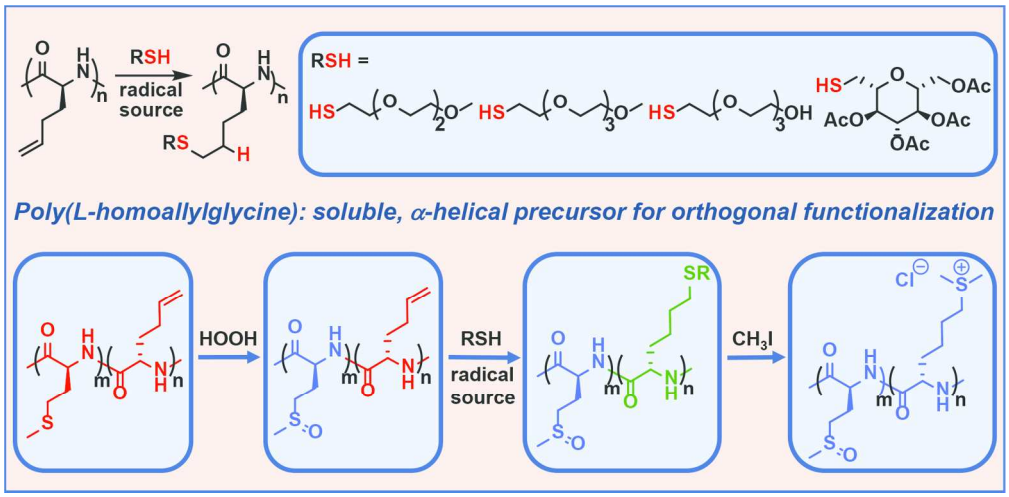
ChemComm

Homoallylglycine residues are superior precursors to orthogonally modified thioether containing polypeptides

Journal:	*ChemComm*
Manuscript ID	CC-COM-04-2018-003048.R1
Article Type:	Communication



161x78mm (300 x 300 DPI)

Journal Name

COMMUNICATION

Homoallylglycine residues are superior precursors to orthogonally modified thioether containing polypeptides†

Pesach Perlin,[a] Eric G. Gharakhanian,[a] and Timothy J. Deming[a,b]

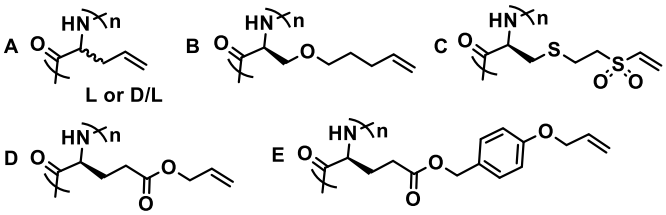
Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Homoallylglycine N-carboxyanhydride, Hag NCA, monomers were synthesized and used to prepare polypeptides containing Hag segments with controllable lengths of up to 245 repeats. Poly(L-homoallylglycine), G^{HA}, was found to adopt an α-helical conformation, which provided good solubility in organic solvents and allowed high yield functionalization of its alkene side-chains via radical promoted addition of thiols. The conformations of these derivatives were shown to be switchable between α-helical and disordered states in aqueous media using thioether alkylation or oxidation reactions. Incorporation of G^{HA} segments into block copolymers with poly(L-methionine), M, segments provided a means to orthogonally modify thioether side-chains different ways in separate copolypeptide domains. This approach allows preparation of functional polypeptides containing discrete domains of oxidized and alkylated thioether containing residues, where chain conformation and functionality of each domain can be independently modified.

There has been considerable recent interest in the development of methods to selectively introduce functional tags into peptide, protein, and polypeptide sequences.[1] Among these, the thiol-ene reaction has been used extensively,[2] since it can provide modifications with high yields and high functional group selectivity. For polypeptides, many unnatural alkene containing residues have been employed for subsequent thiol-ene modification (Scheme 1).[3] In these examples, the side-chain structures of these alkene amino acid residues have substantial influence on resulting polypeptide chain lengths, conformations, solubility, and consequently the efficiency of thiol-ene conjugations. We sought to optimize the design of alkene containing residues to enable robust polypeptide and block copolypeptide synthesis, high efficiency in subsequent thiol-ene modifications, and control of chain conformations.



Scheme 1. (A-E) Alkene containing homopolypeptides used for thiol-ene conjugation.

