



Cite this: *Polym. Chem.*, 2021, **12**, 5279

Received 6th August 2021,  
Accepted 10th September 2021  
DOI: 10.1039/d1py01052b

rsc.li/polymers

## Synthesis and sequencing of informational poly(amino phosphodiester)s†

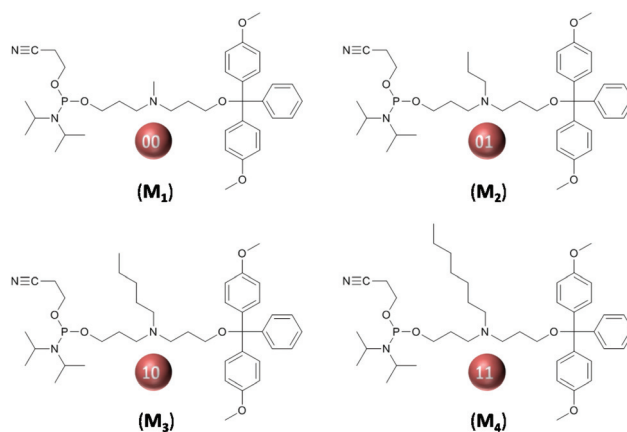
Ian Roszak,<sup>a</sup> Laurence Oswald,<sup>a</sup> Abdelaziz Al Ouahabi,<sup>a</sup> Annabelle Bertin,<sup>a,b,c</sup> Eline Laurent,<sup>a</sup> Olivier Felix,<sup>id</sup> <sup>a</sup> Isaure Carvin-Sergent,<sup>id</sup> <sup>d</sup> Laurence Charles<sup>id</sup> <sup>d</sup> and Jean-François Lutz<sup>id</sup> \*<sup>a</sup>

**Sequence-defined poly(amino phosphodiester)s containing main-chain tertiary amines were synthesized by automated solid-phase phosphoramidite chemistry. These polymers were prepared using four monomers with different substituents. The formed polymers were characterized by HPLC and mass spectrometry. These methods evidenced preparation of molecularly-defined polymers. Furthermore, the presence of tertiary amines in the polymer backbones facilitates sequencing by tandem mass spectrometry.**

The stepwise synthesis of sequence-defined polymers has recently received increasing attention in fundamental polymer science.<sup>1–4</sup> Indeed, the precise control of comonomer sequences enables new macromolecular properties to be achieved. For instance, it allows preparation of synthetic informational polymers, in which comonomers are used as a molecular language.<sup>5</sup> Such information-containing polymers are used in various applications,<sup>6</sup> in which they are often decoded by a sequencing analytical method.<sup>7</sup> We<sup>8–10</sup> and others<sup>11–16</sup> have reported examples of such oligomers and polymers. These macromolecules have a uniform chain-length (*i.e.* apparent polydispersity close to 1.0) and a perfectly controlled comonomer sequence. Yet, they are not necessarily stereoregular. Since informational polymers are often decoded by analytical methods that are not sensitive to tacticity (*e.g.* tandem mass spectrometry, MS/MS), racemic synthons can be used to construct them. For instance, our group has reported the synthesis of digital poly(phosphodiester)s (d-PPDE),<sup>17–19</sup> in which a comonomer alphabet is used to write binary information.<sup>20</sup>

These polymers are prepared by solid-phase phosphoramidite polymer chemistry (PPC); a versatile method that can be used for synthesis of biological (*i.e.* DNA and RNA),<sup>21</sup> abiological<sup>19,22–24</sup> and bio-hybrid macromolecules.<sup>25–28</sup> In all d-PPDE we have studied so far, the coded side-chains are attached to a main-chain carbon atom.<sup>19</sup> In order to avoid the formation of atactic polymers, symmetric carbons were used in our initial design.<sup>17</sup> Still, we have recently developed expanded alphabets, in which racemic building blocks were also used.<sup>29,30</sup> Although these polymers can be accurately deciphered by MS/MS, the formed backbones contain asymmetric carbons and are therefore atactic. This could be more problematic for other sequencing methods such as nanopore sensing.<sup>31</sup>

A convenient design to avoid tacticity issues is to install coded substituents on a main-chain nitrogen atom. This strategy is used for preparing achiral sequence-defined polymers such as peptide nucleic acids,<sup>32</sup> peptoids<sup>33</sup> and poly(*N*-substituted urethanes).<sup>34</sup> Various amine-containing phosphoramidite monomers have been reported in the literature. Behr and coworkers have reported the synthesis of oligospermine by



