



# Automated Synthesis of Monodisperse Oligomers, Featuring Sequence Control and Tailored Functionalization

Steven Martens, Jos Van den Begin, Annemieke Madder, Filip E. Du Prez,\* and Pieter Espeel

Department of Organic and Macromolecular Chemistry, Ghent University, Krijgslaan 281 S4, 9000 Ghent, Belgium

## Supporting Information

**ABSTRACT:** Long, multifunctional sequence-defined oligomers were obtained on solid support from a protecting-group-free two-step iterative protocol, based on the inherent reactivity of a readily available molecule containing an isocyanate and a thiolactone. Aminolysis of the latter entity with an amino alcohol liberates a thiol that reacts with an acrylate or acrylamide, present in the same medium. Subsequently, a new thiolactone can be reinstated by means of an  $\alpha$ -isocyanato- $\gamma$ -thiolactone. Different acrylic compounds were used to incorporate diverse functionalities in the oligomers, which were built up to the level of decamers. The reaction conditions were closely monitored in order to fine-tune the applied strategy as well as facilitate the translation to an automated protocol.

Polymer chemists have made great progress in the last decades in the preparation of functionalized and complex polymer architectures with defined structure–property relationships, yet these structures do not reach the same potential as their natural counterparts (DNA, RNA and proteins).<sup>1–3</sup> As a result, in the past decade, efforts have increased to achieve control over the primary structure of macromolecules in hopes of creating a stronger and/or unique correlation between properties and structure.<sup>4</sup> Many different methods have been explored for the synthesis of such sequence-controlled macromolecules.<sup>5–25</sup>

(Semi)automated protocols, which facilitate iterative solid phase synthesis (SPS) of oligonucleotides<sup>26</sup> and oligopept(o)-ides,<sup>27,28</sup> have gained much attention for the preparation of defined/controlled sequences.<sup>29–31</sup> SPS revolutionized the fields of peptide and oligonucleotide chemistry.<sup>26,27</sup> This versatile method enabled the fast synthesis and isolation of oligomers. The efficient automation of SPS allowed even faster synthesis and parallel approaches to build a wide variety of natural oligomers, nowadays even up to a ton scale.<sup>32</sup> Automated synthesis of oligopeptoids has been achieved via an efficient submonomer strategy, which does not require any protecting groups.<sup>7,28</sup> Given these successes, adapted protocols have been employed for the synthesis of sequence-defined polymers using phosphoramidite chemistry, which is classically used for the synthesis of poly- and oligonucleotides on DNA synthesizers.<sup>30</sup> Lutz et al. applied this chemistry to make sequences with high DP (>100), using a limited monomer alphabet consisting of two custom-made monomers, with the aim to produce binary code type (1/0) sequences.

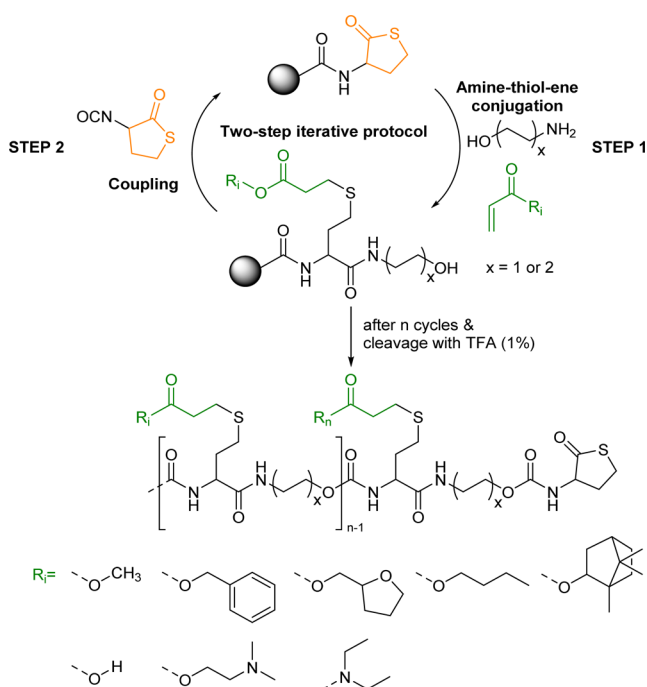
This paper aims to present a robust and versatile protocol for the synthesis of sequence-defined multifunctional oligomers based on thiolactone (Tla) chemistry<sup>23,33,34</sup> and its successful automation. Recently, we have shown that immobilizing a Tla unit on a solid support enables chain extension after on-resin aminolysis.<sup>33</sup> In this previous submonomer strategy, the monomer alphabet, encompassing amines and thiolactone units, is broad and protecting groups are not necessary. However, the reported approach was limited to the synthesis on very small scale of tetramers with moderate crude purity. As a result of disulfide formation during on-resin aminolysis, treatment with phosphines as reducing agents was necessary, promoting the accumulation of side products. Therefore, to allow for the preparation of much longer, highly functional sequences, with the additional possibility to translate it to an automated protocol, a new protecting-group-free strategy was developed. Remarkable features are much higher versatility in side chain functionality and backbone composition, and elevated crude purity.

