

A Simple Post-Polymerization Modification Method for Controlling Side-Chain Information in Digital Polymers

Niklas Felix König, Abdelaziz Al Ouahabi, Salomé Poyer, Laurence Charles, Jean-François Lutz

▶ To cite this version:

Niklas Felix König, Abdelaziz Al Ouahabi, Salomé Poyer, Laurence Charles, Jean-François Lutz. A Simple Post-Polymerization Modification Method for Controlling Side-Chain Information in Digital Polymers. Angewandte Chemie International Edition, 2017, 56 (25), pp.7297-7301. 10.1002/anie.201702384. hal-01957112

HAL Id: hal-01957112 https://hal.science/hal-01957112v1

Submitted on 17 Dec 2018

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: N. F. König, A. A. Ouahabi, S. Poyer, L. Charles, J.-F. Lutz, Angew. Chem. Int. Ed. **2017**, 56, 7297-7301, which has been published in final form at <u>https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.201702384</u>. This article may be used for non-commercial purposes in accordance with Wiley Terms

and Conditions for Self-Archiving."

A simple post-polymerization modification method for controlling sidechain information in digital polymers

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

Abstract: A three-step post-polymerization modification procedure was studied for the design of digitally-encoded poly(phosphodiester)s with controllable side-groups. Sequence-defined precursors were synthesized, either manually on polystyrene resins or automatically on controlled pore glass supports, using two phosphoramidite monomers containing either terminal alkynes or triisopropylsilyl (TIPS)-protected alkyne side-groups. Afterwards, these polymers were modified by stepwise copper-catalyzed azide-alkyne cycloaddition (CuAAC). The terminal alkynes were first reacted with a model azide compound and, after deprotection of the TIPS groups, the remaining alkynes were reacted with another organic azide. This simple method allows quantitative side-chain modification, thus opening up interesting avenues for the preparation of a wide variety of digital polymers. Yet, optimized protocols reported herein shall be followed. For instance, the use of iodine in iterative phosphoramidite chemistry may result in the formation of unwished iodo-acetylenes. Practical solutions to bypass that problem are reported herein.

 MSc N. F. König, Dr. A. Al Ouahabi, Dr. J.-F. Lutz ; Université de Strasbourg, CNRS, Institut Charles Sadron UPR22, 23 rue du Loess, 67034 Strasbourg Cedex 2 (France)
 E-mail: jflutz@unistra.fr

 ^[2] Prof. L. Charles, Dr. S. Poyer Aix-Marseille Univ., CNRS, ICR UMR7273, 13397 Marseille (France) E-mail: laurence.charles@univ-amu.fr

Supporting information for this article is given via a link at the end of the document.

Digital polymers are natural or non-natural information-containing macromolecules that store binary information in their monomer sequences.^[1] For instance, DNA, which is the natural storage medium of genetic information, can be engineered to store digital data.^[2] Alternatively, digital information can be stored in a wide variety of synthetic polymers, including poly(triazole amide)s,^[3] poly(phosphodiester)s,^[4] poly(alkoxyamine phosphodiester)s,^[5] poly(alkoxyamine amide)s,^[6] and polyurethanes.^[7] Digital polymers are currently studied for different applications including data storage,^[8] long-term storage^[9] and anti-counterfeiting tags.^[10] In all cases, information is usually written in the polymer chains using a step-by-step iterative process and is read using a sequencing technology.^[1c] Numerous sequencing methods are available for deciphering the sequences of information-containing macromolecules.^[11] For instance, binary information stored on non-natural polymer backbones may be potentially deciphered using tandem mass spectrometry (MS/MS), protein nanopore sequencing, solidstate nanopore sequencing and controlled depolymerization. It was pointed out that the molecular structure of non-natural polymers can be intentionally tuned to facilitate sequencing.^[1c] For instance, it was observed that the presence of repeated alkoxyamine and carbamate linkages facilitates MS/MS sequencing of information-containing macromolecules.^[5-7] However, polymers that are tailor-made for one particular sequencing technique are not necessarily optimal for another one. For instance, nanopore sequencing,^[12] which is a non-destructive analytical technique, may require specific readable side-chain motifs rather than breakable main-chain motifs. In this approach, sequence-defined polyelectrolytes are subjected to an ionic current and thread through nanopores of defined diameter. The interactions of the monomers with the pore lead to voltage variations that can be monitored and related to a sequence. Yet, this technique is not trivial and DNA pore sequencing required about twenty years of optimization.^[12] This is, in part, due to the fact that DNA nucleobases have closely-related structures (i.e. purine and pyrimidine heterocycles) that are difficult to distinguish from another. In comparison, the molecular bits of non-natural sequence-coded polymers could be intentionally engineered to interact to a higher or smaller extent with a pore, thus rendering the analytes more "readable" than DNA. More generally speaking, the molecular structure of monomer bits may greatly influence the storage density and readability of digital polymers.^[1c] In this context, it seems crucial to develop synthetic methods that permit to tune easily side-chain information in digital polymers. Here, we describe a universal approach, which allows side-chain modification in digitally-encoded poly(phosphodiester)s.



