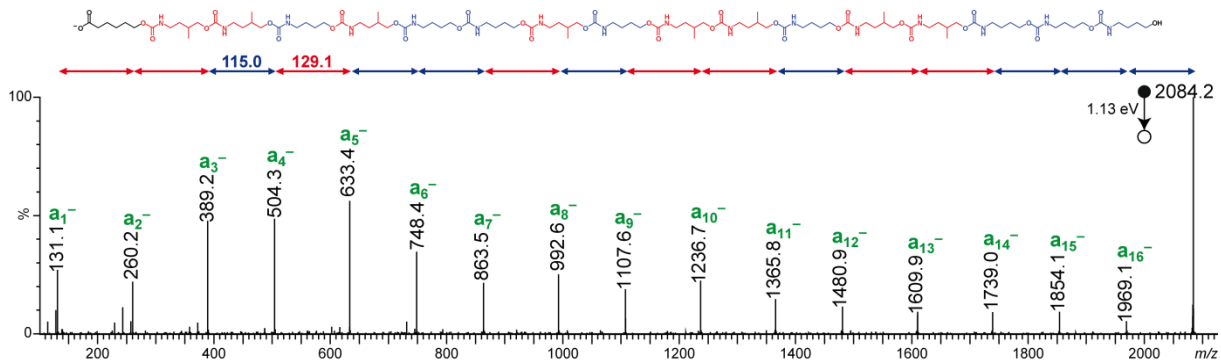


Fig. S2. ESI(–)-MS/MS of the deprotonated P2 oligomer of sequence **1101001011011000** (top structure) detected at *m/z* 2084.2. Activation energy is given in the center-of-mass frame. This experiment was performed with the Synapt G2-HDMS mass spectrometer.

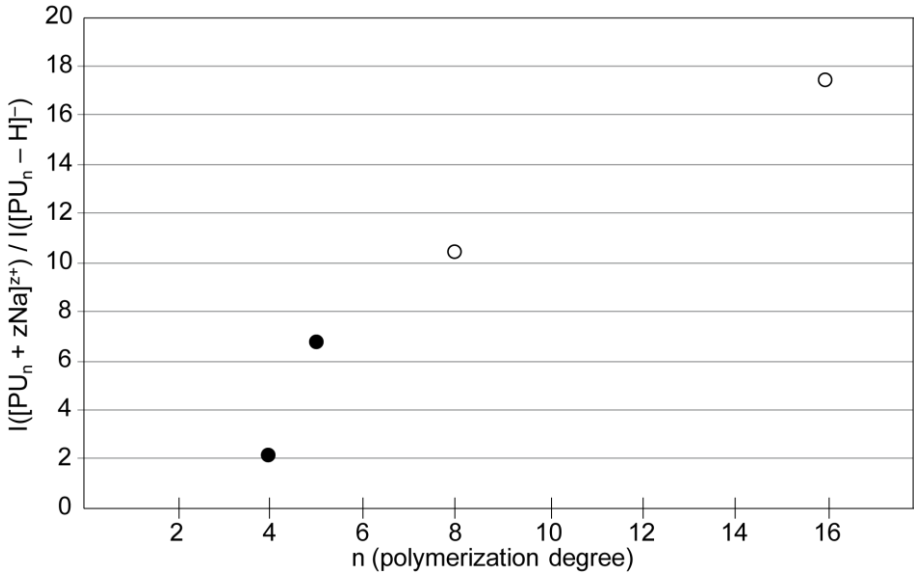


Fig. S3. Relative intensity of some PUs when adducted to one (black dots) or two (white dots) sodium cations compared to deprotonated molecules obtained in the negative ion mode, as a function of their polymerization degree, n. These experiments were performed with the Synapt G2-HDMS mass spectrometer.

		before HDX			*after HDX*		
		composition	m/z_{th}	m/z_{exp}	composition	m/z_{th}	m/z_{exp}
α-containing fragments	a_1^+	$C_6H_{13}O_3^+$	133.0859	133.0860	$C_6H_{10}D_3O_3^+$	136.1048	136.1051
	a_2^+	$C_{11}H_{22}NO_5^+$	248.1492	248.1489	$C_{11}H_{18}D_4NO_5^+$	252.1744	252.1749
	a_3^+	$C_{16}H_{31}N_2O_7^+$	363.2126	363.2124	$C_{16}H_{26}D_5N_2O_7^+$	368.2440	368.2439
	a_4^+	$C_{21}H_{40}N_3O_9^+$	478.2759	478.2756	$C_{21}H_{34}D_6N_3O_9^+$	484.3136	484.3138
	α_1^+	$C_6H_{11}O_2^+$	115.0754	115.0755	$C_6H_{10}D_1O_2^+$	116.0816	116.0827
	α_2^+	$C_{11}H_{20}NO_4^+$	230.1387	230.1386	$C_{11}H_{18}D_2NO_4^+$	232.1512	232.1515
	α_3^+	$C_{16}H_{29}N_2O_6^+$	345.2020	345.2020	$C_{16}H_{26}D_3N_2O_6^+$	348.2208	348.2204
	α_4^+	$C_{21}H_{38}N_3O_8^+$	460.2653	460.2651	$C_{21}H_{34}D_4N_3O_8^+$	464.2904	464.2906
ω-containing fragments	x_1^+	$C_6H_{12}NO_2^+$	130.0863	130.0852	$C_6H_{10}D_2NO_2^+$	132.0988	132.1008
	x_2^+	$C_{11}H_{21}N_2O_4^+$	245.1496	245.1495	$C_{11}H_{18}D_3N_2O_4^+$	248.1684	248.1691
	x_3^+	$C_{16}H_{30}N_3O_6^+$	360.2129	360.2123	$C_{16}H_{26}D_4N_3O_6^+$	364.2380	364.2381
	x_4^+	$C_{21}H_{39}N_4O_8^+$	475.2762	475.2762	$C_{21}H_{34}D_5N_4O_8^+$	480.3076	480.3080
	y_1^+	$C_5H_{14}NO^+$	104.1070	104.1068	$C_5H_{10}D_4NO^+$	108.1321	108.1315
	y_2^+	$C_{10}H_{23}N_2O_3^+$	219.1703	219.1700	$C_{10}H_{18}D_5N_2O_3^+$	224.2017	224.2014
	y_3^+	$C_{15}H_{32}N_3O_5^+$	334.2336	334.2331	$C_{15}H_{26}D_6N_3O_5^+$	340.2713	340.2716
	y_4^+	$C_{20}H_{41}N_4O_7^+$	449.2970	449.2966	$C_{20}H_{34}D_7N_4O_7^+$	456.3409	456.3432

