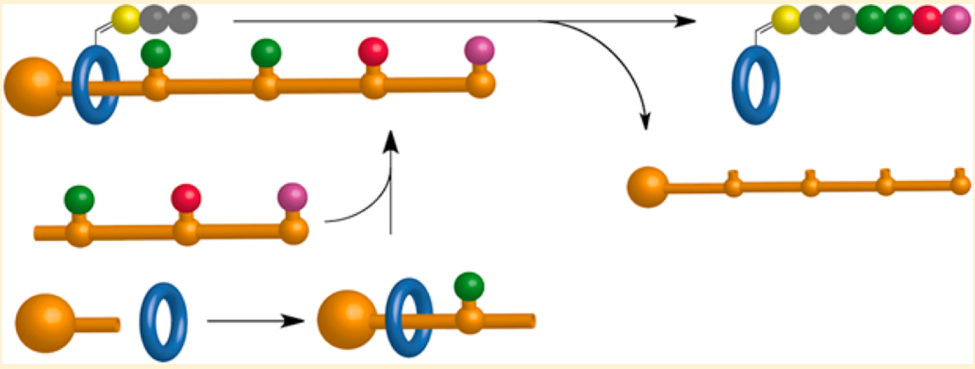
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

Efficient Assembly of Threaded Molecular Machines for Sequence-Specific Synthesis

Guillaume De Bo, Sonja Kuschel, David A. Leigh,* Bartosz Lewandowski, Marcus Papmeyer, and John W. Ward

School of Chemistry, University of Manchester, Oxford Road, Manchester, M13 9PL United Kingdom

Supporting Information



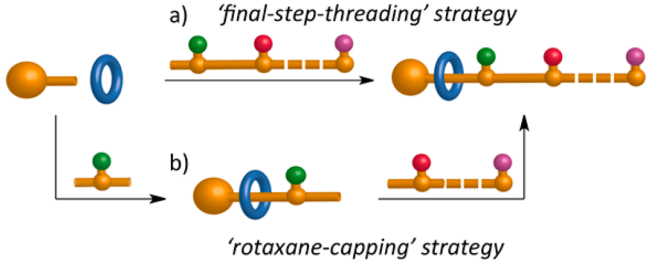
ABSTRACT: We report on an improved strategy for the preparation of artificial molecular machines that can pick up and assemble reactive groups in sequence by traveling along a track. In the new approach a preformed rotaxane synthon is attached to the end of an otherwise fully formed strand of building blocks. This "rotaxane-capping" protocol is significantly more efficient than the "final-step-threading" method employed previously and enables the synthesis of threaded molecular machines that operate on extended oligomer, and potentially polymer, tracks. The methodology is exemplified through the preparation of a machine that adds four amino acid building blocks from a strand in sequence, featuring up to 20-membered ring native chemical ligation transition states.

INTRODUCTION

Many biological molecular machines responsible for sequence-specific synthesis, including the ribosome[1] and various DNA polymerases,[2] travel along tracks in order to assemble reactive building blocks in a precise order.[3] An artificial rotaxane-based molecular machine[4] was recently described that successively adds three amino acids to a growing peptide chain in such a fashion.[5,6] The macrocycle moves along a thread derivatized with amino acid phenolate esters, picking up the groups that block its path and linking them together through successive native chemical ligation[7] (NCL) reactions to form a new peptide oligomer. The sequence of the product corresponds to the order of the building blocks on the original track.

The key reaction to form the interlocked machine architecture involves threading of the ring onto the full strand of building blocks (Scheme 1a).[5] Unfortunately, the active template[8,9] reaction employed to do this is low yielding and comes as the last step of a long synthetic route, limiting the utility of this approach for longer tracks. The low efficiency of the threading is probably due to the number of amide groups on the completed track that can sequester the Cu(I) catalyst from the pyridine group of the macrocycle used to promote the active template threading reaction.[9a] In order to produce machines that can operate on extended strands with more building blocks it is clearly necessary to find a more efficient strategy to assemble such machine-track conjugates. Here we report on a molecular peptide synthesizer containing four amino acids loaded onto the building block strand, prepared using a "rotaxane-capping" strategy (Schemes 1b). The operation of the molecular device yields an oligopeptide with four amino acids added from the track in the desired sequence (Scheme 3).