Scheme 1. General concept herein for the selective functionalization of sequence-defined poly(phosphodiester)s: (1) CuAAC modification of the 0-bit units, (2) deprotection of the 1-bit units; (3) CuAAC modification of the 1-bit units.

Scheme 1 shows the strategy studied in this work. Several interesting routes have been described during the last few years for the synthesis of sequence-defined polymers.^[13] Among them, iterative phosphoramidite chemistry, which was initially introduced for the synthesis of oligonucleotides, has been shown to be also a robust method for the preparation of non-natural sequence-defined poly(phosphodiester)s.^[4a, 14] This method can be performed manually but can also be automated, thus allowing preparation of long coded chains.^[4b]

Phosphoramidite chemistry was therefore selected herein for the synthesis of digitally-encoded polymers of different lengths. These precursors contained two reactive bits 0 and 1 as shown in Scheme 1; one of them bears terminal alkynes and the other TIPS-protected alkynes. Although a variety of reactions can be performed on acetylenes, CuAAC was chosen for modifying these precursors. Indeed, this reaction was shown to be versatile for stepwise modification of DNA and synthetic polymers.^[15] The terminal alkynes in 0-bits were first reacted with an organic azide R_1 - N_3 to afford substituted triazoles. Afterwards, the protecting TIPS groups in 1-bits were removed and the remaining alkynes were reacted by CuAAC with another azide R_2 - N_3 .



Figure 1. (a) Monomer synthesis: i) NaH, (protected) propargylbromide; ii) LiAlH₄; iii) DMTr chloride; iv) 2-cyanoethyl-N,N-diisopropylchlorophosphoramidite. (b) Functionalization of polystyrene resin: v) succinic anhydride, DMAP; vi) aminomethylated polystyrene, DMAP, DCC. (c) Commercial dT-CPG support modified with deoxythymidine. (d) Example of iterative synthesis on polystyrene support: vii) DMT deprotection with TCA; viii) coupling: **5a** or **5b**, 1H-tetrazole; ix) oxidation: I₂; x) phosphate deprotection with piperidine; xi) iodine reduction with sodium thiosulfate; xii) cleavage from resin with aqueous methylamine/ammonia. (e) Example of post-polymerization modification: xiii) R₁-N₃**10**, CuBr, TBTA, sodium ascorbate; xiv) TBAF; xv) R₂-N₃**11**, CuBr, TBTA, sodium ascorbate.

Two phosphoramidite monomers were used to build the information-containing polymers: **5a** that bears two pendant terminal alkynes and **5b** which contains two TIPS-protected alkyne functions (Figure 1a). The synthesis of **5a** begins with the reaction of diethylmalonate and propargyl bromide (*i*) and is followed by a reduction to the corresponding diol **3a** (*ii*). For **5b**, the corresponding TIPS-functionalized propargyl bromide was employed in step (*i*). For both monomers, one alcohol function of the diols is then mono-protected by a 4,4'-dimethoxytrityl (DMT) group (*iii*) and the remaining alcohol is converted into a reactive phosphoramidite (*iv*). Both **5a** and **5b** were obtained in excellent yields. The mono-protected intermediates **4a** and **4b** were also used to prepare polystyrene supports **7a** and **7b** for manual iterative synthesis (Figure 1b). They were first transformed into base-labile succinidyl ester linkers **6a** and **6b** (*v*) and then coupled to an amino-methylated polystyrene resin (*vi*). Besides modified polystyrene resins, a commercial controlled pore glass support **8** was also used for automated synthesis (Figure 1c).

