Ugi multicomponent reaction to prepare peptide–peptoid hybrid structures with diverse chemical functionalities

Article

Accepted Version

Chemical Science

ARTICLE

Iterative synthesis of sequence-defined peptide-peptoid hybrid structures with diverse chemical functionalities

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Manuel Hartweg,[1] Charlotte J. C. Edwards-Gayle,[2,3] Elham Radvar,[1] Dominic Collis,[1] Mehedi Reza,[4] Michael Kaupp,[5] Jan Steinkönig,[5] Janne Ruokolainen,[4] Robert Rambo,[3] Christopher Barner-Kowollik,[5] Ian W. Hamley,[2] Helena S. Azevedo[1] and C. Remzi Becer[1*]

Monodisperse sequenced peptides and peptoids present unique nano-structures based on their self-assembled secondary and tertiary structures. However, the generation of peptide and peptoid hybrid structures in a sequence-defined manner has not yet been studied. Herein, we pioneer a synthetic strategy that enables both the modification of peptides as well as the generation of sequence-defined peptide-peptoid hybrid structures. Our synthetic methodology rests on the fusion of solid phase peptide synthesis with Ugi multicomponent reactions. We evidence that a diversity of chemical functionalities can be inserted into peptides or used in the design of peptide-peptoid hybrids exploiting a wide functional array including amines, carboxylic acids, hydrocarbons, carbohydrates as well as polymers, introducing a sequence-defined synthetic platform technology for precision peptoid hybrids.

Introduction

Nature has created sequence defined peptides to carry out specific functions, such as molecular recognition and biocatalysis, in the course of evolution.[1] Today, even though synthetic macromolecules are widely used for a large number of applications, achieving an absolute control over their primary structure still remains as a key challenge for chemists.[2,3,4] Therefore, our fundamental understanding of the role of monomer sequence and its influence on the secondary and tertiary structures of peptides and proteins has to be enhanced dramatically.

Amongst a range of methods for the synthesis of such sequence-defined macromolecules, Merrifield solid phase peptide synthesis (SPPS) method has been the most successful.[5-7,8-17] Solid-phase protocols typically use excess of reagents and require two steps, which are the coupling and deprotection steps, for incorporation of each repeat unit. Besides, extremely efficient reactions are required to ensure monodisperse formation of peptide chains in high yields.

Complementary to the synthesis of peptides, Zuckermann *et al*. reported the synthesis of peptoids *via* solid-phase submonomer synthesis. Polypeptoids are accepted as peptidomimetic polymers[18] with a substituent on the amide nitrogen, which significantly influences the formation of higher order structures due to suppression of a hydrogen bonding.[19, 20] Moreover, the *N*-substitution eliminates the chirality of the more flexible peptoid backbone,[21] which transforms peptoids to a generally side-chain dominated system, dissimilar to peptides in which conformation and packing are dictated primarily by the chemical structure of the backbone. However, compared to peptides, engineered peptoids display a range of interesting properties, such as protein-mimetic secondary and tertiary structures.[22, 23] They have been demonstrated to form stable nanosheets *in vivo*,[24-26] as well as to improve tissue accumulation for reduced excretion rates.[27,28] Further successful examples for the peptoid synthesis can be listed as solid-phase synthesis *via* Fmoc-strategy,[29] bromoacetyl-bromide method in solution-phase,[30] *N*-heterocyclic carbene mediated ring-opening polymerization of *N*-substituted *N*-carboxy-anhydrides[31], and Ugi 4-component reaction (Ugi-4CR).[32] However, the combination of peptide and peptoids in a sequence defined way has not yet been reported, except a few studies have reported the synthesis of peptide-peptoid block conjugates,[33, 34] using native chemical ligation (NCL)[11], fragment condensation,[35] and copper-catalyzed azide-alkyne [3 + 2] cycloaddition.[36] Moreover, Ugi-4CR was adopted in this study as it displays the desired high efficiency,[37-39] compatibility with amino acid chemistry[40], while allowing incorporation of a range of chemical diversity into peptoid side