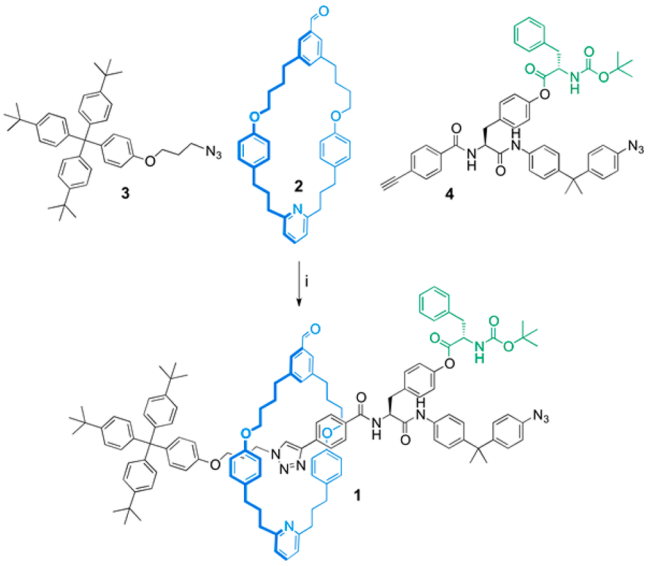
Scheme 1. Assembling Threaded Molecular Machines for Sequence-Specific Synthesis



Received: March 4, 2014
Published: March 28, 2014

   dx.doi.org/10.1021/ja5022415 | J. Am. Chem. Soc. 2014, 136, 5811−5814

Scheme 2. Synthesis of Rotaxane Synthon 1[a]



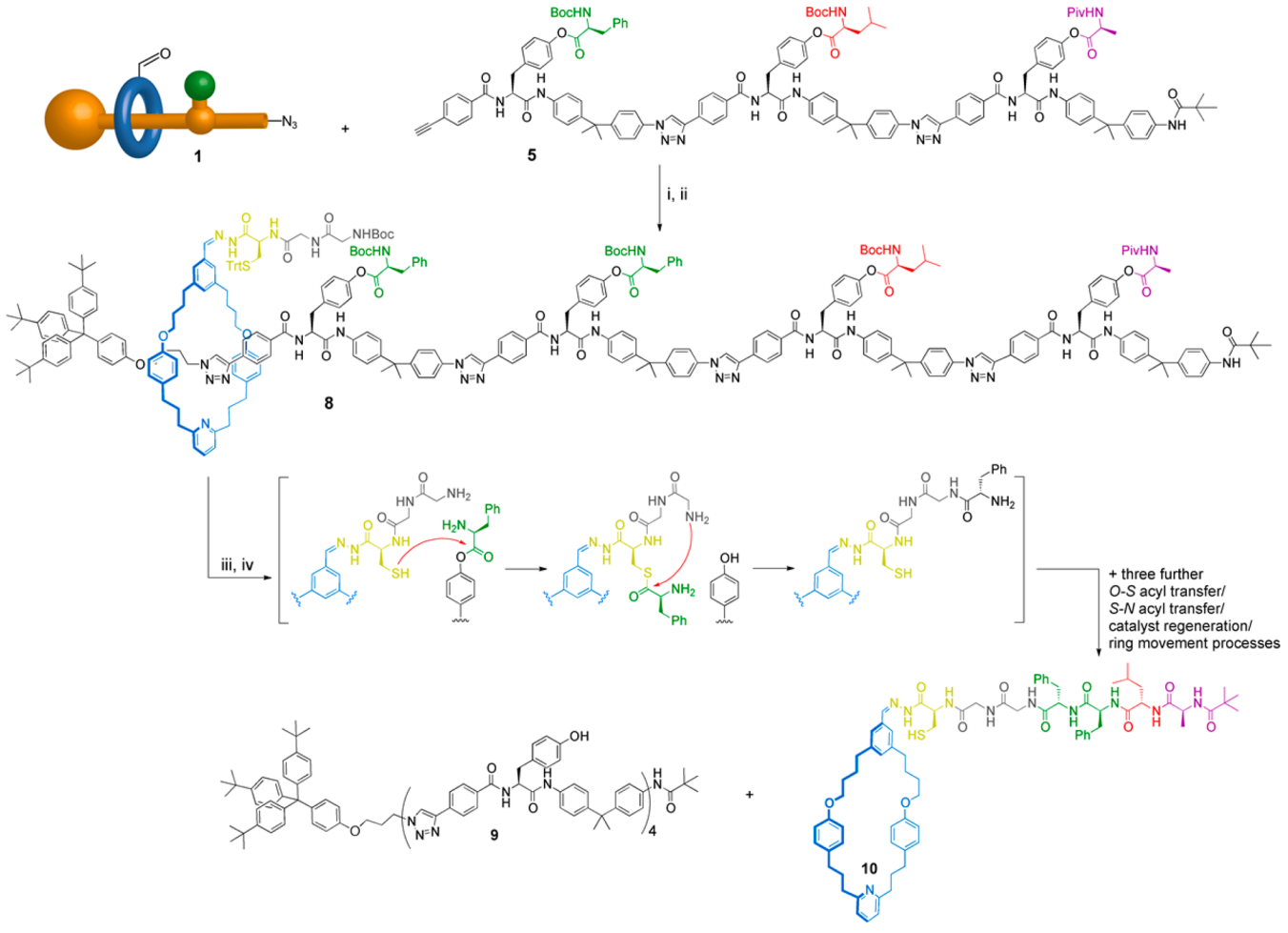
[a]Reagents and conditions: (i) **2** (2 equiv), **3** (10 equiv), **4** (1 equiv), $Cu(CH_3CN)_4.PF_6$ (0.5 equiv), 6:1 CH_2Cl_2/*t*-BuOH, RT, 24 h, 50%.