The sequence-coded precursors were synthesized by iterative phosphoramidite chemistry. Classical cycles involving DMT-deprotection (vii), alcohol-phosphoramidite coupling (viii) and oxidation of the formed phosphites into phosphates (ix) were used. In order to illustrate the versatility of our approach, some polymers were prepared manually on polystyrene supports 7a and 7b (Figure 1d), whereas some others were prepared on 8 using an automated DNA synthesizer. In all cases, readable binary sequences were implemented through the use of 5a and 5b. However, robot-made sequences contain a thymine residue that comes from the use of thymine-loaded support 8. After reaching the desired sequence and length, the 2-cyanoethyl phosphate protective groups (x) and the ester linker were cleaved (xii) (Figure 1d). All the digital polymers were characterized by negative mode high resolution electrospray mass spectrometry (ESI-HR-MS) (Table 1 and Figures S1-S8). In all cases, the targeted structures were detected as deprotonated molecules under different charge states. Yet, it should be noted that the polymers prepared manually on 7a and 7b (P1-P4 in Table 1) exhibited impurities and slight polydispersity due to missed coupling steps, whereas those prepared on 8 using the automated synthesizer (P5-P8 in Table 1) were monodisperse. This is probably due to the fact that the automated procedure employs two capping states per cycle that were not performed during the manual synthesis of short oligomers. Moreover, it should be remarked that conventional inverse-phase chromatography purification cannot be applied to these sequence-coded precursors because of the apolar character of the TIPS-protecting groups.

	Supp.	Sequence	Yield	m/z_{th}	m/z _{exp}
P1	7a	01100110 ^[b-c]	53%	966.1367 ^[d]	966.1375
P2	7a	0110 ^[b-c]	73%	708.8620 ^[e]	708.8626
P3	7b	1001 ^[b-c]	68%	708.8620 ^[e]	708.8638
P4	7b	1001 ^[b]	69%	708.8620 ^[e]	708.8628
P5	8	T 01000001	54%	1288.9641 ^[e]	1288.9659
P6	8	T 01001110 01001011	70%	880.4022 ^[f]	880.4030
P7	8	T 01000001 01000001 01000011	49%	945.2324 ^[g]	945.2326
P8	8	T 01001000 01000001 01010000 01010000 01011001 0100000 01010000 01001000 0100000		1079.0947 ^[h]	1079.0873 ^[i]

Table 1. ESI-HR-MS characterization of the sequence-coded precursors.^[a]

^[a] The numbers 0 and 1 denote monomers **5a** and **5b**, respectively. ^[b] Contains missed-steps. ^[c] Contains iodo-acetylene units. ^[d] [M-3H]³⁻. ^[e] [M-2H]²⁻. ^[f] [M-7H]⁷⁻. ^[g] [M-8H]⁸⁻. ^[h] [M-20H]²⁰⁻. ^[i] Lack of standards at the appropriate charge state accounts for the larger error (- 7 ppm) associated with measurement of this highly charged oligomer. All ions were measured at the isotopic maximum.



Figure 2. ESI-HR-MS spectrum (zoom of the 950-1200 m/z range) recorded for the sequence-coded precursor **P1** *with the sequence 01100110, detected as a triply deprotonated molecule (m/z 966.2) together with iodinated species.*