Table S2. Accurate mass measurement of fragments formed in CID of protonated P1 using the precursor ion ($C_{27}H_{51}N_4O_{11}^+$, *m/z* 607.3549 before H/D exchanges, left; $C_{27}H_{44}D_7N_4O_{11}^+$, *m/z* 614.3988 after H/D exchanges, right) as an internal standard.

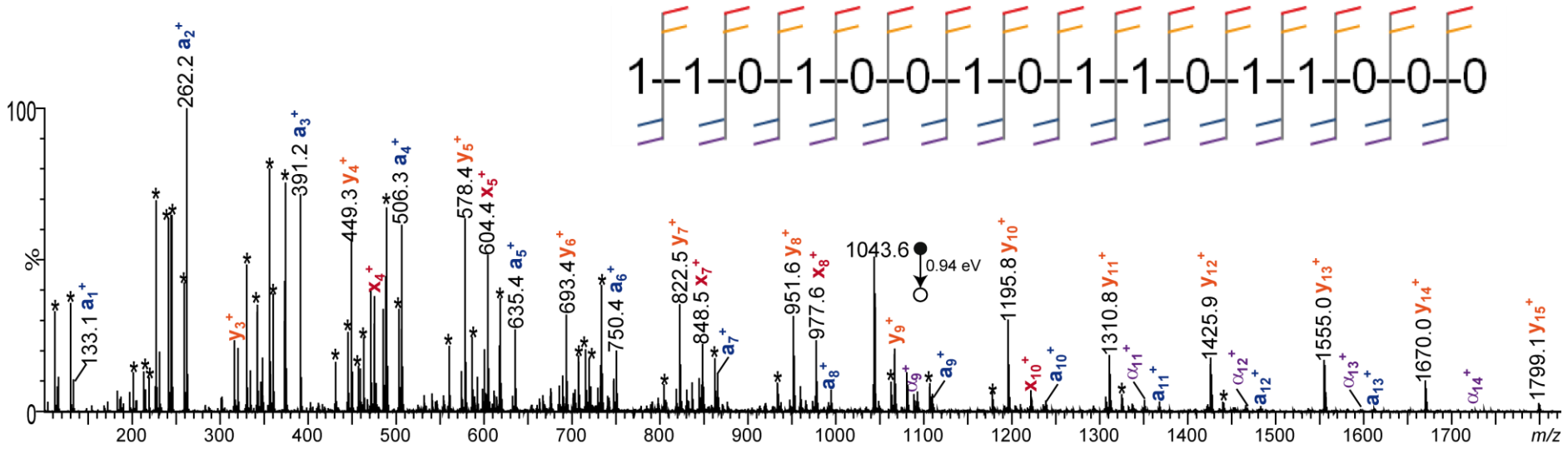


Fig. S4. ESI(+)-MS/MS of the doubly protonated P2 detected at *m/z* 1043.6, allowing full coverage of the **1101001011011000** sequence, as depicted by the inset scheme. However, the sequencing task was made difficult by numerous secondary fragments (designated by asterisks) of particularly high abundance in the low *m/z* range of the MS/MS spectrum. All peaks were not annotated for the sake of clarity but fragments useful for sequencing are all reported in Table S3. Activation energy is given in the center-of-mass frame. This experiment was performed with the Synapt G2-HDMS mass spectrometer.