RESULTS AND DISCUSSION

A simple [2]rotaxane synthon, **1**, bearing no persistent recognition elements between the macrocycle and sites on the thread, which would retard free movement of the macrocycle toward the reactive building blocks, and a *N*-Boc-Phe group, which is sufficiently large to prevent dethreading, was prepared from macrocycle **2**, azide **3**, and alkyne **4** by an active template[9] Cu(I)-catalyzed alkyne−azide cycloaddition (CuAAC) reaction[10] (Scheme 2). Rotaxane **1** was isolated in 50% yield in a procedure that could be conveniently run on a scale that provided hundreds of milligrams of the threaded building block. Rotaxane **1** was subsequently attached to a molecular strand containing a further three amino acids in its structure, **5**, using a conventional CuAAC ligation to afford **6** (Scheme 3). The CysGlyGly "arm", responsible for removing and joining together the track amino acid groups, was then added (in protected form) to the extended machine assembly through hydrazone exchange of **6** with **7**, under aniline catalysis.[11]

The fully assembled molecular machine, **8**, had the acid-labile protecting groups removed (20% CF_3CO_2H in CH_2Cl_2; Scheme 2 (iii)) and was allowed to operate at 60 °C in 3:1 CH_3CN:$(CH_3)_2NCHO$, in the presence of iPr_2NEt

Scheme 3. Synthesis and Operation of Four-Barrier Molecular Machine 8[a]



[a]Reagents and conditions: (i) TentaGel-TBTA/$Cu(CH_3CN)_4{\cdot}PF_6$ (0.5 equiv), 4:1 CH_2Cl_2/*t*-BuOH, RT, 48 h, 70%. (ii) $PhNH_2$, BocGlyGlyCys(*S*-Trt)NHN=$CHC_6H_4OCH_3$ (**7**) in 3:1 dimethyl sulfoxide: 2-(*N*-morpholino)ethanesulfonic acid aqueous buffer (pH 6.0), 60 °C, 2 days, 90%. (iii) CF_3CO_2H (20% in CH_2Cl_2), RT, 2 h. (iv) $((CH_3)_2CH)_2NEt$, $(HO_2CCH_2CH_2)_3P$ in 3:1 CH_3CN:$(CH_3)_2NCOH$, 60 °C, microwave, 48 h, **9** (80%) and **10** (53%).