The implementation of the nucleophilic amine-thiol-ene conjugation<sup>35</sup> is a breakthrough development for Tla-based SPS. In the first step of the synthetic cycle (STEP 1, Scheme 1), the immobilized Tla is selectively opened through the primary amine function of an amino alcohol (ethanolamine or 4-amino-1-butanol), releasing a thiol that subsequently reacts with an acrylate or acrylamide through a nucleophilic thiol–ene reaction. A wide set of functionalities can be introduced by using abundantly available acrylics whereas the nature of the backbone can be varied through the selected amino alcohol. Moreover, it is anticipated that disulfide formation can be completely avoided in this mild additive-free approach.<sup>33,35</sup> In the next step, chain extension is performed via the reaction of the readily available and stable  $\alpha$ -isocyanato- $\gamma$ -Tla (Scheme S2) with the remaining alcohol function (STEP 2, Scheme 1).

This two-step protocol can be repeated until a unique oligomer with targeted length, sequence and functionality is obtained. In the present study, the prepared structures feature amides and urethanes in the backbone and ester- and amide-linked side chains. The  $\alpha$ - and  $\omega$ -end groups of these sequences depend on two parameters: the functional group needed to initially immobilize the Tla moiety on the solid phase ( $\alpha$ -terminus) and the stage at which the iterative protocol is terminated at the other extremity ( $\omega$ -terminus). There are two possibilities as far as the latter is concerned. The sequence can be terminated after step 1, which results in an alcohol, or after step 2, which results in a Tla. The nature of the  $\alpha$ -end group, on the

Received: July 10, 2016

Published: October 14, 2016

**Scheme 1. Two-Step Iterative SPS Yielding Monodisperse, Multifunctional Oligomers Based on Thiolactone Chemistry<sup>a</sup>**


<sup>a</sup>The immobilized thiolactone group is opened with an amino alcohol and subsequently reacted with a functional acrylate or acrylamide (STEP 1) whereas  $\alpha$ -isocyanato- $\gamma$ -Tla serves to reinstate the Tla moiety (STEP 2).