	α-containing product ions			**ω-containing product ions**		
	a_i^{z+} ions			x_i^{z+} ions		
i	composition	m/z_{th}	m/z_{exp}	composition	m/z_{th}	m/z_{exp}
1	$C_6H_{13}O_3^+$	133.0859	133.0850	$C_5H_{10}NO_2^+$	116.0706	116.0710
2	$C_{12}H_{24}NO_5^+$	262.1649	262.1658	$C_{10}H_{19}N_2O_4^+$	231.1339	231.1334
3	$C_{18}H_{35}N_2O_7^+$	391.2439	391.2432	$C_{15}H_{28}N_3O_6^+$	346.1973	346.1968
4	$C_{23}H_{44}N_3O_9^+$	506.3072	506.3080	$C_{21}H_{39}N_4O_8^+$	475.2762	475.2773
5	$C_{29}H_{55}N_4O_{11}^+$	635.3862	635.3874	$C_{27}H_{50}N_5O_{10}^+$	604.3552	604.3557
6	$C_{34}H_{64}N_5O_{13}^+$	750.4495	750.4488	$C_{32}H_{59}N_6O_{12}^+$	719.4185	719.4221
7	$C_{39}H_{73}N_6O_{15}^+$	865.5128	865.5103	$C_{38}H_{70}N_7O_{14}^+$	848.4975	848.4976
8	$C_{45}H_{84}N_7O_{17}^+$	994.5918	994.5902	$C_{44}H_{81}N_8O_{16}^+$	977.5765	977.5791
9	$C_{50}H_{93}N_8O_{19}^+$	1109.6552	1109.6582	$C_{49}H_{90}N_9O_{18}^+$	1092.6398	1092.6423
	$C_{50}H_{94}N_8O_{19}^{2+}$	555.3312	555.3334	$C_{49}H_{91}N_9O_{18}^{2+}$	546.8236	546.8214
10	$C_{56}H_{104}N_9O_{21}^+$	1238.7341	1238.7306	$C_{55}H_{101}N_{10}O_{20}^+$	1221.7188	1221.7178
	$C_{56}H_{105}N_9O_{21}^{2+}$	619.8707	619.8729	$C_{55}H_{102}N_{10}O_{20}^{2+}$	611.3630	611.3624
11	$C_{62}H_{115}N_{10}O_{23}^+$	1367.8131	1367.8116	$C_{60}H_{110}N_{11}O_{22}^+$	1336.7821	1336.7800
	$C_{62}H_{116}N_{10}O_{23}^{2+}$	684.4102	684.4101	$C_{60}H_{111}N_{11}O_{22}^{2+}$	668.8947	668.8962
12	$C_{67}H_{124}N_{11}O_{25}^+$	1482.8764	1482.8787	$C_{65}H_{119}N_{12}O_{24}^+$	1451.8455	1451.8480
	$C_{67}H_{125}N_{11}O_{25}^{2+}$	741.9419	741.9420	$C_{65}H_{120}N_{12}O_{24}^{2+}$	726.4264	726.4264
13	$C_{73}H_{135}N_{12}O_{27}^+$	1611.9554	1611.9526	$C_{71}H_{130}N_{13}O_{26}^+$	1580.9244	1580.9332
	$C_{73}H_{136}N_{12}O_{27}^{2+}$	806.4813	806.4844	$C_{71}H_{131}N_{13}O_{26}^{2+}$	790.9659	790.9698
14	$C_{79}H_{146}N_{13}O_{29}^+$	1741.0344	1741.0328	$C_{76}H_{139}N_{14}O_{28}^+$	1695.9878	1695.9958
	$C_{79}H_{147}N_{13}O_{29}^{2+}$	871.0208	871.0197	$C_{76}H_{140}N_{14}O_{28}^{2+}$	848.4975	848.4976
15	$C_{84}H_{156}N_{14}O_{31}^{2+}$	928.5525	928.5551	$C_{82}H_{151}N_{15}O_{30}^{2+}$	913.0370	913.0364
16	$C_{89}H_{165}N_{15}O_{33}^{2+}$	986.0842	986.0862	$C_{88}H_{162}N_{16}O_{32}^{2+}$	977.5765	977.5791
	α_i^{z+} ions			y_i^{z+} ions		
1	$C_6H_{11}O_2^+$	115.0754	115.0763	$C_4H_{12}NO^+$	90.0913	90.0913
2	$C_{12}H_{22}NO_4^+$	244.1543	244.1539	$C_9H_{21}N_2O_3^+$	205.1547	205.1541
3	$C_{18}H_{33}N_2O_6^+$	373.2333	373.2327	$C_{14}H_{30}N_3O_5^+$	320.218	320.2183
4	$C_{23}H_{42}N_3O_8^+$	488.2966	488.2977	$C_{20}H_{41}N_4O_7^+$	449.297	449.2979
5	$C_{29}H_{53}N_4O_{10}^+$	617.3756	617.3787	$C_{26}H_{52}N_5O_9^+$	578.376	578.3771
6	$C_{34}H_{62}N_5O_{12}^+$	732.4389	732.4427	$C_{31}H_{61}N_6O_{11}^+$	693.4393	693.4413
7	$C_{39}H_{71}N_6O_{14}^+$	847.5023	847.5001	$C_{37}H_{72}N_7O_{13}^+$	822.5183	822.5180
8	$C_{45}H_{82}N_7O_{16}^+$	976.5813	976.5745	$C_{43}H_{83}N_8O_{15}^+$	951.5972	951.5962
9	$C_{50}H_{91}N_8O_{18}^+$	1091.6446	1091.6497	$C_{48}H_{92}N_9O_{17}^+$	1066.6606	1066.6587
	$C_{50}H_{92}N_8O_{18}^{2+}$	546.3259	546.3278	$C_{48}H_{93}N_9O_{17}^{2+}$	533.8339	533.8307
10	$C_{56}H_{102}N_9O_{20}^+$	1220.7236	1220.7241	$C_{54}H_{103}N_{10}O_{19}^+$	1195.7395	1195.7408
	$C_{56}H_{103}N_9O_{20}^{2+}$	610.8654	610.8681	$C_{54}H_{104}N_{10}O_{19}^{2+}$	598.3734	598.3763
11	$C_{62}H_{113}N_{10}O_{22}^+$	1349.8025	1349.8071	$C_{59}H_{112}N_{11}O_{21}^+$	1310.8029	1310.8035
	$C_{62}H_{114}N_{10}O_{22}^{2+}$	675.4049	675.4178	$C_{59}H_{113}N_{11}O_{21}^{2+}$	655.9051	655.8975
12	$C_{67}H_{122}N_{11}O_{24}^+$	1464.8659	1464.8668	$C_{64}H_{121}N_{12}O_{23}^+$	1425.8662	1425.8635
	$C_{67}H_{123}N_{11}O_{24}^{2+}$	732.9366	732.9335	$C_{64}H_{122}N_{12}O_{23}^{2+}$	713.4367	713.4415
13	$C_{73}H_{133}N_{12}O_{26}^+$	1593.9449	1593.9575	$C_{70}H_{132}N_{13}O_{25}^+$	1554.9452	1554.9419
	$C_{73}H_{134}N_{12}O_{26}^{2+}$	797.4761	797.4739	$C_{70}H_{133}N_{13}O_{25}^{2+}$	777.9762	777.9720
14	$C_{79}H_{144}N_{13}O_{28}^+$	1723.0238	1723.0527	$C_{75}H_{141}N_{14}O_{27}^+$	1670.0085	1670.0164
	$C_{79}H_{145}N_{13}O_{28}^{2+}$	862.0156	862.0173	$C_{75}H_{142}N_{14}O_{27}^{2+}$	835.5079	835.5067
15	$C_{84}H_{154}N_{14}O_{30}^{2+}$	919.5472	919.5516	$C_{81}H_{153}N_{15}O_{29}^{2+}$	900.0474	900.0472
16	$C_{89}H_{163}N_{15}O_{32}^{2+}$	977.0789	977.0735	$C_{87}H_{164}N_{16}O_{31}^{2+}$	964.5869	964.5881