[1] *School of Engineering and Materials Science, Queen Mary University, London E1 4NS, United Kingdom*
[2] *Department of Chemistry, University of Reading, Whiteknights, Reading, RG6 6AD, UK*
[3] *Diamond Light Soruce Ltd., Harwell Science and Innovation Campus, Didcot, OX11 0DE, UK*
[4] *Department of Applied Physics, Aalto University, Finland*
[5] *Preparative Macromolecular Chemistry, Institut für Technische Chemie und Polymerchemie, Karlsruhe Institute of Technology (KIT), Engesserstraße 18, 76128 Karlsruhe, Germany and Institute of Biological Interfaces, Karlsruhe Institute of Technology (KIT), Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshaven, Germany*
School of Chemistry, Physics and Mechanical Engineering, Queensland University of Technology (QUT), 2 George Street, QLD 4000, Brisbane, Australia
Electronic Supplementary Information (ESI) available: [polymerization kinetics and analysis, competition binding assays]. See DOI: 10.1039/x0xx00000x

chains.[44] The method benefits from its simplicity and remarkably high diversity for modification of peptides or synthesis of novel peptoid and hybrid structures.

Herein, pioneer an iterative synthesis method towards highly functional, sequence-defined, and monodisperse peptides, peptoids, and peptide-peptoid hybrids (**Fig. 1**). The methodology fuses SPPS and Ugi-4CR for the incorporation of peptide and peptoid units, respectively. Furthermore, we investigate the self-assembly behavior and defragmentation pattern of selected compounds to demonstrate the diversity of these hybrid structures.