**Chart 1** Molecular structure of the monomers prepared and used in this work.

<sup>a</sup>Université de Strasbourg, CNRS, Institut Charles Sadron UPR22, 23 rue du Loess, 67034 Strasbourg Cedex 2, France. E-mail: jflutz@unistra.fr

<sup>b</sup>BAM Federal Institute for Materials Research and Testing, Unter den Eichen 87, 12205 Berlin, Germany

<sup>c</sup>Institute of Chemistry and Biochemistry – Organic Chemistry, Free University Berlin, Takustraße 3, 14195 Berlin, Germany

<sup>d</sup>Aix Marseille Université, CNRS, UMR 7273, Institute of Radical Chemistry, 13397 Marseille Cedex 20, France

†Electronic supplementary information (ESI) available. See DOI: 10.1039/d1py01052b

phosphoramidite chemistry.<sup>35,36</sup> More recently, Sleiman and coworkers have described the synthesis of poly(phosphodiester)s containing main-chain tertiary amines that were used as

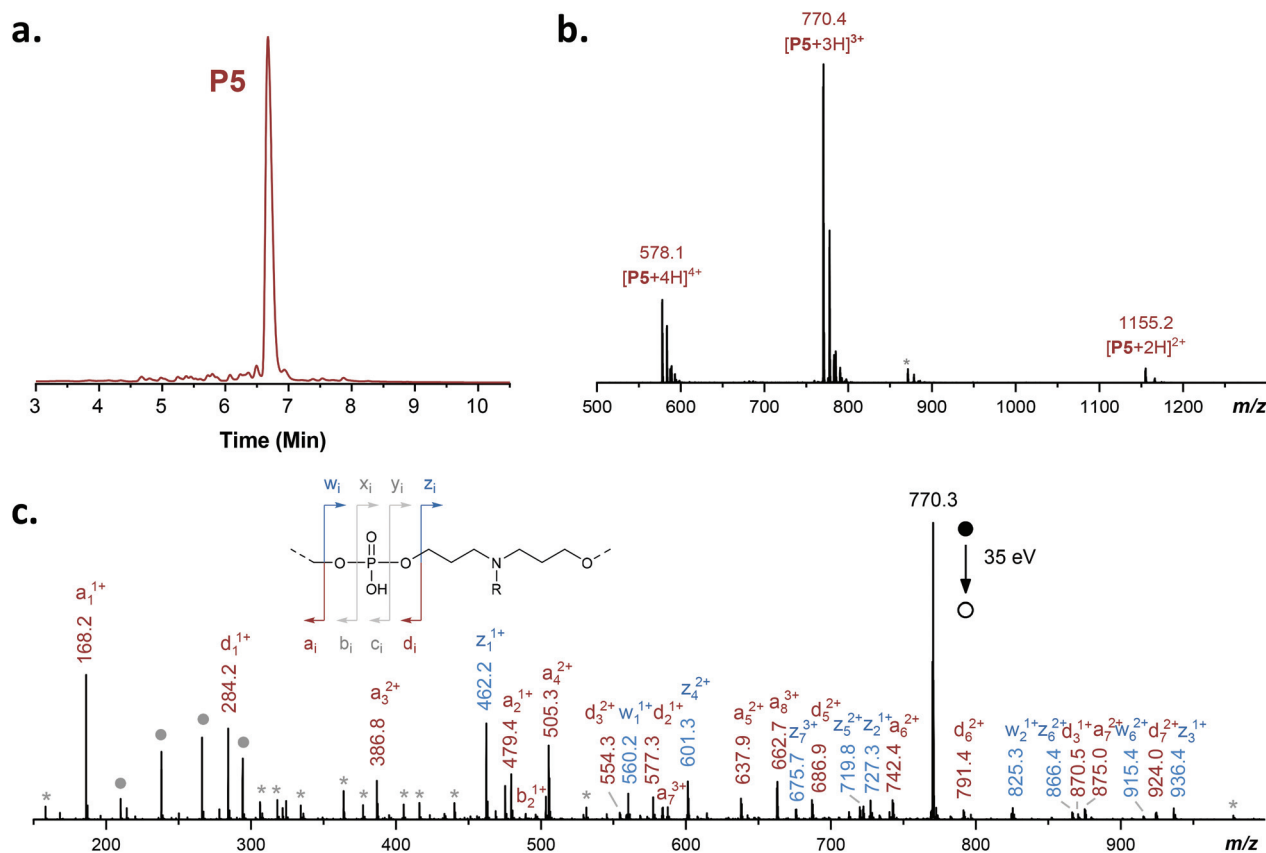
a platform for side-chain functionalization.<sup>37</sup> In the present work, a comparable design was applied for the first time to the synthesis of informational d-PPDE. Four phosphoramidite monomers containing main-chain nitrogen atom were synthesized and tested for the preparation of digitally-encoded poly(phosphodiester)s (Chart 1).