Table S3. Accurate mass measurement of fragments formed in CID of $[P2 + 2H]^{2+}$ (Figure S4) and used to recover the **1101001011011000** sequence, using the precursor ion ($C_{94}H_{174}N_{16}O_{35}^{2+}$, *m/z* 1043.6158) as an internal standard.

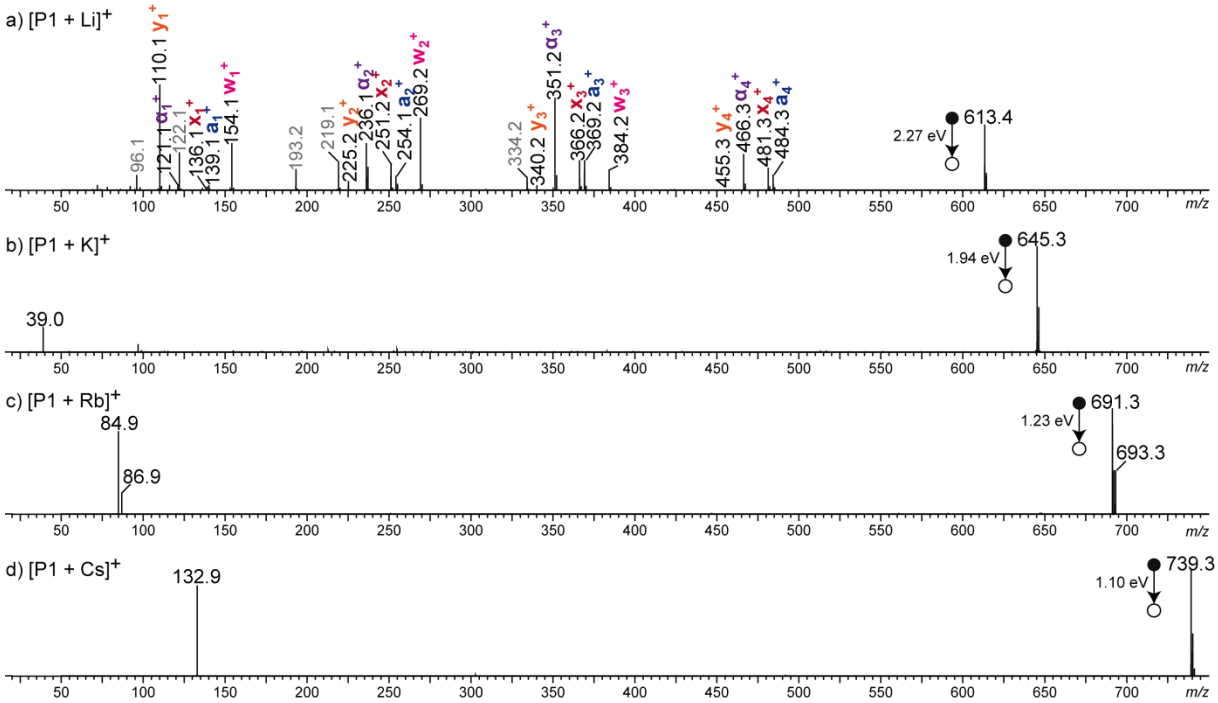


Fig. S5. ESI(+)-MS/MS of P1 when adducted with (a) Li^+ at *m/z* 613.4, (b) K^+ at *m/z* 645.3, (c) Rb^+ at *m/z* 691.3 and (d) Cs^+ at *m/z* 739.3. Activation energies are given in the center-of-mass frame. Peaks annotated in grey correspond to secondary fragments. These experiments were performed with the Synapt G2-HDMS mass spectrometer.