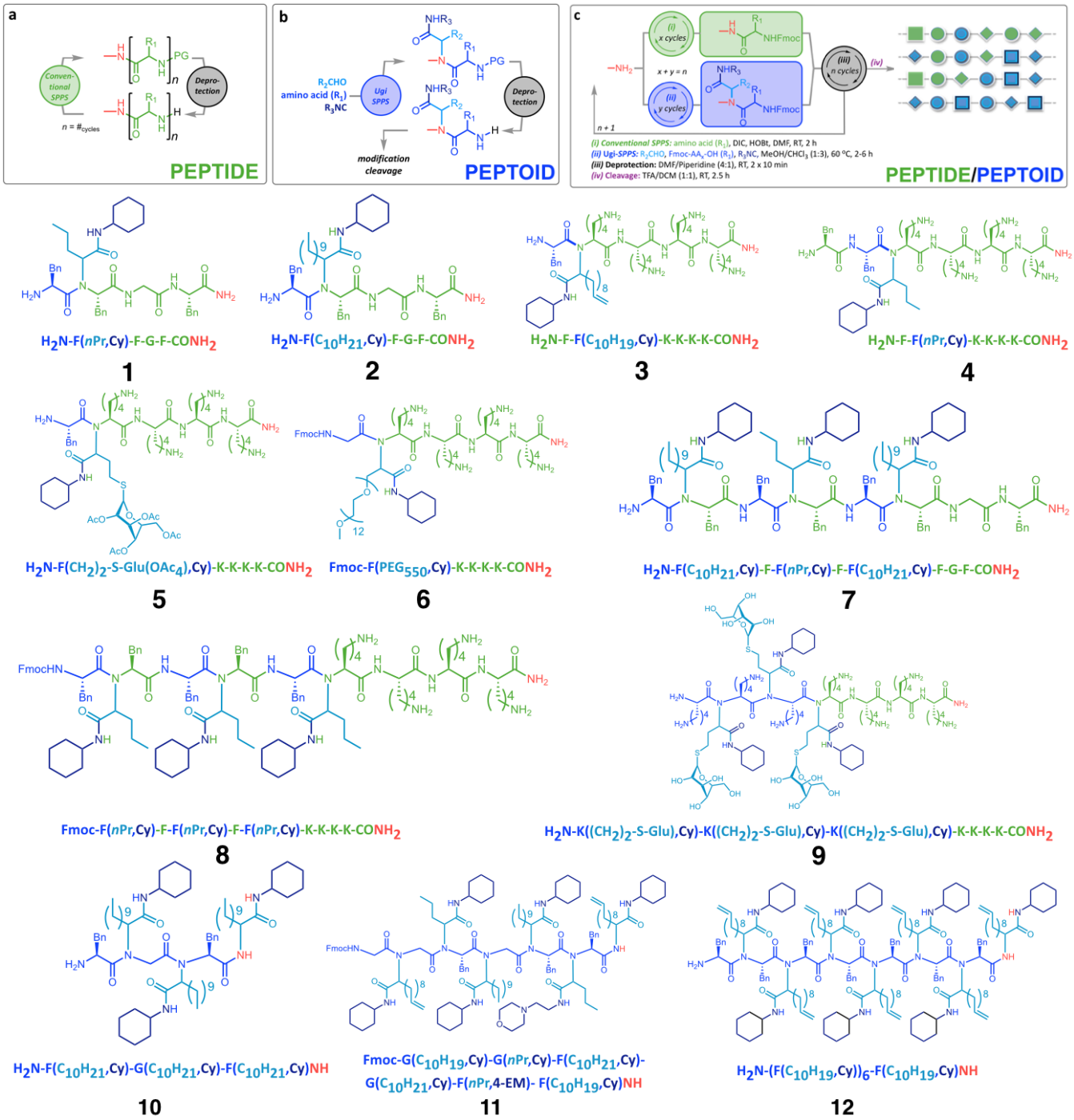

Figure 1. (a) Conventional solid-phase peptide synthesis (SPPS), (b) Solid-phase peptoid synthesis via Ugi reaction, and (c) the iterative process for the synthesis of peptide-peptoid hybrids *via* combination of SPPS and Ugi-SPPS (AA = amino acid; R_A = $R_{aldehyde}$; R_{IC} = $R_{isocyanide}$). (1-12) Chemical structures of the peptide/peptoid hybrid structures synthesized in this study.

RESULTS AND DISCUSSION

The preliminary considerations indicated that the distinct reaction protocols of conventional SPPS and Ugi-SPPS are suitable for iterative, random, or sequence-defined solid phase synthesis. For the Ugi-SPPS process, the combinations of several amino acids (*i.e.* glycine, phenylalanine and lysine), aldehydes (*i.e.* butyraldehyde, undecanal, undecenal, glucose aldehyde, and PEG-aldehyde), and isocyanides (cyclohexyl, and 2-morpholinoethyl isocyanide) were investigated. An Fmoc-Rink-amide coupling strategy was employed throughout the process for both coupling methodologies. In order to validate the successful combination compounds **1-6** (**Fig. 1**) were prepared. Thus, peptide anchors (H_2N-FGF-$CONH_2$ in **1**-**2**: H_2N-KKKK-$CONH_2$ in **3**-**6**) were prepared *via* automated SPPS on Rink-amide resin. Anchors were then deprotected to generate the free amines for incorporation of subsequent peptide or peptoid units.