All the digital polymers investigated herein were synthesized by automated PPC on controlled pore glass solid supports.<sup>18</sup> In this multistep synthesis, a cycle of four successive reactions, namely dimethoxytrityl (DMT) deprotection, phosphoramidite coupling, capping and oxidation, is repeated a certain number of times until a desired chain-length is reached. In order to store digital information in the polymer chains, a set of four phosphoramidite monomers was developed in this work (Chart 1). They all contain a main-chain nitrogen atom and alkyl side-chain substituents of increasing size, namely methyl (**M1**), propyl (**M2**), pentyl (**M3**) and heptyl (**M4**). Following our laboratory convention,<sup>19</sup> monomers of increasing molar mass **M1**, **M2**, **M3** and **M4** represent binary dyads 00, 01, 10 and 11, respectively. Main-chain propyl spacers were intentionally placed between oxygen and nitrogen atoms in all these monomers. Attempts were also made with

**Table 1** Sequence-defined poly(amino phosphodiester)s prepared in this work<sup>a</sup>

|    | Sequence <sup>a</sup>  | $m/z_{th}$            | $m/z$    |
|----|--|-----------------------|----------|
| H1 | ( <b>M1</b> ) <sub>10</sub> -T   | 584.2341 <sup>b</sup> | 584.2339 |
| H2 | ( <b>M2</b> ) <sub>10</sub> -T   | 654.5631 <sup>b</sup> | 654.5637 |
| H3 | ( <b>M3</b> ) <sub>10</sub> -T   | 724.3906 <sup>b</sup> | 724.3914 |
| H4 | ( <b>M4</b> ) <sub>10</sub> -T   | 794.4688 <sup>b</sup> | 794.4683 |
| P1 | ( <b>M2-M1</b> ) <sub>4</sub> -T   | 676.6303 <sup>c</sup> | 676.6305 |
| P2 | <b>M1</b> -( <b>M2</b> ) <sub>3</sub> -( <b>M1</b> ) <sub>4</sub> -T   | 667.2865 <sup>c</sup> | 667.2880 |
| P3 | <b>M2-M1</b> -( <b>M2</b> ) <sub>2</sub> - <b>M1</b> -( <b>M2</b> ) <sub>3</sub> -T                          | 695.3178 <sup>c</sup> | 695.3181 |
| P4 | <b>M1</b> -( <b>M2</b> ) <sub>4</sub> -( <b>M1</b> ) <sub>6</sub> -( <b>M2</b> ) <sub>4</sub> - <b>M1</b> -T | 763.3368 <sup>d</sup> | 763.3375 |
| P5 | <b>M3</b> -( <b>M4</b> ) <sub>2</sub> - <b>M2</b> - <b>M3</b> - <b>M1</b> - <b>M3</b> - <b>M2</b> -T         | 770.0680 <sup>c</sup> | 770.0690 |
| P6 | ( <b>M4</b> ) <sub>2</sub> - <b>M3</b> - <b>M1</b> - <b>M2</b> -( <b>M3</b> ) <sub>2</sub> - <b>M4</b> -T    | 788.7555 <sup>c</sup> | 788.7568 |
| P7 | ( <b>M1</b> - <b>M2</b> - <b>M3</b> - <b>M4</b> ) <sub>2</sub> -T  | 751.3804 <sup>c</sup> | 751.3813 |
| P8 | <b>M1</b> - <b>M2</b> - <b>M3</b> -( <b>M4</b> ) <sub>2</sub> - <b>M3</b> - <b>M2</b> - <b>M1</b> -T         | 751.3804 <sup>c</sup> | 751.3819 |

<sup>a</sup>The letter T denotes a terminal thymidine that comes from the pre-loaded solid-support. <sup>b</sup>Detected as  $[M + 4H]^{4+}$ . <sup>c</sup>Detected as  $[M + 3H]^{3+}$ . <sup>d</sup>Detected as  $[M + 5H]^{5+}$ .