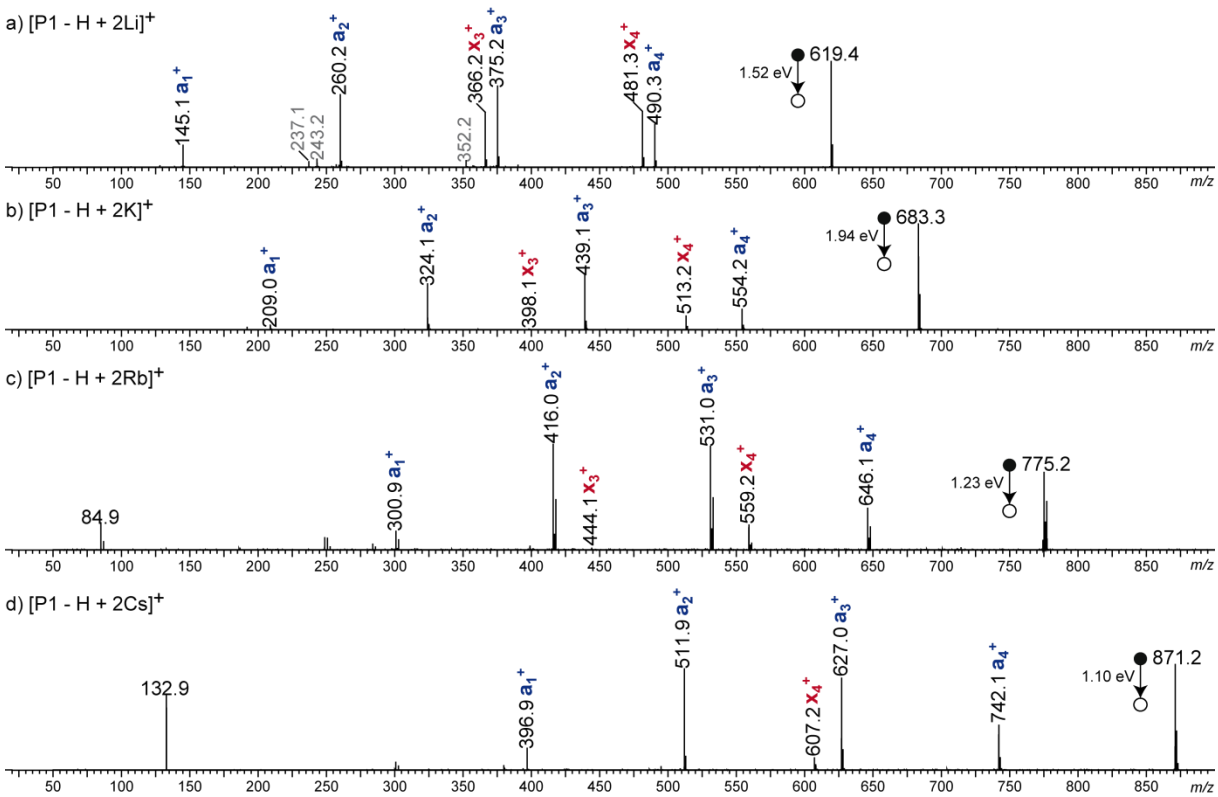


Fig S6. ESI(+)-MS/MS of (a) $[P1 - H + 2Li]^+$ at *m/z* 619.4, (b) $[P1 - H + 2K]^+$ at *m/z* 683.3, (c) $[P1 - H + 2Rb]^+$ at *m/z* 775.2 and (d) $[P1 - H + 2Cs]^+$ at *m/z* 871.2. Activation energies are given in the center-of-mass frame. Peaks annotated in grey correspond to secondary fragments. These experiments were performed with the Synapt G2-HDMS mass spectrometer.

		before HDX			*after HDX*		
		composition	m/z_{th}	m/z_{exp}	composition	m/z_{th}	m/z_{exp}
α-containing fragments	a_1^+	$C_6H_{12}O_3Na^+$	155.0679	155.0686	$C_6H_{10}D_2O_3Na^+$	157.0804	*n.d.*
	a_2^+	$C_{11}H_{21}NO_5Na^+$	270.1312	270.1318	$C_{11}H_{18}D_3NO_5Na^+$	273.1500	273.1488
	a_3^+	$C_{16}H_{30}N_2O_7Na^+$	385.1945	385.1941	$C_{16}H_{26}D_4N_2O_7Na^+$	389.2196	389.2193
	a_4^+	$C_{21}H_{39}N_3O_9Na^+$	500.2579	500.2575	$C_{21}H_{34}D_5N_3O_9Na^+$	505.2892	505.2888
	α_1^+	$C_6H_{10}O_2Na^+$	137.0573	137.0578	$C_6H_{10}O_2Na^+$	137.0573	137.0585
	α_2^+	$C_{11}H_{19}NO_4Na^+$	252.1206	252.1201	$C_{11}H_{18}DNO_4Na^+$	253.1269	253.1265
	α_3^+	$C_{16}H_{28}N_2O_6Na^+$	367.1840	367.1839	$C_{16}H_{26}D_2N_2O_6Na^+$	369.1965	369.1962
	α_4^+	$C_{21}H_{37}N_3O_8Na^+$	482.2473	482.2474	$C_{21}H_{34}D_3N_3O_8Na^+$	485.2661	485.2662
ω-containing fragments	y_1^+	$C_5H_{13}NONa^+$	126.0889	126.0893	$C_5H_{10}D_3NONa^+$	129.1078	129.1077
	y_2^+	$C_{10}H_{22}N_2O_3Na^+$	241.1523	241.1524	$C_{10}H_{18}D_4N_2O_3Na^+$	245.1774	245.1772
	y_3^+	$C_{15}H_{31}N_3O_5Na^+$	356.2156	356.2150	$C_{15}H_{26}D_5N_3O_5Na^+$	361.2470	361.2455
	y_4^+	$C_{20}H_{40}N_4O_7Na^+$	471.2789	471.2796	$C_{20}H_{34}D_6N_4O_7Na^+$	477.3166	477.3193
	x_1^+	$C_6H_{11}NO_2Na^+$	152.0682	152.0681	$C_6H_{10}DNO_2Na^+$	153.0745	153.0753
	x_2^+	$C_{11}H_{20}N_2O_4Na^+$	267.1315	267.1320	$C_{11}H_{18}D_2N_2O_4Na^+$	269.1441	269.1445
	x_3^+	$C_{16}H_{29}N_3O_6Na^+$	382.1949	382.1948	$C_{16}H_{26}D_3N_3O_6Na^+$	385.2137	385.2129
	x_4^+	$C_{21}H_{38}N_4O_8Na^+$	497.2582	497.2583	$C_{21}H_{34}D_4N_4O_8Na^+$	501.2833	501.2834
	w_1^+	$C_6H_{13}NO_3Na^+$	170.0788	170.0788	$C_6H_{10}D_3NO_3Na^+$	173.0976	173.0975
	w_2^+	$C_{11}H_{22}N_2O_5Na^+$	285.1421	285.1424	$C_{11}H_{18}D_4N_2O_5Na^+$	289.1672	289.1668
	w_3^+	$C_{16}H_{31}N_3O_7Na^+$	400.2054	400.2052	$C_{16}H_{26}D_5N_3O_7Na^+$	405.2368	405.2362
	w_4^+	$C_{21}H_{40}N_4O_9Na^+$	515.2688	515.2689	$C_{21}H_{34}D_6N_4O_9Na^+$	521.3064	*n.d.*