Afterwards, in contrast to conventional SPPS, where the amine moieties are reacted with coupling reagents and an amino acid, amines were reacted with an excess of aldehyde (4.00 eq) in 1:3 MeOH/$CHCl_3$. It is important to note that the use of both MeOH and $CHCl_3$ was crucial to achieve full and selective conversion of the terminal amine into a peptoid sequence. $CHCl_3$ is essential to ensure sufficient swelling of the resin, whereas MeOH is required to promote Ugi reaction. After complete imine formation, the respective Fmoc-protected amino acid (5.00 eq) and isocyanide (6.00 eq) were added to react for a further 2-6 hours. Terminal free amine has been formed as a result of the subsequent deprotection step in 1:4 Pyridine/DMF. Successive coupling steps were performed accordingly to obtain the desired sequence of the final peptoid or peptide/peptoid hybrids. Following the last coupling reaction or the final deprotection step, oligomers were cleaved off the solid support with 1:1 TFA/DCM mixture and crude mixtures were analyzed. Electrospray ionization mass spectrometer (ESI-MS) or matrix assisted laser desorption ionization time of flight mass spectrometer (MALDI-ToF/MS) were used to investigate the conversion and chemoselectivity of oligomers. Furthermore, reverse phase high performance liquid chromatography (RP-HPLC) or size exclusion chromatography (SEC) was conducted to confirm the purity of monodisperse oligomers. It should be noted that these reactions are highly efficient and not even traces of the conventional peptide derivatives were observed as side products. SEC analysis of **1** and **2** have displayed narrow peaks and indicate the formation of the desired sequences (**Fig. S6**-**S7**).

Subsequently, to obtain more polar hybrids, H_2N-KKKK-$CONH_2$ was used as an anchor. The reactions with undecenal or butyraldehyde, Fmoc-Phe-OH, and cyclohexyl isocyanide, as well as successive SPPS with Fmoc-Phe-OH, yielded oligomeric **3** and **4** at high purity based on RP-HPLC traces (**Fig. S10**-**S12**). Compound **3** has two main isomers, one from *cis* and one from the *trans,* of the *bis*-amide formed as a result of Ugi-4CR. Furthermore, ESI-MS analysis was carried out for **3** and **4**, where mainly singly and doubly charged cations were observed (**Fig. S10-S12**). Subsequently, the addition of a glucose aldehyde resulted in a chemoselective formation of glycopeptoid -peptide hybrid **5** that was confirmed by RP-HPLC and ESI-MS (**Fig. S14**). An acetyl-protected sugar-derived aldehyde was synthesized *via* base catalyzed thiol-Michael addition using tetra-acetylated thioglucose and acrolein (ESI). Additionally, Ugi-4CR was used as a ligation tool in the synthesis of the hybrid **6**. Therefore, PEG-aldehyde was synthesized by Swern oxidation reaction (ESI).[47] Thus, PEG_{500} was used as a starting material and ligated to the anchor compound **6**. The success of the reaction was determined by MALDI-ToF, whereby the PEG signals shift by $\Delta m/z$ = 917.2 Da (**Fig. S15**), which corresponds exactly to the mass of peptoid/aldehyde.