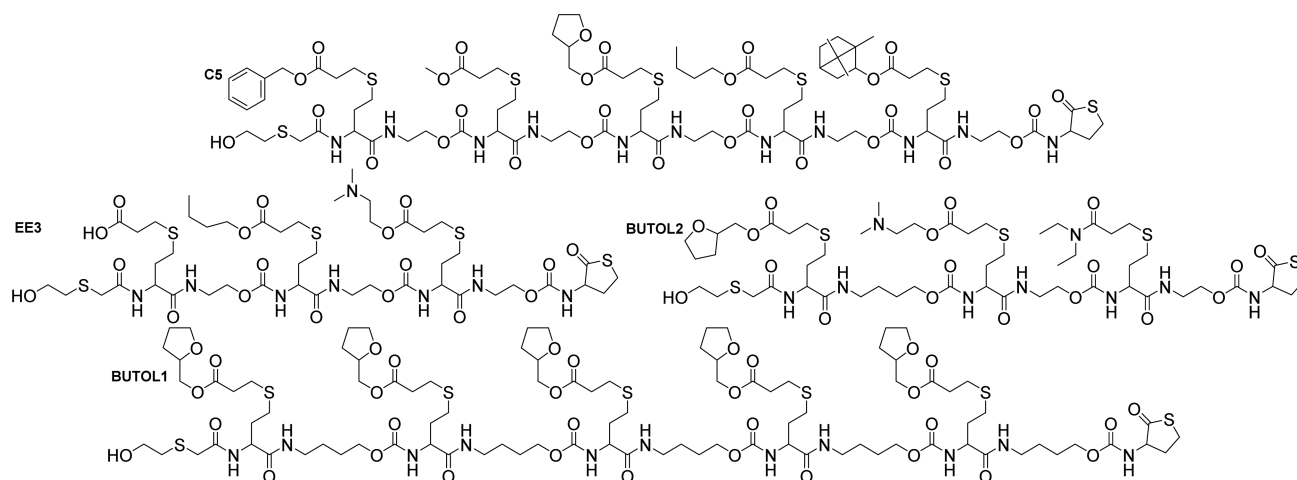
other hand, depends on the functionality used to connect the Tla on the cross-linked polystyrene resin. In addition to a Tla-containing acid<sup>33</sup> as linker, a Tla-containing alcohol was evaluated. The loading of the hydroxyl-functionalized Tla linker was 0.85 mmol Tla units per gram resin (53% loading efficiency, Scheme S1 and Figure S4).<sup>36</sup>

The reaction conditions were carefully investigated to prepare sequence-defined oligomers efficiently and to transfer subsequently this protocol to an automated approach. Optimization parameters included the reaction stoichiometry, the choice of

solvent and the reaction times used for each step. The amine/acrylate ratio used in the first step is important, as this influences the extent of disulfide formation. A ratio of one to two was found to prevent this side reaction. Next, the chosen solvent must promote the reaction between the three reaction partners as well as ensure the swelling of the resin. Although DMF and  $\text{CH}_2\text{Cl}_2$  are commonly applied in SPS, they would cause difficulties in the present protocol. Indeed, DMF promotes the side reaction between an acrylate and an amine,<sup>37</sup> while we reported earlier that  $\text{CH}_2\text{Cl}_2$  can react with the thiol formed after aminolysis.<sup>38</sup> In  $\text{CHCl}_3$ , the occurrence of aza-Michael addition was found to be less problematic. The last important parameter for Step 1 is reaction time. A series of measurements demonstrated that the reaction is completed after 15 min (Figure S5). However, the first step is performed twice to ensure full conversion for longer sequences.

The second step of the iterative protocol consists of the reaction between the  $\alpha$ -isocyanato- $\gamma$ -Tla and an alcohol incorporated in the first step via aminolysis of the Tla. The reaction is performed with 10 equiv of the isocyanate and 0.025 equiv of dibutyltin dilaurate (DBTL). Subsequently, three dry solvents were tested: DMF, ethyl acetate and  $\text{CHCl}_3$  (Figures S6–S8). This screening revealed that the reaction is completed within 30 min when using  $\text{CHCl}_3$ , after 2 h for ethyl acetate and still not finished after 8 h for DMF. In the optimized protocol, the chain extension in  $\text{CHCl}_3$  is, however, made to last 1 h to ensure full conversion.