Table S4. Accurate mass measurement of fragments formed in CID of sodiated P1 using the precursor ion ($C_{27}H_{50}N_4O_{11}Na^+$, *m/z* 629.3368 before H/D exchanges, left; $C_{27}H_{44}D_6N_4O_{11}Na^+$, *m/z* 635.3745 after H/D exchanges, right) as an internal standard. *n.d.*: not detected.

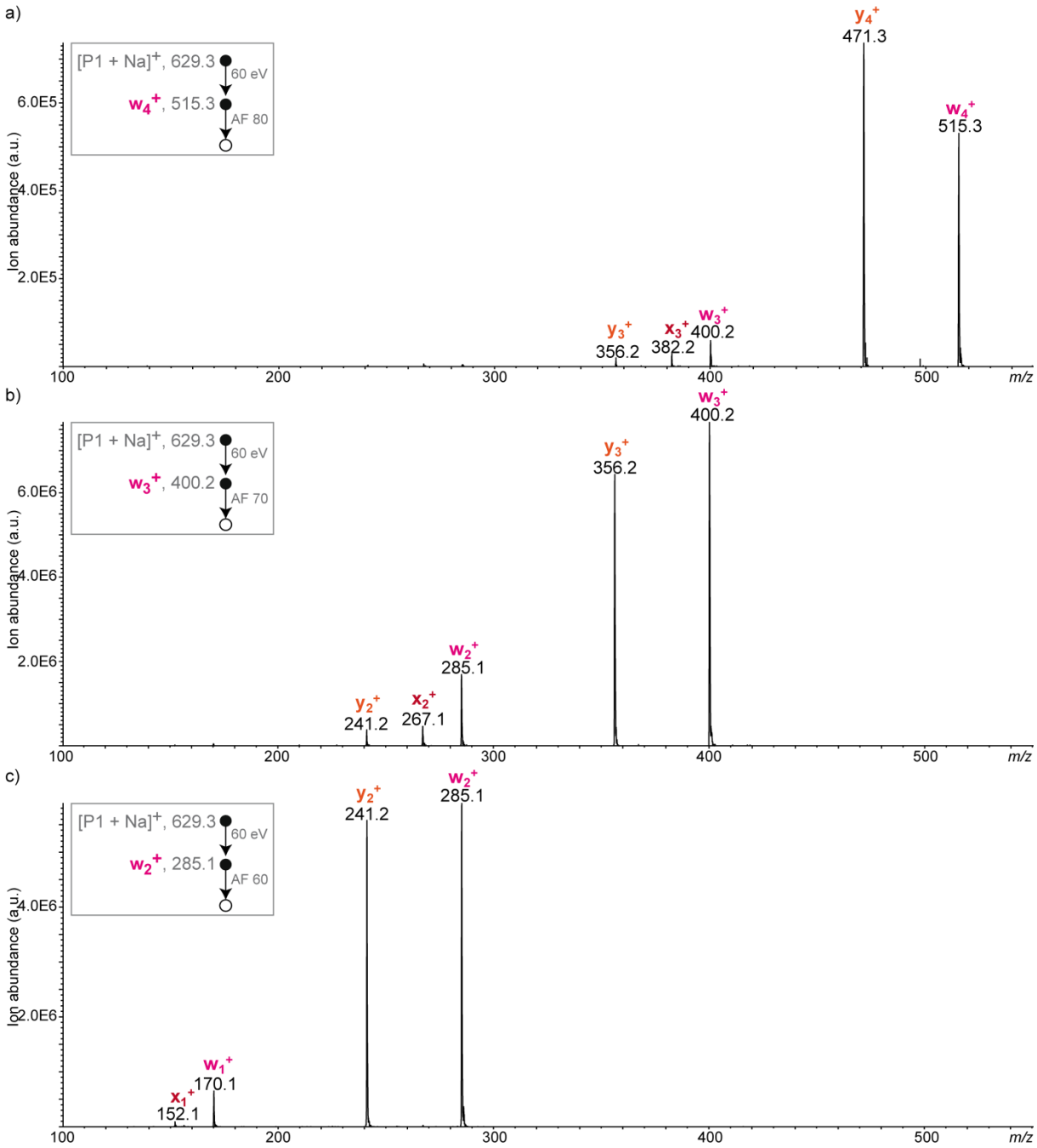


Fig. S7. MS^3 experiments performed for $[P1 + Na]^+$ at *m/z* 629.3, showing that consecutive dissociation of w_j^+ fragments (in pink) at (a) *m/z* 515.3 (w_4^+), (b) *m/z* 400.2 (w_3^+) and (c) *m/z* 285.1 (w_2^+) led to w_{j-1}^+ lower congeners as well as y_j^+ ions (in orange) and x_{j-1}^+ ions (in red). These experiments were performed with the QTrap 3200 mass spectrometer.