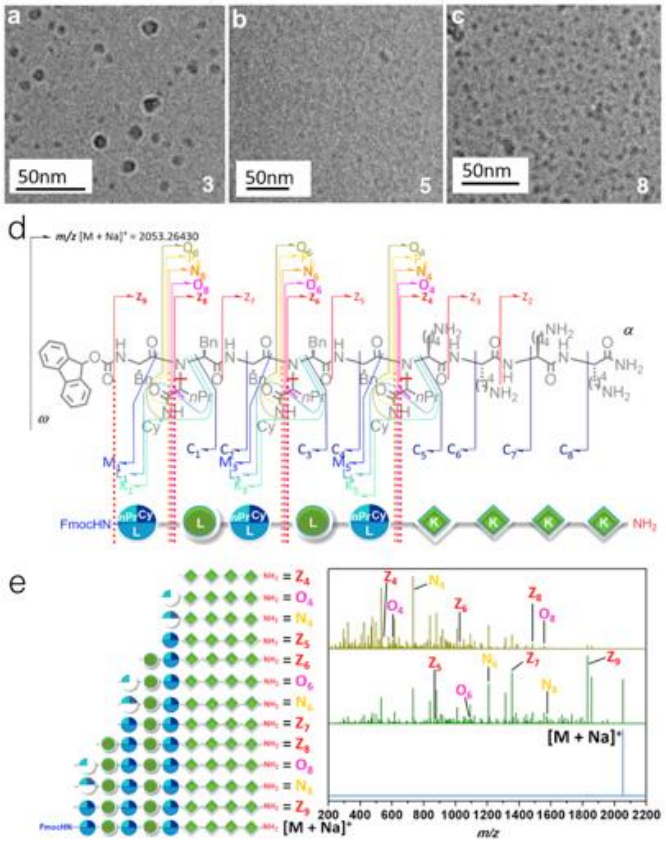

Figure 2. (a-c) Cryo-TEM and (d-e) MS/MS analysis data. (a) Compound **3**, 1 w% solution, 20 v% MeOH. (b) Compound **5**, 1 w% solution. (c) Compound **8**, 10 v% THF.(d) Schematic presentation of compound **8** and its assignment into different fragmentation patterns. (e) MS/MS analysis of compound **8**. **Blue trace:** isolated parent ion $[M + Na]^+$. **Green trace:** MS/MS at lower energy. **Gold trace:** MS/MS at higher energy.

To further expand the versatility of this synthetic process, several hybrids with different backbone architectures, such as alternating **7**, and **8** and block **9** co-oligomers, were prepared. Starting from the same H_2N-FGF-$CONH_2$ peptide anchor, three peptoid sequences were attached with a phenylalanine peptide sequence in between the peptoids. The MS analysis of compound **7** (**Fig. S17**) demonstrated high purity even in the absence of any purification steps. Also, H_2N-KKKK-$CONH_2$ was

designed as a more polar anchor and used in the synthesis of **8** and **9**. ESI-MS (**Fig. S19**) shows the formation of the 9-mer displaying the singly, doubly and triply charged cations. Compound **9** is designed as a glycopeptoid containing three glycan units **9**. Acetyl-protected glucose aldehyde was used in the synthesis and following to the deacetylation of glucose units, ESI-MS revealed signals for the doubly and triply charged cations (**Fig. S21**). RP-HPLC analysis of compound **9** also indicated the formation of an oligomer in high purity.

In the final set of hybrid structures, homopeptoids were studied to obtain highly functional oligopeptoids. Initially, **10** was prepared and analyzed as a trimer. MS analysis (**Fig. S23**) indicated a small amount of a side product, which was assigned as the trimer incorporating only two peptoid and one peptide sequence instead of the third peptoid unit. The main reason for this effect is ascribed as the steric hindrance as the reaction rate significantly suffers when bulky reagents are used in the Ugi reaction. The isotopic pattern of the main peak (**Fig. S24**) is in good agreement with the theoretical.

Moreover, hexamer compound **11** using different starting compounds and heptamer compound **12** with identical units and molecular weights of approximately 3 kDa were synthesized. In the case of **11**, the MALDI-ToF/MS analysis indicated that when 2-morpholinoethyl isocyanide was used alternatively, the reaction was not chemoselective and thus only one single peptide sequence was incorporated in the backbone (**Fig. S25**). MALDI-ToF/MS spectrum of **12** shows a series of peaks with a low intensity and $\Delta m/z$ of one single Ugi-repetition unit was detected (**Fig. S26**). Nevertheless, the major peak was assigned to be the desired compound **12** and the isotopic pattern of **8** and **9** are found to be in good agreement.