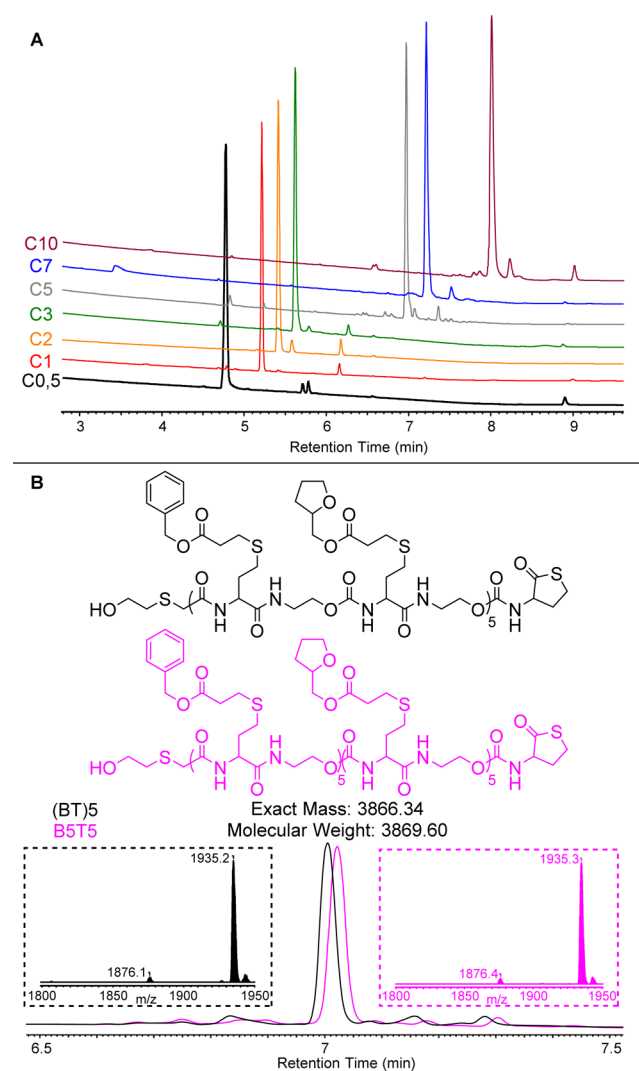
Subsequently, the protocol was closely followed while first making a trimer sequence using benzyl, methyl and tetrahydrofurfuryl acrylate (C0,5-C3,  $C_x$  = performed cycles). HR-MS, LC-ESI-MS and NMR spectra were recorded to verify sample purity and incorporation of side chain residues (Tables S1–S5 and Figures S9–S16). The LC traces of crude reaction mixtures consistently displayed one major peak that corresponded with the expected product, demonstrating that the products display no deletions (Figure 2A) and that 5% of the hydroxyl end group is transformed into the corresponding TFA ester (Figures S9 and S11). Next, two extra functionalities were incorporated using butyl and isobornyl acrylate, yielding a highly pure pentamer (C5) (Figure 1, Figures S17–S18 and Table S6). Continuing with this pentamer, a heptamer (C7) was made by incorporating



**Figure 1.** Overview of different sequences, illustrating the diversity of incorporated acrylics. C5 and EE3 were made from ethanolamine and various acrylates, while BUTOL1 and BUTOL2 are examples of a pentamer and a trimer in which 4-amino-1-butanol is incorporated in the backbone. For BUTOL2, there is also an acrylamide incorporated in the side chain.

additional benzyl and tetrahydrofurfuryl residues (Figures S19–S20 and Table S7).

Next, five decamers were prepared, the first one (C10) by starting from the heptamer with the incorporation of methyl, butyl and isobornyl groups. The second (T10) one was a homodecamer with only tetrahydrofurfuryl groups as functionality. The third (B5T5) and fourth ((BT)5) were built with the latter acrylate and benzyl acrylate to make two decamers, respectively consisting of two pentamer blocks and an alternating structure. The small retention time difference in the LC traces of both decamers originates from their different behavior toward the column (Figure 2B). Finally, the fifth decamer had a sequence similar to the first one, but the acid linker was applied to incorporate a different  $\alpha$ -end group (Figures S21–S30 and Tables S8–S12).



**Figure 2.** Overlay of HPLC traces of sequences from an opened Tla (C0,5) to a decamer (C10) ( $C_x$  = performed cycles) (A) and of two decamers with the same mass but with a different sequence (B).