**Fig. 1** Characterization of copolymer P5. (a) HPLC trace. (b) Positive mode ESI-MS. The grey star indicates traces of DMT-protected polymer. All other unlabeled peaks are cation adducts. (c) MS/MS sequencing. This spectrum was obtained by collision-induced dissociation of  $[P5 + 3H]^{3+}$ . Grey stars denote internal fragments. The full grey circles indicate internal fragments corresponding to the four different coded subunits. The inset schematizes the fragmentation pattern of a phosphate repeat unit. Light grey arrows indicate fragmentation series that are nearly detectable in these measurements.

ethyl-based monomers that can be more easily derived from commercially-available products. However, as previously described,<sup>37</sup> low yields and contaminated oligomers were obtained in these syntheses (data not shown). Table 1 and Scheme S1† list all the polymers that were synthesized in this work. At first homopolymers **H1–H4** were synthesized in order to verify that all monomers **M1–M4** can be efficiently used in PPC. In all cases, HPLC, ESI-MS and MS/MS measurements (Fig. S1–S4† and Table 1) indicated formation of uniform polymers with the expected chain-length and homopolymer sequence.

Based on these promising results, different copolymers **P1–P8** were synthesized and characterized (Table 1). Copolymers **P1–P4** were synthesized using only comonomers **M1** and **M2**, whereas **P5–P8** were prepared using all four comonomers **M1–M4**. Overall, all copolymers appeared near-monodisperse, as evidenced by HPLC and ESI-MS (Fig. 1 and Fig. S5–S11†). Furthermore, all copolymers could be sequenced and decrypted by MS/MS (Fig. 1 and Fig. S5–S11†). While d-PPDE are usually exclusively sequenced by negative-mode MS/MS,<sup>38</sup> the presence of a main-chain nitrogen atom in copolymers **P1–P8** enables sequencing in both negative and positive mode. Moreover, it was found that dissociation of protonated species formed in the latter mode is highly beneficial to sequencing compared to the former one. Indeed, as schematized in the inset of Fig. 1c and detailed in Fig. S12,† phosphate repeat units usually lead to eight fragment ion series once deprotonated in the negative ion mode.<sup>38</sup> Yet, in positive mode MS/MS, four of them (noted in grey in the inset) are either weakly or not detected. As a consequence, signal dilution is minimized and the other four series are detected with an unusually high intensity, thus allowing a perfect sequence coverage. As compared to conventional negative mode (Fig. S12†), it permits to envisage accurate sequencing of poly(amino phosphodiester)s from data recorded at very low concentrations or with very short acquisition times.

In summary, digitally-encoded poly(amino phosphodiester)s were successfully synthesized and decoded by MS/MS. The presence of a main-chain nitrogen atom in these informational copolymers brings substantial advantages. First of all, it is a convenient site for side-chain functionalization. In addition, it enables positive mode MS analysis, thus facilitating sequencing. Last but not least, such polymers, containing both main chain phosphates and tertiary amines, are potential polyampholytes and could therefore be useful for electrostatic self-assembly.<sup>39</sup>

## Conflicts of interest

There are no conflicts to declare.