Having demonstrated the versatility of the synthetic process to prepare sequence-defined peptide-peptoid hybrid structures, their ability for information storage and self-assembly into nanostructures have been investigated. Biopolymers are usually decoded fast and reliably *via* tandem mass spectroscopy (MS/MS). However, no universal rules apply for the MS/MS dissociation patterns of non-natural polymers. As illustrated (**Fig. 2**), the composition of amino acids of peptide units can be assigned using their one letter code (*i.e.* K = lysine, F = phenylalanine). In order to demonstrate the power of data storage on designed hybrid structures, compound **8** was analyzed *via* ESI-MS/MS. The parent ion ($[M + Na]^+$ (m/z = 2053.26 g·mol^{-1})) was isolated and activated with increasing collisional energy (**Fig. 2** blue trace). As expected, even at lower energies, a fragment (Z_9) was detected initially, which displays the cleavage of the urethane bond of the Fmoc-group ($\Delta m/z$ = 222 Da). When the collisional energy was increased (**Fig. 2** green and gold spectrum), further fragmentation patterns were observed. This powerful technique allowed not only observation of fragmentation patterns from the main chain (C, M, N and Z), but also the identification of specific side chains. These outcomes provided structural information about both R-groups of the *N*-substitution of the peptoid units and thus revealing the location of every starting compound used in the respective Ugi-reaction. For instance, the fragmentation of the peptoid sequence situated at the eighth sequence can be envisioned (N_8 vs O_8 vs Z_8). The fragment of the oligomer including sequence eight, N_8 (m/z = 1685.09 Da) displays $\Delta m/z$ = 127.12 Da compared to O_8 (m/z = 1557.97 Da), which corresponds to CyNHC(H)=O. Z_8 (m/z = 1500.95 Da) provides information about the aldehyde derived R-group. Morevover, a $\Delta m/z$ = 57.02 Da value compared to Z_8 belongs to $CH_3(CH_2)_2CH_2$ and displays the use of butyraldehyde as a monomer in the Ugi-reaction.

Secondary structures of **3**, **5** and **8** were determined using circular dichroism (CD) spectroscopy. It was found that the structures were intrinsically disordered, identified through a negative peak around 200 nm (**Fig. S29**). The self-assembly properties of these three compounds were then investigated using cryogenic transmission electron microscopy (Cryo-TEM) and small-angle X-ray scattering (SAXS). Cryo-TEM images of **3** (**Fig. 2a**) revealed the presence of slightly larger globular structures with an average diameter of 14.95 nm. Compound **5** showed a small population of irregular sized aggregates or clusters (**Fig. 2b**). Finally, for compound **8** mixture of larger globular like structures and smaller micelle like assemblies with an average diameter of 3.125 nm are observed (**Fig. 2c**). SAXS was used to further probe the self-assembled structures of the hybrid oligomers, providing information on the structure and dimensions of the assemblies. Scattering intensities and form factor models are shown (**Fig. S30**). Scattering intensities of **3** and **5** were fitted based on a coexistence of clusters (described by a mass fractal form factor) and monomers (described by generalized Gaussian coil form factor). The models are in good agreement with cryo-TEM results, which showed a population of irregular aggregated structures. SAXS data for compound **8** fitted to a 'spherical shell' form factor model, corresponding to a core-shell micelle structure, also consistent with Cryo-TEM results. Parameters of fits are listed in **Tables S2**-**S3**.

Conclusions

In summary, two chemical synthesis techniques, i.e. solid-phase peptide synthesis (SPPS) and Ugi multicomponent reaction were combined to obtain peptide/peptoid hybrid structures with defined sequences and functionalities. Thus, a series of model peptide/peptoid hybrids were prepared. As evident from the HPLC and MS spectra of the hybrids, both reactions proceeded at very high conversions and no purification was required at all, demonstrating the power of the combined iterative approach and the diversity of the chemical groups that can be inserted in peptide sequences via Ugi reactions.

Acknowledgements

This work was partially supported by the European Commission Horizon2020 programme (EU-ITN EuroSequences Proposal No: 642083). C.R.B. also acknowledges EPSRC

(EP/P009018/1) for the financial support. D.C. and H.S.A. thank the support from the EU-funded project "SuprHApolymers" (PCIG14-GA-2013-631871). I.W.H. thanks EPSRC (EP/L020599/1) for support and Diamond Light Source and the University of Reading for co-funding the PhD studentship of C.E.G.. Diamond Light Source is thanked for the award of beamtime on beamline B21. C.B.-K. acknowledges the Sonderforschungsbereich 1176 (project A3) funded by the German Research Council for continued support.