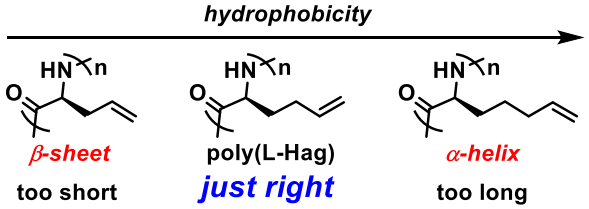
The simplest alkene containing polypeptides used for thiol-ene functionalization are based on allylglycine (Scheme 1A). Both poly(L-allylglycine) and poly(DL-allylglycine) have been prepared and were found to adopt β-sheet conformations, which result in aggregation and limit the ability to prepare high molecular weight chains.[4] Consequently, efficient thiol-ene functionalization of these polymers was restricted to samples with short chain lengths (i.e. typically < 20 residues), and often required incorporation of comonomers or segments (i.e. PEG) to promote solubility.[5] Related polypeptides have been prepared based on alkene functionalized serine[6] and cysteine[7] residues (Scheme 1B,C) that also adopt β-sheet conformations leading to poorly controlled chain aggregation, which would be problematic for downstream use as segments in block copolypeptide assemblies.

Additional studies have utilized functionalized glutamate esters as components for preparation of alkene containing polypeptides (Scheme 1D,E).[8] These polypeptides have the advantage of adopting α-helical conformations, which possess good solubility and allow formation of high molecular weight chains. Homopolypeptides up to 100 residues long were prepared and could be efficiently modified with different thiols yielding α-helical derivatives. While this strategy is useful for preparation of homopolypeptides, the labile side-chain ester linkages would be problematic in preparation of multifunctional copolypeptides.[9] Also, this strategy only allows for preparation of polypeptides with α-helical conformations, which cannot be

[a] Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095.
[b] Department of Bioengineering, University of California, Los Angeles, CA 90095. Corresponding author: demingt@seas.ucla.edu.

Electronic Supplementary Information (ESI) available: [Supporting figures S1-8, tables S1-2, schemes S1-2, experimental procedures and spectral data for all new compounds]. See DOI: 10.1039/x0xx00000x

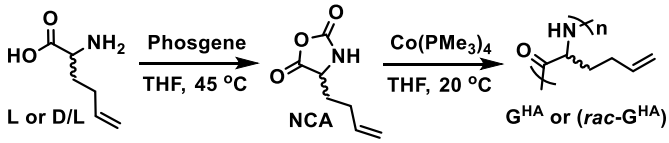
switched due to their long, hydrophobic side-chains.[8] Polypeptides with conformations that can be switched reversibly under mild conditions are desirable for use in development of self-assembled materials such as nanocarriers and hydrogels that can actively respond to biological and chemical cues.



Scheme 2. Comparison of allyl, homoallyl, and pentenyl glycine homopolypeptides.

To take full advantage of the thioether functionality introduced by thiol-ene conjugation, we sought to develop polypeptides containing alkene side-chains of minimal length so that modifications of product thioether groups would induce switchable chain conformations.[10] Further, to enable preparation of soluble, high molecular weight chains, α-helical conformations were desired for the initial alkene bearing polypeptides. While poly(allylglycine)s are known to form β-sheets, it has also been reported that poly(L-pentenylglycine) adopts an α-helical conformation (Scheme 2).[4] While poly(L-pentenylglycine) has not been used for thiol-ene conjugation, its hydrophobic side-chains might be too long to allow conformational switching, similar to the glutamate derivatives described above. Since single carbon homologation of side-chain functional groups in β-sheet forming polypeptides can result in polypeptides that adopt α-helical conformations, such as homologation of cysteine to homocysteine,[11] we hypothesized that the intermediate side-chain length of Hag would be sufficiently long to stabilize α-helical conformations in **G^{HA}** and yet be short enough to allow introduced thioether groups to influence chain conformations (Scheme 2). Notably, Hag has also been utilized as an artificial residue for efficient thiol-ene modification in proteins.[12]

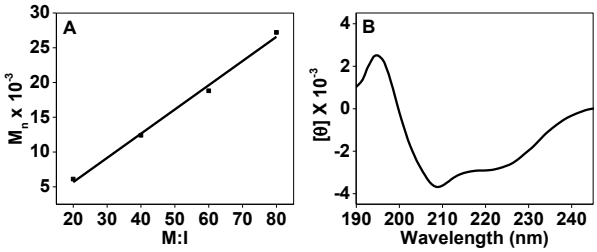
Consequently, we sought to develop procedures for preparation of new NCA monomers of L-Hag and *rac*-Hag, and synthesize their corresponding new homopolypeptides (Scheme 3). To enhance the ability to prepare multifunctional polypeptides with stimulus responsive conformations, we also sought to prepare block copolypeptides of Hag with L-methionine, Met. Specifically, we aimed to utilize Hag residues as "masked" precursors of thioether groups, which could be functionalized orthogonally to the thioether groups in Met residues. The goal of this approach being the preparation of block copolypeptides where discrete domains can be functionalized and conformationally switched independent of one another.