dx.doi.org/10.1021/ja5022415 | *J. Am. Chem. Soc.* 2014, 136, 5811−5814

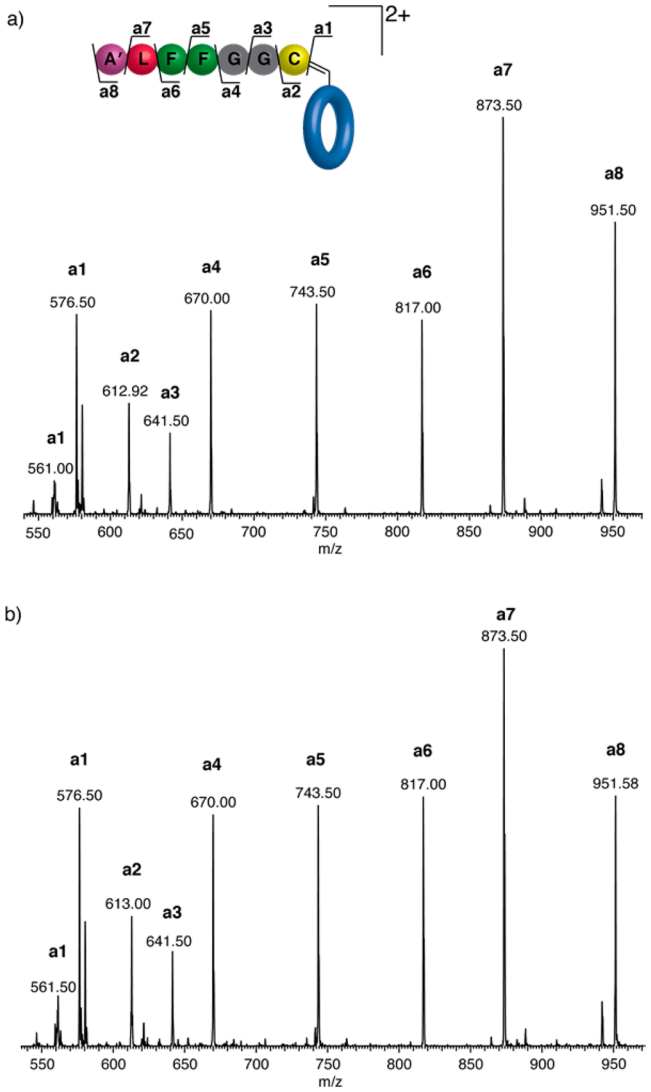


Figure 1. Tandem mass spectra of 2+ ions of *S*,*N*-acetal-derivatized macrocycles bearing the peptide sequence (Piv)-AlaLeuPhePheGlyGlyCys: (a) Product of the operation of molecular machine **8** in Scheme 2 (2+ ion isotope-selected m/z = 951.42). (b) The identical compound prepared unambiguously by conventional peptide synthesis (2+ ion isotope-selected m/z = 951.58). The m/z values observed correspond to the loss of peptide fragments at various points along the chain (**a7**, **a6**, **a5**, ... etc.), and therefore the peptide sequence can be read directly simply by considering the fragmentation from high m/z value to low. Fragment **a1** is observed as a singly charged adduct of the aldehyde macrocycle (m/z = 561) and the corresponding methanol or hydrazone adduct (m/z = 576).

(to promote NCL reactions[12]) and $(HO_2CCH_2CH_2)_3P$ (to reduce any cysteine disulfides formed *in situ* back to thiols[13]) (Scheme 3 (iv)). After 48 h mass spectrometry showed two major products present in the postoperation mixture which were isolated by preparative thin layer chromatography. One of the products was identified as the noninterlocked thread with four phenol groups, **9** (80% isolated yield); the other was the macrocycle bearing a peptide chain that mass spectrometry showed had a molecular mass (m/z = 1355 $[M + H]^+$) corresponding to the tripeptide of the original arm with four further amino acids added, one from each site of the thread, **10** (53% isolated yield).