Notes and references

1 J.-F. Lutz, J.-M. Lehn, E. W. Meijer and K. Matyjaszewski, *Nat. Rev. Mater.*, **2016**, *1*, 16024.
2 J.-F. Lutz, *Polym. Chem.*, **2010**, *1*, 55-62.
3 J.-F. Lutz, M. Ouchi, D. R. Liu and M. Sawamoto, *Science*, **2013**, *34, 1238149.*
4 Y. Brudno and D. R. Liu, *Chem. & Biol.*, **2009**, *16*, 265-276.
5 R. B. Merrifield, *J. Am. Chem. Soc.*, **1963**, *85*, 2149-2154.
6 R. B. Merrifield, *Angew. Chem.*, **1985**, *97*, 801-812.
7 R. B. Merrifield, *Angew. Chem. Int. Ed. Engl.*, **1985**, *24*, 799-810.
8 R. N. Zuckermann, J. M. Kerr, S. B. H. Kent and W. H. Moos, *J. Am. Chem. Soc.*, **1992**, *114*, 10646-10647.
9 S. J. Danishefsky, K. F. McClure, J. T. Randolph and R. B. Ruggeri, *Science*, **1993**, *260*, 1307-1309.
10 M. Schuster, P. Wang, J. C. Paulson and C.-H. Wong, *J. Am. Chem. Soc.*, **1994**, *116*, 1135-1136.
11 P. E. Dawson, T. W. Muir, S. B. Kent, *Science*, **1994**, *266*, 776-779.
12 O. J. Plante and P. H. Seeberger, *Science*, **2001**, *291*, 1523-1527.
13 P. Espeel, L. L. G. Carrette, S. Capenberghs, J. C. Martins, F. E. Du Prez and A. Madder, *Angew. Chem. Int. Ed.*, **2013**, *52*, 13261-13264.
14 T. G. W. Edwardson, K. M. M. Carneiro, C. J. Serpell and H. F. Sleiman, *Angew. Chem. Int. Ed.*, **2014**, *53*, 4567-4571.
15 D. Ponader, S. Igde, M. Wehle, K. Märker, M. Santer, D. Bléger and L. Hartmann, *Beilstein J. Org. Chem.*, **2014**, *10*, 1603-1612.
16 R. K. Roy, A. Meszynska, C. Laure, L. Charles, C. Verchin and J.-F. Lutz, *Nat. Commun.*, **2015**, *6*.
17 A. A. Ouahabi, M. Kotera, L. Charles and J.-F. Lutz, *ACS Macro Lett.*, **2015**, *4*, 1077-1080.
18 J. Sun and R. N. Zuckermann, *ACS Nano*, **2013**, *7*, 4715-4732.
19 B. Laufer, A. O. Frank and H. Kessler, *J. Pept. Sci.*, **2009**, *15*, 141-146.
20 N. Gangloff, J. Ulbricht, T. Lorson, H. Schlaad and R. Luxenhofer, *Chem. Rev.*, **2016**, *116*, 1753-1802.
21 A. M. Rosales, H. K. Murnen, S. R. Kline, R. N. Zuckermann and R. A. Segalman, *Soft Matter*, **2012**, *8*, 3673-3680.
22 B.-C. Lee, R. N. Zuckermann and K. A. Dill, *J. Am. Chem. Soc.*, **2005**, *127*, 10999-11009.
23 B.-C. Lee, T. K. Chu, K. A. Dill and R. N. Zuckermann, *J. Am. Chem. Soc.*, **2008**, *130*, 8847-8855.
24 B. Sanii, R. Kudirka, A. Cho, N. Venkateswaran, G. K. Olivier, A. M. Olson, H. Tran, R. M. Harada, L. Tan and R. N. Zuckermann, *J. Am. Chem. Soc.*, **2011**, *133*, 20808-20815.