For some samples prepared on polystyrene resins, ESI-HR-MS also evidenced the presence of several species with masses corresponding to increments of 126 Da higher than the expected mass of the poly(phosphodiester)s (Figure 2 and Figures S1-S3). Such signals are most probably due to the fact that some terminal alkynes get randomly iodinated during the iterative synthesis. Since this side-reaction is not observed for polymers prepared on **8**, a logical explanation would be that traces of iodine used in *(ix)* stay entrapped in the resins despite extensive washing. These traces may accumulate and react with terminal alkynes when ammonia is used in step (*xii*).^[16] Even if acetylene iodination does not prevent post-modification by alkyne-azide cycloaddition,^[17] the concomitant use of both terminal alkynes and iodo-alkynes may lead to heterogeneously-modified, thus unreadable, polymers. In order to avoid this side-reaction, a sodium thiosulfate wash (*xi*) was conducted to reduce residual iodine before performing (*xii*) (Figure 1d). Sample **P4**, which was prepared following that procedure, did not exhibit any iodinated species in ESI-HR-MS (Figure S4), whereas comparable sample **P3**, which was cleaved in the presence of iodine, contained iodo-acetylene units (Figure S3). Altogether, the results of Table 1 suggest that any binary sequence can be synthesized using **5a** and **5b**.

Sequential CuAAC modification was then performed on some non-iodinated sequence-defined precursors. Although modification steps (*xiii*), (*xiv*) and (*xv*) could be potentially directly performed on solid-supports, step (*xiv*) may interfere with the base-labile linkers used in the present work. Therefore, sequential modification of the cleaved polymers was investigated in solution. Figure 3 shows for example the characterization of the post-polymerization modification of resin-made polymer **P4**. As displayed in Figure 1e, the terminal alkynes of the resin-made precursors were first reacted by CuAAC with a model azide, namely 1-azido-2-(2-(2-methoxyethoxy)ethoxy)ethane **10** (*xiii*). This reaction was performed in a mixture of DMSO and *tert*-butanol using Cu(I)Br as copper source, tris(benzyltriazolyl-methyl)amine (TBTA) as ligand following conditions previously optimized by Carrell and coworkers for DNA modification.^[15b] The resulting mono-modified polymer **P4**' was purified by dialysis in methanol and characterized by ESI-HR-MS (Figure S9) and ¹H NMR (Figure 3c). Both methods evidenced successful

CuAAC modification. In particular, quantitative modification was supported by the complete disappearance of terminal alkyne protons signals in the ¹H NMR spectrum at 2.3 ppm and the apparition of new peaks due to triazole proton (8.1 ppm), methylene protons adjacent to the triazole rings (4.5 ppm) and 2.8 ppm) and methylenoxy protons (3.75-3.4 ppm). Afterwards, the TIPS protecting groups of the 1units in P4' were cleaved with TBAF in THF solution (xiv) to obtain P4" (Figure S10). Quantitative deprotection was evidenced by the complete disappearance of the triisopropylsilyl protons at 1 ppm in 1 H NMR and the re-appearance of terminal alkyne protons at 2.3 ppm (Figure 3c). The second CuAAC azidoethoxy)ethoxy)ethoxy)ethyl)-2,2,2-trifluoroacetamide 11 as a model azide. After methanol dialysis, the fully modified digitally-encoded polymer P4" was characterized by ESI-HR-MS (Figure 3a), ¹⁹F NMR (Figure 3b) and ¹H NMR (Figure 3c). The latter technique evidenced the quantitative disappearance of terminal alkyne protons and a significant increase of the intensity of the methylenoxy protons broad peak. Successful modification was also confirmed by the appearance of a signal due to trifluoroacetamide moieties at - 77 ppm in ¹⁹F NMR. Perhaps more importantly, ESI-HR-MS evidenced the formation of polymer P4" in which all 0 and 1 units were quantitatively modified in steps (xiii) and (xv), respectively. In addition, incompletely-modified sequences could not be detected, thus suggesting that the stepwise CuAAC protocol is valid for modification of information-containing macromolecules.



Figure 3. Characterization of the binary-coded poly(phosphodiester)s **P4** at different stages of its modification: after solid-phase synthesis (light grey), after CuAAC transformation of the 0-units **P4'** (blue), after TIPS-removal on the 1-units **P4''** (dark grey) and after CuAAC modification of the 1-units **P4'''** (red). (a) ESI-HR-MS spectrum of the fully modified polymer **P4'''**. (b) ¹⁹F NMR spectrum of the fully modified polymer **P4'''**. (b) ¹⁹F NMR spectrum of the fully modified polymer **P4'''**. (c) ¹H NMR spectra recorded in methanol-d₄ for **P4**, **P4'**, **P4''**, and **P4'''**. The displayed regions are 8.5-7.5 ppm and 4.7-0.5 ppm. The cut region 7.5-4.7 ppm contains the water peak of the deuterated solvent.