To emphasize the almost unlimited choice for side chain functionalities, two trimers (EE3 and DMAE3) were made that contained a *n*-butyl group, an acid and a tertiary amine in the side chains (Figure 1). The acid and amine functionalities could be incorporated by using respectively 1-ethoxyethyl and 2-(dimethylamino)ethyl acrylate.<sup>39</sup> Acid cleavage from the resin

also removed the ethoxyethyl group, resulting in an acid residue. To extend further the functionality scope from acrylates to acrylamides, *N,N*-diethylacrylamide was also incorporated in a trimer (DEAA3) (Figures S31–S36 and Tables S13–S15).

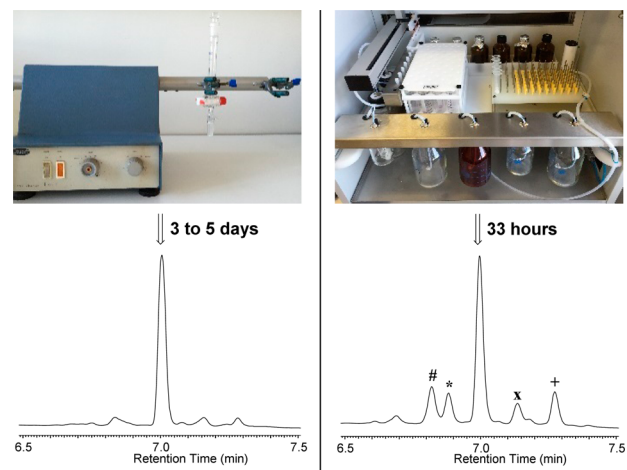
Finally, to illustrate the high versatility, also the backbone has been modified by using 4-amino-1-butanol instead of ethanolamine. In previous research,<sup>35</sup> the effect of the amine on the kinetics of aminolysis was investigated and taken into account in the protocol (Figure S37). Although the first synthesis of a trimer indicated a deletion (Figure S38), the protocol was modified for this specific aminolysis to four times 15 min. With this adjustment, a high purity pentamer (BUTOL1) was made with only tetrahydrofurfuryl groups as side chain functionalities and a trimer (BUTOL2) with three side chain functionalities and based on two alcohol amines (Figure 1, Figures S39–S42 and Tables S16–17).

The next target was the automatization of the protocol to facilitate the iterative work and synthesize several sequences in parallel, thus significantly reducing the overall process time. To achieve this goal, a peptide synthesizer was adapted to our needs and all sequences were prepared in a single nonstop run. In comparison with a standard synthesizer, extra bottles were installed for washing solvents and closed vessels were required for the inert storage of the  $\alpha$ -isocyanato- $\gamma$ -Tla, DBTL and dried  $\text{CHCl}_3$ . This last adaptation was absolutely necessary because of the water-sensitive urethane synthesis during each second step of the protocol (see movie and picture with additional information in SI).

First, a trimer (R1) was made to adjust instrument settings (Figure S31). Next, six random hexamers (R2–R7) were synthesized in parallel using benzyl, butyl and tetrahydrofurfuryl acrylate. This second screening was equally very effective, since every hexamer was characterized by a high level of purity (Figures S32–S37). A nonamer (R8) was also prepared with a high abundance, but traces of deletions could be noticed (Figure S38).

The final test was a comparison between (BT)5 decamers, made in a manual and in an automated way (R9) (Figure 3).

The LC trace of the robot-made decamer indicated that the decamer was formed and displayed the highest abundance, but it



**Figure 3.** Comparison between conventional (left) and automated (peptide synthesizer) synthesis (right) of decamer (BT)5. Peak # represents a nonamer missing a benzyl residue while peak \* is a decamer in the absence of the terminal Tla unit. Peaks x and + are TFA-ester adducts of \* and the main compound.

also shows the presence of a nonamer (#), missing a benzyl functionality, and a decamer without the presence of a Tla unit (\*). In comparison with the storage vessels, the reactors for the automated synthesis are open, which can cause problems during the chain extension, because isocyanates are sensitive to humidity in the air. The formation of a white precipitate, which corresponds to dithiolactone urea (Figures S40–S41), can confirm this air sensitivity. On the other hand, it should be emphasized that the time for the synthesis of a decamer is obviously much shorter with the automated protocol ( $\pm 33$  h versus 3 to 5 days).