25 K. T. Nam, S. A. Shelby, P. H. Choi, A. B. Marciel, R. Chen, L. Tan, T. K. Chu, R. A. Mesch, B.-C. Lee, M. D. Connolly, C. Kisielowski and R. N. Zuckermann, *Nat. Mater.*, **2010**, *9*, 454-460.
26 R. V. Mannige, T. K. Haxton, A. Battigelli, G. L. Butterfoss, R. N. Zuckermann and S. Whitelam, *Nature*, **2015**, *526*, 415-420.
27 J. Seo, G. Ren, H. Liu, Z. Miao, M. Park, Y. Wang, T. M. Miller, A. E. Barron and Z. Cheng, *Bioconjugate Chem.*, **2012**, *23*, 1069-1079.
28 B.-C. L. J. Seo, R. N. Zuckermann, in *Comprehensive Biomaterials*, Elsevier, **2011**, vol. *2*, pp. 53-76.
29 R. J. Simon, R. S. Kania, R. N. Zuckermann, V. D. Huebner, D. A. Jewell, S. Banville, S. Ng, L. Wang, S. Rosenberg and C. K. Marlowe, *Proc. Natl. Acad. Sci*, **1992**, *89*, 9367-9371.
30 T. Hjelmgaard, S. Faure, C. Caumes, E. De Santis, A. A. Edwards and C. Taillefumier, *Org. Lett.*, **2009**, *11*, 4100-4103.
31 L. Guo and D. Zhang, *J. Am. Chem. Soc.*, **2009**, *131*, 18072-18074.
32 L. A. Wessjohann, O. E. Vercillo, *Chem. Rev.*, **2009**, *109*, 796-814.
33 K. Pels and T. Kodadek, *ACS Comb. Sci.*, **2015**, *17*, 152-155.
34 Y. Gao and T. Kodadek, *Chem. & biol.*, **2013**, *20*, 360-369.
35 P. M. Levine, T. W. Craven, R. Bonneau and K. Kirshenbaum, *Org. Biomol. Chem.*, **2013**, *11*, 4142-4146.
36 D. Pasini, *Molecules*, **2013**, *18*, 9512.
37 I. Ugi, A. Dömling and W. Hörl, *Endeavour*, **1994**, *18*, 115-122.
38 A. Dömling, K. Wang, *Chem. Rev.*, **2012**, *112*, 3083-3135.
39 E. Ruijter, R. Scheffelaar and R. V. A. Orru, *Angew. Chem. Int. Ed.*, **2011**, *50*, 6234-6246.
40 A. D. F. S. Barreto, O. E. Vercillo, J. C. Caulfield, L. A. Wessjohann and C. K. Z. Andrade, *Org. Biomol. Chem.*, **2011**, *9*, 5024-5027.
41 A. K. Szardenings, T. S. Burkoth, H. H. Lu, D. W. Tien and D. A. Campbell, *Tetrahedron*, **1997**, *53*, 6573-6593.
42 K. M. Short and A. M. M. Mjalli, *Tetrahedron Lett.*, **1997**, *38*, 359-362.
43 A. M. L. Hoel and J. Nielsen, *Tetrahedron Lett.*, **1999**, *40*, 3941-3944.
44 S. C. Solleder, K. S. Wetzel and M. A. R. Meier, *Polym. Chem.*, **2015**, *6*, 3201-3204.
45 D. G. Udugamasooriya, S. P. Dineen, R. A. Brekken and T. Kodadek, *J. Am. Chem. Soc.*, **2008**, *130*, 5744-5752.
46 J. E. Murphy, T. Uno, J. D. Hamer, F. E. Cohen, V. Dwarki and R. N. Zuckermann, *Proc. Natl. Acad. Sci*, **1998**, *95*, 1517-1522.
47 K. Omura and D. Swern, *Tetrahedron*, **1978**, *34*, 1651-1660.
48 M. Ciaccia and S. Di Stefano, *Org. Biomol. Chem.*, **2015**, *13*, 646-654.