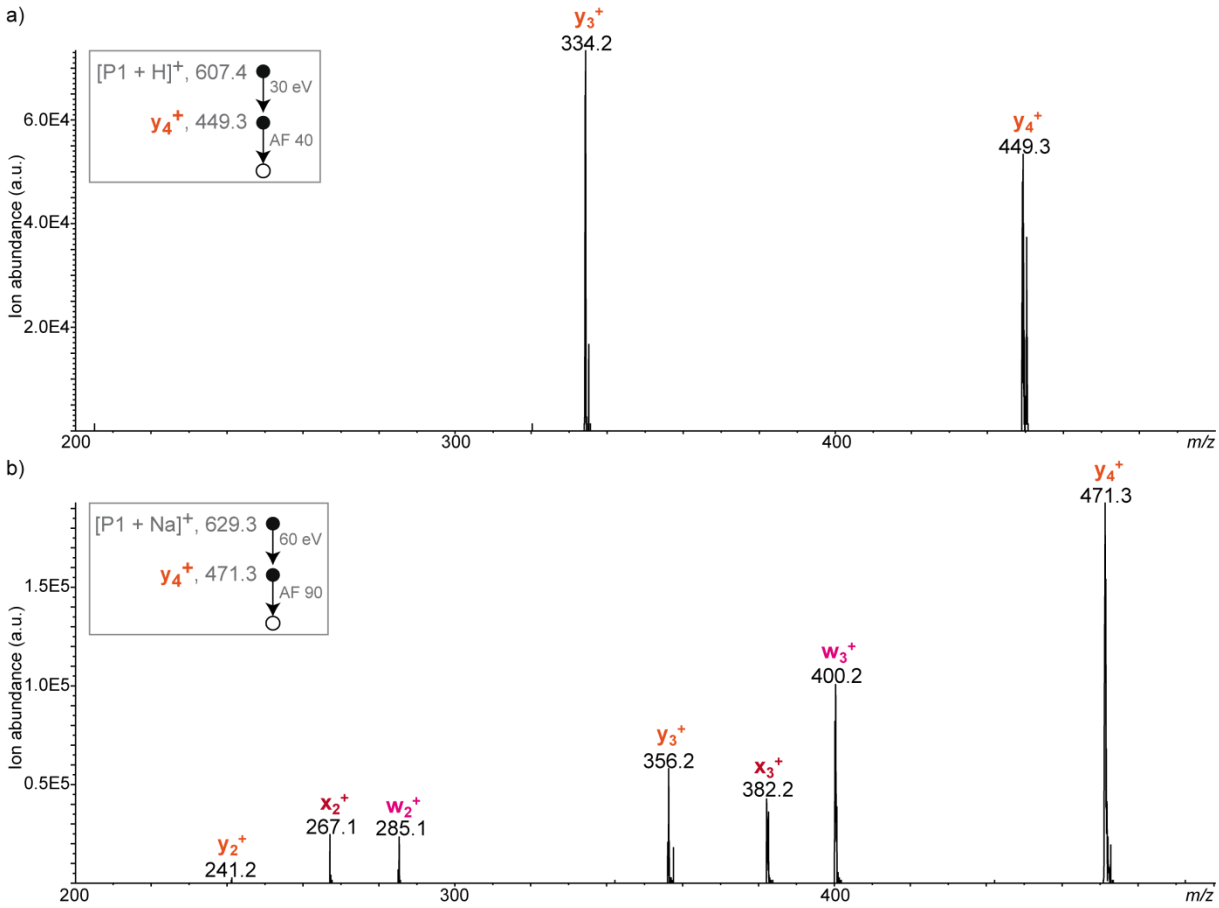
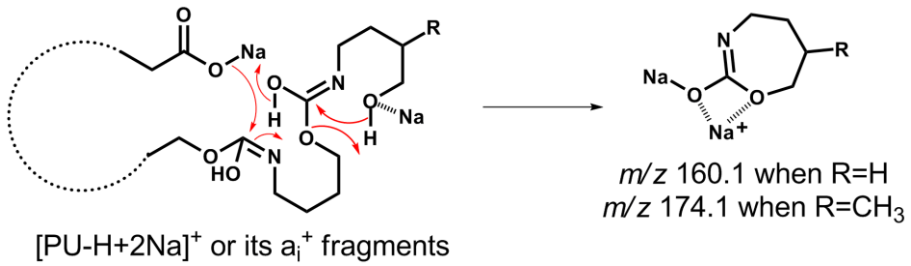


Fig. S8. MS^3 experiments showing that (a) dissociation of y_j^+ produced from for $[P1 + H]^+$ at *m/z* 607.4 mainly led to y_{j-1}^+, whereas (b) y_j^+ produced from $[P1 + Na]^+$ at *m/z* 629.3 required highly energetic activation to generate w_{j-1}^+ (in pink) and, to a lower extent, x_{j-1}^+ (in red) and y_{j-1}^+ (in orange) ions. These experiments were performed with the QTrap 3200 mass spectrometer.



Scheme S1. Mechanism proposed to account for internal fragments generated upon CID of $[PU - H + 2Na]^+$ precursor ions.