## Notes and references

- J.-F. Lutz, M. Ouchi, D. R. Liu and M. Sawamoto, *Science*, 2013, **341**, 1238149.
- T. T. Trinh, C. Laure and J.-F. Lutz, *Macromol. Chem. Phys.*, 2015, **216**, 1498–1506.
- S. C. Solleder, R. V. Schneider, K. S. Wetzels, A. C. Boukiss and M. A. R. Meier, *Macromol. Rapid Commun.*, 2017, **38**, 1600711.
- P. Nanjan and M. Porel, *Polym. Chem.*, 2019, **10**, 5406–5424.
- J.-F. Lutz, *Macromolecules*, 2015, **48**, 4759–4767.
- M. A. R. Meier and C. Barner-Kowollik, *Adv. Mater.*, 2019, **31**, 1806027.
- H. Mutlu and J.-F. Lutz, *Angew. Chem., Int. Ed.*, 2014, **53**, 13010–13019.
- T. T. Trinh, L. Oswald, D. Chan-Seng and J.-F. Lutz, *Macromol. Rapid Commun.*, 2014, **35**, 141–145.
- R. K. Roy, A. Meszynska, C. Laure, L. Charles, C. Verchin and J.-F. Lutz, *Nat. Commun.*, 2015, **6**, 7237.
- U. S. Gunay, B. E. Petit, D. Karamessini, A. Al Ouahabi, J.-A. Amalian, C. Chendo, M. Bouquey, D. Gigmès, L. Charles and J. F. Lutz, *Chem*, 2016, **1**, 114–126.
- A. C. Boukiss and M. A. R. Meier, *Eur. Polym. J.*, 2018, **104**, 32–38.
- S. Martens, A. Landuyt, P. Espeel, B. Devreese, P. Dawyndt and F. Du Prez, *Nat. Commun.*, 2018, **9**, 4451.
- J. O. Holloway, F. Van Lijsebetten, N. Badi, H. A. Houck and F. E. Du Prez, *Adv. Sci.*, 2020, **7**, 1903698.
- J. M. Lee, M. B. Koo, S. W. Lee, H. Lee, J. Kwon, Y. H. Shim, S. Y. Kim and K. T. Kim, *Nat. Commun.*, 2020, **11**, 56.
- B. Liu, Q. Shi, L. Hu, Z. Huang, X. Zhu and Z. Zhang, *Polym. Chem.*, 2020, **11**, 1702–1707.
- K. S. Wetzels, M. Frölich, S. C. Solleder, R. Nickisch, P. Treu and M. A. R. Meier, *Commun. Chem.*, 2020, **3**, 63.
- A. Al Ouahabi, L. Charles and J.-F. Lutz, *J. Am. Chem. Soc.*, 2015, **137**, 5629–5635.
- A. Al Ouahabi, M. Kotera, L. Charles and J.-F. Lutz, *ACS Macro Lett.*, 2015, **4**, 1077–1080.
- L. Charles and J.-F. Lutz, *Acc. Chem. Res.*, 2021, **54**, 1791–1800.
- H. Colquhoun and J.-F. Lutz, *Nat. Chem.*, 2014, **6**, 455–456.
- M. H. Caruthers, *Science*, 1985, **230**, 281–285.
- N. Appukkutti and C. J. Serpell, *Polym. Chem.*, 2018, **9**, 2210–2226.
- M. Vybornyi, Y. Vyborna and R. Häner, *Chem. Soc. Rev.*, 2019, **48**, 4347–4360.
- N. Appukkutti, J. R. Jones and C. J. Serpell, *Chem. Commun.*, 2020, **56**, 5307–5310.
- T. G. W. Edwardson, K. M. M. Carneiro, C. J. Serpell and H. F. Sleiman, *Angew. Chem., Int. Ed.*, 2014, **53**, 4567–4571.
- T. Mondal, M. Nerantzaki, K. Flesch, C. Loth, M. Maaloum, Y. Cong, S. S. Sheiko and J.-F. Lutz, *Macromolecules*, 2021, **54**, 3423–3429.
- C. Loth, L. Charles, J.-F. Lutz and M. Nerantzaki, *ACS Macro Lett.*, 2021, **10**, 481–485.
- M. Nerantzaki, C. Loth and J.-F. Lutz, *Polym. Chem.*, 2021, **12**, 3498–3509.
- E. Laurent, J.-A. Amalian, M. Parmentier, L. Oswald, A. Al Ouahabi, F. Dufour, K. Launay, J.-L. Clément, D. Gigmès, M.-A. Delsuc, L. Charles and J.-F. Lutz, *Macromolecules*, 2020, **53**, 4022–4029.

- 30 E. Laurent, J. A. Amalian, T. Schutz, K. Launay, J.-L. Clément, D. Gigmes, A. Burel, C. Carapito, L. Charles, M.-A. Delsuc and J. F. Lutz, *C. R. Chim.*, 2021, **24**, 69–76.
- 31 C. Cao, L. F. Krapp, A. Al Ouahabi, N. F. König, N. Cirauqui, A. Radenovic, J.-F. Lutz and M. D. Peraro, *Sci. Adv.*, 2020, **6**, eabc2661.
- 32 M. Egholm, O. Buchardt, P. E. Nielsen and R. H. Berg, *J. Am. Chem. Soc.*, 1992, **114**, 1895–1897.
- 33 J. Sun and R. N. Zuckermann, *ACS Nano*, 2013, **7**, 4715–4732.
- 34 T. Mondal, V. Greff, B. É. Petit, L. Charles and J.-F. Lutz, *ACS Macro Lett.*, 2019, **8**, 1002–1005.
- 35 E. Voirin, J.-P. Behr and M. Kotera, *Nat. Protocols*, 2007, **2**, 1360–1367.
- 36 R. Noir, M. Kotera, B. Pons, J.-S. Remy and J.-P. Behr, *J. Am. Chem. Soc.*, 2008, **130**, 13500–13505.
- 37 D. de Rochambeau, Y. Sun, M. Barlog, H. S. Bazzi and H. F. Sleiman, *J. Org. Chem.*, 2018, **83**, 9774–9786.
- 38 J. A. Amalian, A. Al Ouahabi, G. Cavallo, N. F. König, S. Poyer, J. F. Lutz and L. Charles, *J. Mass Spectrom.*, 2017, **52**, 788–798.
- 39 R. Szweda, M. Tschopp, O. Felix, G. Decher and J.-F. Lutz, *Angew. Chem.*, 2018, **130**, 16043–16047.