In conclusion, a robust chemical platform was developed to perform the protecting-group-free SPS of multifunctional sequence-defined oligomers using Tla chemistry. Various functional groups could be incorporated by means of acrylics during aminolysis. Furthermore, both the nature of the backbone and the end-groups at both termini has been adapted. Functionalized oligomers, up to decamers, could be prepared with high purity. Finally, the protocol could be efficiently translated to an automated approach on an adapted peptide synthesizer for an accelerated synthesis of libraries of sequence-defined oligomers.

In the near future, we hope to be able to show that the available racemic structures can be applied in the field of molecular data storage. The option to store information both in the multifunctional side and/or main chains can result in a quite significant progress in terms of data capacity on a single oligomer. Future work will also focus on the stereocontrol of the applied thiolactone unit, because this is expected to have an impact on the folding behavior of the sequence-defined structures.<sup>40</sup>

## ■ ASSOCIATED CONTENT

### 📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b07120.

Experimental procedures; synthesis of different Tla's and sequences; kinetic study; NMR and LC-ESI-MS data (PDF)

Adapted peptide synthesizer setup (AVI)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*Filip.DuPrez@ugent.be

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

Jan Goeman is acknowledged for the MS-measurements. S.M. is indebted to the Research Foundation-Flanders (FWO) for a position as research assistant. F.D.P. and A.M. gratefully acknowledge project support from the Research Foundation-Flanders (FWO), grant agreement G.0204.15N). F.D.P. acknowledges the Belgian Program on Interuniversity Attraction Poles initiated by the Belgian State, the Prime Minister's office (P7/05) and the H2020 program of the European Union (project Euro-Sequences, H2020-MSCA-ITN-2014, grant agreement No 642083).

## ■ REFERENCES

(1) Hawker, C. J.; Wooley, K. L. *Science* **2005**, *309*, 1200.  
(2) Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry*; W.H. Freeman: Basingstoke, 2012.