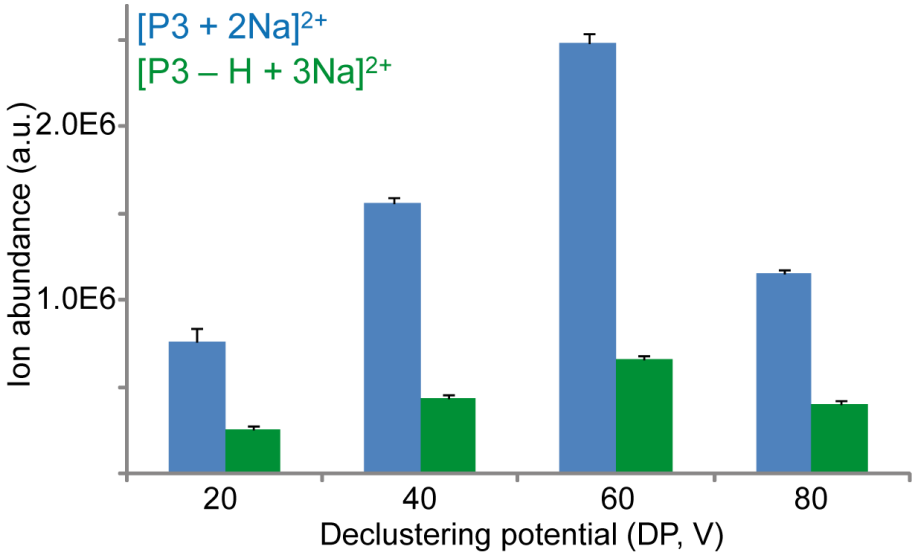


Fig. S9. Influence of the declustering potential (DP) on the abundance of $[P3 + 2Na]^{2+}$ at *m/z* 563.3 (in blue) and of $[P3 – H +3Na]^{2+}$ at *m/z* 574.3 (in green). Error bars as standard deviation of 3 replicates. These experiments were performed with the Synapt G2-HDMS mass spectrometer.

a_i^{z+} ions					x_j^{z+} ions			
composition	m/z_{th}	m/z_{exp}	i↓	**P2**	j↑	m/z_{exp}	m/z_{th}	composition
$C_6H_{11}O_3Na_2^+$	177.0498	177.0486	1	**α**				
$C_{12}H_{22}NO_5Na_2^+$	306.1288	306.1263	2	**1**	16	999.5544	999.5585	$C_{88}H_{160}N_{16}O_{32}Na_2^{2+}$
$C_{18}H_{33}N_2O_7Na_2^+$	435.2078	435.2040	3	**1**	15	935.0143	935.0190	$C_{82}H_{149}N_{15}O_{30}Na_2^{2+}$
$C_{23}H_{42}N_3O_9Na_2^+$	550.2711	550.2673	4	**0**	14	870.4756	870.4795	$C_{76}H_{138}N_{14}O_{28}Na_2^{2+}$
$C_{29}H_{53}N_4O_{11}Na_2^+$	679.3501	679.3465	5	**1**	13	812.9430	812.9478	$C_{71}H_{129}N_{13}O_{26}Na_2^{2+}$
$C_{34}H_{62}N_5O_{13}Na_2^+$	794.4134	794.4096	6	**0**	12	1473.8174	1473.8274	$C_{65}H_{118}N_{12}O_{24}Na^+$
$C_{39}H_{71}N_6O_{15}Na_2^+$	909.4767	909.4727	7	**0**	11	1358.7594	1358.7641	$C_{60}H_{109}N_{11}O_{22}Na^+$
$C_{45}H_{82}N_7O_{17}Na_2^+$	1038.5557	1038.5612	8	**1**	10	1243.6989	1243.7008	$C_{55}H_{100}N_{10}O_{20}Na^+$
$C_{50}H_{91}N_8O_{19}Na_2^+$	1153.6190	1153.6168	9	**0**	9	1114.6183	1114.6218	$C_{49}H_{89}N_9O_{18}Na^+$
$C_{56}H_{102}N_9O_{21}Na_2^+$	1282.6980	1282.7000	10	**1**	8	999.5544	999.5585	$C_{44}H_{80}N_8O_{16}Na^+$
$C_{62}H_{113}N_{10}O_{23}Na_3^{2+}$	717.3831	717.3796	11	**1**	7	870.4756	870.4795	$C_{38}H_{69}N_7O_{14}Na^+$
$C_{67}H_{122}N_{11}O_{25}Na_3^{2+}$	774.9148	774.9111	12	**0**	6	741.3952	741.4005	$C_{32}H_{58}N_6O_{12}Na^+$
$C_{73}H_{133}N_{12}O_{27}Na_3^{2+}$	839.4543	839.4505	13	**1**	5	626.3334	626.3372	$C_{27}H_{49}N_5O_{10}Na^+$
$C_{79}H_{144}N_{13}O_{29}Na_3^{2+}$	903.9938	903.9894	14	**1**	4	497.2544	497.2582	$C_{21}H_{38}N_4O_8Na^+$
$C_{84}H_{153}N_{14}O_{31}Na_3^{2+}$	961.5254	961.5212	15	**0**	3	368.1763	368.1792	$C_{15}H_{27}N_3O_6Na^+$
$C_{89}H_{162}N_{15}O_{33}Na_3^{2+}$	1019.0571	1019.0555	16	**0**	2	253.1169	253.1159	$C_{10}H_{18}N_2O_4Na^+$
-	-	-	*n.e.*	**0**	1	*n.d.*	138.0525	$C_5H_9NO_2Na^+$

Table S5. Accurate mass measurement of a_i^{z+} and x_j^{z+} fragments formed in CID of [P2 – H +3Na]$^{2+}$ (Figure 3) using the precursor ion ($C_{94}H_{171}N_{16}O_{35}Na_3^{2+}$, *m/z* 1076.5887) as an internal standard. Sequence of P2 is 1101001011011000, with 0: $C_5H_9NO_2$ and 1: $C_6H_{11}NO_2$. *N.e.*: not expected. *N.d.*: not detected.