A slightly modified protocol was used for the modification of polymers prepared on 8 that were typically obtained in smaller quantities that those prepared on 7a and 7b. The first CuAAC step was not performed in solution but directly on the controlled-pore glass cartridge. Subsequently, the mono-modified polymer was cleaved from the support and TIPS deprotection and the second CuAAC steps

were performed in solution. Following this methodology, sample **P7** with sequence T 01010000 01010000 01000100 was modified with azides **10** and **11**. Figures S11 and S12 show the ESI-HR-MS and ¹H NMR characterization of the intermediates **P7'** and **P7''**. Ultimately, Figure S13 shows the ¹H NMR of the final polymer **P7'''**. The complete disappearance of terminal alkyne protons in that spectrum confirmed the successful stepwise CuAAC modification of the digital polymer

In summary, stepwise CuAAC modification is an efficient strategy for the side-chain modification of digitally-encoded polymers. Importantly, all modification steps studied herein appeared to be quantitative, thus allowing synthesis of defect-free readable sequences. Consequently, this work opens up interesting avenues for the design of "universally" readable non-natural information-containing macromolecules.

Acknowledgements

J.F.L. thanks the H2020 program of the European Union (project Euro-Sequences, H2020-MSCA-ITN-2014, grant agreement n°642083), the European Research Council (ERC Proofs-of-Concept project Sequence Barcodes, grant agreement n°680097) and the CNRS for financial support. The PhD position of N.F.K. is supported by the ITN Euro-Sequences. The manager position of A.A.O. is supported by the ERC.

Keywords: Sequence-controlled polymers • information-containing macromolecules • digitally-encoded polymers • polymer modification • solid-phase synthesis

- [1] a) J.-F. Lutz, M. Ouchi, D. R. Liu, M. Sawamoto, *Science* 2013, 341, 1238149; b) H. Colquhoun, J.-F. Lutz, *Nat. Chem.* 2014, 6, 455-456; c) J.-F. Lutz, *Macromolecules* 2015, 48, 4759-4767.
- [2] a) G. M. Church, Y. Gao, S. Kosuri, *Science* 2012, *337*, 1628-1628; b) N. Goldman, et al., *Nature* 2013, *494*, 77-80; c) C. Mayer, G. R. McInroy, P. Murat, P. Van Delft, S. Balasubramanian, *Angew. Chem., Int. Ed.* 2016, *55*, 11144-11148; d) Y. Erlich, D. Zielinski, *Science* 2017, *355*, 950.
- [3] a) T. T. Trinh, L. Oswald, D. Chan-Seng, J.-F. Lutz, *Macromol. Rapid Commun.* 2014, 35, 141-145;
 b) T. T. Trinh, L. Oswald, D. Chan-Seng, L. Charles, J.-F. Lutz, *Chem. Eur. J.* 2015, 21, 11961-11965.
- [4] a) A. Al Ouahabi, L. Charles, J.-F. Lutz, J. Am. Chem. Soc. 2015, 137, 5629-5635; b) A. Al Ouahabi, M. Kotera, L. Charles, J.-F. Lutz, ACS Macro Lett. 2015, 1077-1080.
- [5] G. Cavallo, A. Al Ouahabi, L. Oswald, L. Charles, J.-F. Lutz, J. Am. Chem. Soc. 2016, 138, 9417-9420.
- [6] a) R. K. Roy, et al., *Nat. Commun.* 2015, 6, 7237; b) R. K. Roy, C. Laure, D. Fischer-Krauser, L. Charles, J.-F. Lutz, *Chem. Commun.* 2015, 51, 15677-15680; c) C. Laure, D. Karamessini, O. Milenkovic, L. Charles, J.-F. Lutz, *Angew. Chem., Int. Ed.* 2016, 55, 10722-10725.
- [7] U. S. Gunay, et al., *Chem* **2016**, *1*, 114-126.
- [8] V. Zhirnov, R. M. Zadegan, G. S. Sandhu, G. M. Church, W. L. Hughes, Nat. Mater. 2016, 15, 366-370.
- [9] R. N. Grass, R. Heckel, M. Puddu, D. Paunescu, W. J. Stark, Angew. Chem., Int. Ed. 2015, 54, 2552-2555.
- [10] a) D. Paunescu, W. J. Stark, R. N. Grass, *Powder Technol.* 2016, 291, 344-350; b) D. Karamessini, B. E. Petit, M. Bouquey, L. Charles, J.-F. Lutz, *Adv. Funct. Mater.* 2017, 27, 1604595.
- [11]H. Mutlu, J.-F. Lutz, Angew. Chem., Int. Ed. 2014, 53, 13010-13019.