The sequence of the newly formed heptapeptide was confirmed as AlaLeuPhePheGlyGlyCys by tandem mass spectrometry on an *S*,*N*-acetal derivative of the operation product (Figure 1a).[5a] Comparison of the fragmentation pattern of the isolated product with an authentic sample of the heptapeptide, prepared by conventional peptide synthesis, confirmed that the operation product had the amino acid groups in the anticipated order (Figure 1b). No other products featuring the macrocycle bearing other peptide sequences or missing amino acid residues were detected, indicating the high sequence specificity of the peptide synthesis by the molecular machine. The efficiency of even the fourth successive NCL reaction, featuring a 20-membered ring *S*,*N*-acyl transfer transition state, without the need for a significant increase in reaction time nor to optimize conditions, suggests that we are not yet approaching the limit of the number of amino acids that can be added through this mechanism.

CONCLUSIONS

The potential utility of artificial polymers made from precise sequences of building blocks is apparent from the diverse properties of the four controlled-sequence macromolecules exploited throughout biology (DNA, RNA, proteins, and oligosaccharides).[3] Currently, however, sequence-specific synthesis using unnatural building blocks or chemistries remains an unsolved problem in polymer science.[14] The facile connection of a prethreaded ring to the end of a strand, without leaving residual recognition elements between the macrocycle and track to hinder the rings movement, should greatly assist the development of artificial molecular machines for this purpose.[5b−d]

ASSOCIATED CONTENT

Supporting Information

Detailed descriptions of synthetic procedures, characterization of new compounds, and spectroscopic data. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

david.leigh@manchester.ac.uk

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was funded by the Engineering and Physical Sciences Research Council (UK). We are grateful to the following organizations for postdoctoral fellowships: Fundacja na Rzecz Nauki Polskiej (to B.L.), Fonds Spécial de Recherche−Fédération Wallonie-Bruxelles and Wallonie-Bruxelles International (to G.D.B.)

REFERENCES

(1) (a) Yonath, A. *Angew. Chem., Int. Ed.* **2010**, *49*, 4340−4354. (b) Ramakrishnan, V. *Angew. Chem., Int. Ed.* **2010**, *49*, 4355−4380. (c) Steitz, T. A. *Angew. Chem., Int. Ed.* **2010**, *49*, 4381−4398.

(2) Bruck, I.; O'Donnell, M. *Genome Biol.* **2001**, *2*, 3001.1−3001.3.

(3) Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry*; W. H. Freeman: New York, 2006.

(4) (a) Kay, E. R.; Leigh, D. A.; Zerbetto, F. *Angew. Chem., Int. Ed.* **2007**, *46*, 72−191. (b) Durot, S.; Reviriego, F.; Sauvage, J.-P. *Dalton Trans.* **2010**, *39*, 10557−10570. (c) Boyle, M. M.; Smaldone, R. A.; Whalley, A. C.; Ambrogio, M. W.; Botros, Y. Y.; Stoddart, J. F. *Chem. Sci.* **2011**, *2*, 204−210. (d) de Juan, A.; Pérez, E. M. *Nanoscale* **2013**, *5*,

7141−7148. (e) Qu, D.-H.; Tian, H. *Chem. Sci.* **2013**, *4*, 3031−3035. (f) Caballero, A.; Zapata, F.; Beer, P. D. *Coord. Chem. Rev.* **2013**, *257*, 2434−2455. (g) Neal, E. A.; Goldup, S. M. *Chem. Commun.*, published online 17 Jan **2014**, DOI: 10.1039/C3CC47842D.

(5) (a) Lewandowski, B.; De Bo, G.; Ward, J. W.; Papmeyer, 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*, 189−193. (b) McGonigal, P. R.; Stoddart, J. F. *Nat. Chem.* **2013**, *5*, 260−262. (c) Bertran-Vicente, J. C.; Hackenberger, P. R. *Angew. Chem., Int. Ed.* **2013**, *52*, 6140−6142. (d) Wilson, C. M.; Gualandi, A.; Cozzi, P. G. *ChemBioChem.* **2013**, *14*, 1185−1187.