- (3) Espeel, P.; Du Prez, F. E. *Macromolecules* **2015**, *48*, 2.  
(4) Lutz, J. F.; Ouchi, M.; Liu, D. R.; Sawamoto, M. *Science* **2013**, *341*, 1238149.  
(5) Hartmann, L. *Macromol. Chem. Phys.* **2011**, *212*, 8.  
(6) Al Ouahabi, A.; Charles, L.; Lutz, J. F. *J. Am. Chem. Soc.* **2015**, *137*, 5629.  
(7) Zuckermann, R. N.; Kerr, J. M.; Kent, S. B. H.; Moos, W. H. *J. Am. Chem. Soc.* **1992**, *114*, 10646.  
(8) Meszynska, A.; Badi, N.; Borner, H. G.; Lutz, J. F. *Chem. Commun.* **2012**, *48*, 3887.  
(9) Amrane, M. L.; Chouikhi, D.; Badi, N.; Lutz, J. F. *Macromol. Chem. Phys.* **2014**, *215*, 1984.  
(10) Vandenberg, J.; Reekmans, G.; Adriaensens, P.; Junkers, T. *Chem. Sci.* **2015**, *6*, 5753.  
(11) De Bo, G.; Kuschel, S.; Leigh, D. A.; Lewandowski, B.; Pappmeyer, M.; Ward, J. W. *J. Am. Chem. Soc.* **2014**, *136*, 5811.  
(12) Lewandowski, B.; De Bo, G.; Ward, J. W.; Pappmeyer, M.; Kuschel, S.; Aldegunde, M. J.; Gramlich, P. M. E.; Heckmann, D.; Goldup, S. M.; D'Souza, D. M.; Fernandes, A. E.; Leigh, D. A. *Science* **2013**, *339*, 189.  
(13) Solleder, S. C.; Meier, M. A. R. *Angew. Chem., Int. Ed.* **2014**, *53*, 711.  
(14) Solleder, S. C.; Wetzel, K. S.; Meier, M. A. R. *Polym. Chem.* **2015**, *6*, 3201.  
(15) Gody, G.; Maschmeyer, T.; Zetterlund, P. B.; Perrier, S. *Nat. Commun.* **2013**, *4*, 3505.  
(16) Srichan, S.; Mutlu, H.; Lutz, J. F. *Eur. Polym. J.* **2015**, *62*, 338.  
(17) Porel, M.; Alabi, C. A. *J. Am. Chem. Soc.* **2014**, *136*, 13162.  
(18) Porel, M.; Thornlow, D. N.; Phan, N. N.; Alabi, C. A. *Nat. Chem.* **2016**, *8*, 590.  
(19) Barnes, J. C.; Ehrlich, D. J. C.; Gao, A. X.; Leibfarth, F. A.; Jiang, Y.; Zhou, E.; Jamison, T. F.; Johnson, J. A. *Nat. Chem.* **2015**, *7*, 810.  
(20) van Genabeek, B.; de Waal, B. F. M.; Gosens, M. M. J.; Pitet, L. M.; Palmans, A. R. A.; Meijer, E. W. *J. Am. Chem. Soc.* **2016**, *138*, 4210.  
(21) Gutekunst, W. R.; Hawker, C. J. *J. Am. Chem. Soc.* **2015**, *137*, 8038.  
(22) Weiss, R. M.; Short, A. L.; Meyer, T. Y. *ACS Macro Lett.* **2015**, *4*, 1039.  
(23) Driessen, F.; Du Prez, F. E.; Espeel, P. *ACS Macro Lett.* **2015**, *4*, 616.  
(24) Pfeifer, S.; Lutz, J. F. *J. Am. Chem. Soc.* **2007**, *129*, 9542.  
(25) Ida, S.; Ouchi, M.; Sawamoto, M. *Macromol. Rapid Commun.* **2011**, *32*, 209.  
(26) Beaucage, S. L.; Iyer, R. P. *Tetrahedron* **1992**, *48*, 2223.  
(27) Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149.  
(28) Knight, A. S.; Zhou, E. Y.; Francis, M. B.; Zuckermann, R. N. *Adv. Mater.* **2015**, *27*, 5665.  
(29) Leibfarth, F. A.; Johnson, J. A.; Jamison, T. F. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 10617.  
(30) Al Ouahabi, A.; Kotera, M.; Charles, L.; Lutz, J.-F. *ACS Macro Lett.* **2015**, *4*, 1077.  
(31) Hartmann, L.; Krause, E.; Antonietti, M.; Borner, H. G. *Biomacromolecules* **2006**, *7*, 1239.  
(32) Bray, B. L. *Nat. Rev. Drug Discovery* **2003**, *2*, 587.  
(33) Espeel, P.; Carrette, L. L. G.; Bury, K.; Capenberghs, S.; Martins, J. C.; Du Prez, F. E.; Madder, A. *Angew. Chem., Int. Ed.* **2013**, *52*, 13261.  
(34) Espeel, P.; Du Prez, F. E. *Eur. Polym. J.* **2015**, *62*, 247.  
(35) Espeel, P.; Goethals, F.; Driessen, F.; Nguyen, L.-T. T.; Du Prez, F. E. *Polym. Chem.* **2013**, *4*, 2449.  
(36) Shankar, B. B.; Yang, D. Y.; Girton, S.; Ganguly, A. K. *Tetrahedron Lett.* **1998**, *39*, 2447.  
(37) Mather, B. D.; Viswanathan, K.; Miller, K. M.; Long, T. E. *Prog. Polym. Sci.* **2006**, *31*, 487.  
(38) Reinicke, S.; Espeel, P.; Stamenovic, M. M.; Du Prez, F. E. *Polym. Chem.* **2014**, *5*, 5461.  
(39) Van Camp, W.; Du Prez, F. E.; Bon, S. A. F. *Macromolecules* **2004**, *37*, 6673.  
(40) Zuckermann, R. N. *Biopolymers* **2011**, *96*, 545.