- [12] a) D. Branton, et al., Nat Biotech 2008, 26, 1146-1153; b) Y.-L. Ying, J. Zhang, R. Gao, Y.-T. Long, Angew. Chem., Int. Ed. 2013, 52, 13154-13161.
- [13] a) P. Espeel, et al., Angew. Chem., Int. Ed. 2013, 52, 13261-13264; b) S. C. Solleder, M. A. R. Meier, Angew. Chem., Int. Ed. 2014, 53, 711-714; c) S. C. Solleder, D. Zengel, K. S. Wetzel, M. A. R. Meier, Angew. Chem., Int. Ed. 2016, 55, 1204-1207; d) J. W. Grate, K.-F. Mo, M. D. Daily, Angew. Chem., Int. Ed. 2016, 55, 3925-3930; e) R. L. Kanasty, et al., Angew. Chem., Int. Ed. 2016, 55, 9529-9533; f) D. Kong, Y. Lei, W. Yeung, R. Hili, Angew. Chem., Int. Ed. 2016, 55, 13164-13168.
- [14]a) R. Häner, F. Garo, D. Wenger, V. L. Malinovskii, J. Am. Chem. Soc. 2010, 132, 7466-7471; b) T. G. W. Edwardson, K. M. M. Carneiro, C. J. Serpell, H. F. Sleiman, Angew. Chem., Int. Ed. 2014, 53, 4567-4571; c) C. J. Serpell, T. G. W. Edwardson, P. Chidchob, K. M. M. Carneiro, H. F. Sleiman, J. Am. Chem. Soc. 2014, 136, 15767-15774; d) D. de Rochambeau, et al., Polym. Chem. 2016, 7, 4998-5003.
- [15]a) J.-F. Lutz, Angew. Chem., Int. Ed. 2007, 46, 1018-1025; b) P. M. E. Gramlich, S. Warncke, J. Gierlich, T. Carell, Angew. Chem., Int. Ed. 2008, 47, 3442-3444; c) P. M. E. Gramlich, C. T. Wirges, A. Manetto, T. Carell, Angew. Chem., Int. Ed. 2008, 47, 8350-8358; d) N. Baradel, S. Fort, S. Halila, N. Badi, J.-F. Lutz, Angew. Chem., Int. Ed. 2013, 52, 2335-2339.
- [16] T. H. Vaughn, J. A. Nieuwland, J. Am. Chem. Soc. 1933, 55, 2150-2153.
- [17]a) J. E. Hein, J. C. Tripp, L. B. Krasnova, K. B. Sharpless, V. V. Fokin, *Angew. Chem., Int. Ed.* 2009, 48, 8018-8021; b) D. N. Barsoum, N. Okashah, X. Zhang, L. Zhu, *J. Org. Chem.* 2015, 80, 9542-9551.

COMMUNICATION



In digital polymers, binary information is written using two comonomers, with slightly different molecular structures, that are defined as 0 and 1 bits. Here, we describe a simple stepwise modification method that permits to tune the molecular structure of the bits after polymerization.

Niklas Felix König, Abdelaziz Al Ouahabi, Salomé Poyer, Laurence Charles,* and Jean-François Lutz*

Page No. – Page No.

A simple post-polymerization modification method for controlling sidechain information in digital polymers