(6) For other artificial threaded processive catalysts, see: (a) Thordarson, P.; Bijsterveld, E. J. A.; Rowan, A. E.; Nolte, R. J. M. *Nature* **2003**, *424*, 915−918. (b) Miyagawa, N.; Watanabe, M.; Matsuyama, T.; Koyama, Y.; Moriuchi, T.; Hirao, T.; Furusho, Y.; Takata, T. *Chem. Commun.* **2010**, *46*, 1920−1922. (c) Monnereau, C.; Ramos, P. H.; Deutman, A. B. C.; Elemans, J. A. A. W.; Nolte, R. J. M.; Rowan, A. E. *J. Am. Chem. Soc.* **2010**, *132*, 1529−1531. (d) Takashima, Y.; Osaki, M.; Ishimaru, Y.; Yamaguchi, H.; Harada, A. *Angew. Chem., Int. Ed.* **2011**, *50*, 7524−7528. (e) van Dongen, S. F. M.; Clerx, J.; Nørgaard, K.; Bloemberg, T. G.; Cornelissen, J. J. L. M.; Trakselis, M. A.; Nelson, S. W.; Benkovic, S. J.; Rowan, A. E.; Nolte, R. J. M. *Nat. Chem.* **2013**, *5*, 945−951.

(7) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776−778.

(8) Crowley, J. D.; Goldup, S. M.; Lee, A.-L.; Leigh, D. A.; McBurney, R. T. *Chem. Soc. Rev.* **2009**, *38*, 1530−1541.

(9) (a) Aucagne, V.; Hänni, K. D.; Leigh, D. A.; Lusby, P. J.; Walker, D. B. *J. Am. Chem. Soc.* **2006**, *128*, 2186−2187. (b) Aucagne, V.; Berná, J.; Crowley, J. D.; Goldup, S. M.; Hänni, K. D.; Leigh, D. A.; Lusby, P. J.; Ronaldson, V. E.; Slawin, A. M. Z.; Viterisi, A.; Walker, D. B. *J. Am. Chem. Soc.* **2007**, *129*, 11950−11963. (c) Goldup, S. M.; Leigh, D. A.; Long, T.; McGonigal, P. R.; Symes, M. D.; Wu, J. *J. Am. Chem. Soc.* **2009**, *131*, 15924−15929. (d) Crowley, J. D.; Goldup, S. M.; Gowans, N. D.; Leigh, D. A.; Ronaldson, V. E.; Slawin, A. M. Z. *J. Am. Chem. Soc.* **2010**, *132*, 6243−6248. (e) Barran, P. E.; Cole, H. L.; Goldup, S. M.; Leigh, D. A.; McGonigal, P. R.; Symes, M. D.; Wu, J.; Zengerle, M. *Angew. Chem., Int. Ed.* **2011**, *50*, 12280−12284. (f) Lahlali, H.; Jobe, K.; Watkinson, M.; Goldup, S. M. *Angew. Chem., Int. Ed.* **2011**, *50*, 4151−4155. (g) Tron, A.; Thornton, P. J.; Rocher, M.; Jacquot de Rouville, H.-P.; Desvergne, J.-P.; Kauffmann, B.; Buffeteau, T.; Cavagnat, D.; Tucker, J. H. R.; McClenaghan, N. D. *Org. Lett.* **2014**, *16*, 1358−1361. (h) Bordolli, R. J.; Goldup, S. M. *J. Am. Chem. Soc.* **2014**, *136*, 4817−4820.

(10) (a) Tornøe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057−3064. (b) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596−2599.

(11) Bhat, V. T.; Caniard, A. M.; Luksch, T.; Brenk, R.; Campopiano, D. J.; Greaney, M. F. *Nat. Chem.* **2010**, *2*, 490−497.

(12) Dawson, P. E.; Kent, S. B. H. *Annu. Rev. Biochem.* **2000**, *69*, 923−960.

(13) Levison, M. E.; Josephson, A. S.; Kirschenbaum, D. M. *Experientia* **1969**, *25*, 126−127.

(14) (a) Lutz, J.-F. *Polym. Chem.* **2010**, *1*, 55−62. (b) Lutz, J.-F.; Ouchi, M.; Liu, D. R.; Sawamoto, M. *Science* **2013**, *341*, 1238149.