Scheme 3. Synthesis of homoallylglycine NCAs and polypeptides.

For NCA monomer preparation, the Hag and *rac*-Hag amino acid precursors were prepared following literature methods (see Scheme S1, ESI†).[12,13] *rac*-Hag was obtained by alkylation of diphenylimino glycine *tert*-butyl ester, which gave the free amino acid directly upon hydrolysis of the protecting groups. Hag was prepared by alkylation of diethyl acetamidomalonate, followed by ester deprotection and decarboxylation to give N-acetyl-*rac*-Hag. This racemic mixture was readily resolved by enantioselective hydrolysis catalyzed by porcine acylase to give multigram quantities of Hag (see Figures S1-2), which possessed an enantiomeric excess of >99% suitable for preparation of NCAs and polypeptides with high optical purity. Analysis of both Hag and *rac*-Hag matched literature data.[12,13] Hag and *rac*-Hag were each subsequently treated with phosgene under standard conditions to obtain the corresponding NCAs that were obtained as high purity colorless solids after column chromatography and recrystallization (see Figures S3-4).[14]

Figure 1. Synthesis and properties of poly(L-homoallylglycine), **G^{HA}**. (A)



Number average molecular weight (M_n) of **G^{HA}** plotted as a function of monomer to initiator ratio (M:I) at complete monomer conversion using $Co(PMe_3)_4$ in THF (r^2 = 0.9874). M_n values were determined by ^1H NMR analysis of PEG end-capped polymers. (B) Circular dichroism spectrum of **G^{HA}** in 15:1:2 cyclohexane:MeCN:IPA (0.5 mg/mL) at 20 °C. Molar ellipticity reported in deg·cm^2/dmol.

Homopolymerizations of Hag and *rac*-Hag NCAs at different monomer to initiator (M:I) ratios were conducted using $Co(PMe_3)_4$ initiator in THF.[15] While Hag NCA polymerizations rapidly went to completion and remained homogeneous up to M:I = 80 (Figure 1A), the *rac*-Hag NCA polymerizations did not go to completion above M:I = 20 (see Tables S1-2, Figure S5). By FTIR we observed the poly(DL-homoallylglycine), **(*rac*-G^{HA})**, forms β-sheet aggregates during polymerization that likely inhibit chain growth (see Figure S6).[4] On the contrary, **G^{HA}** homopolymers were found to be highly α-helical in organic solvents (Figure 1B), which promoted good solubility and enabled the synthesis of polymers up to 245 residues long. Analysis of chain lengths at different M:I showed linear chain growth during Hag NCA polymerization, an indicator of controlled polymerization (Figure 1A). GPC analysis of **G^{HA}** samples, derivatized using thiol-ene reactions to improve solubility (Figure 2), gave unimodal peaks with dispersities of *ca.* 1.1-1.2, confirming the formation of uniform polymers. To further test the ability of Hag NCA to undergo controlled polymerization, diblock copolypeptides with Met NCA were prepared (Table 1). Block copolypeptides of defined sequence and composition were obtained in excellent yields, and GPC analysis of derivatized copolymers (*vide infra*) showed uniform chain length distributions with low dispersity (Figure 2).

Compositions		First Segment[b]			Diblock Copolymer[c]			
1st Monomer[a]	2nd Monomer[a]	M_n	DP	M_w/M_n	M_n	DP	M_w/M_n	Yield (%)[d]
20 Met NCA	10 Hag NCA	6600	50	1.27	9200	74	1.32	99
10 Hag NCA	30 Met NCA	12800	27	1.14	27400	107	1.18	99

Table 1. Synthesis of diblock copolypeptides using $Co(PMe_3)_4$ in THF at 20 °C. [a] First and second monomers added stepwise to the initiator; number indicates equivalents of monomer per $Co(PMe_3)_4$. [b] Molecular weight and dispersity after polymerization of the first monomer determined by ^1H NMR and GPC of derivatized polypeptide. [c] Molecular weight and dispersity after polymerization of the second monomer determined by ^1H NMR and GPC of derivatized copolypeptide. [d] Total isolated yield of diblock copolypeptide. DP = number average degree of polymerization.

After successful polymerization of Hag NCA, the reactivity of **G^{HA}** with a variety of thiols was evaluated. Toward the goal of obtaining water soluble derivatives, oligoethylene glycol and monosaccharide thiols were chosen for these studies (Figure 3). Under optimized conditions, near quantitative thiol-ene functionalization of Hag residues was obtained for all thiols (see SI).[12] For comparison, thiol-ene functionalization of **(*rac*-G^{HA})** was also attempted using similar conditions (see SI). While > 90% thiol conjugation efficiency could be obtained on short **(*rac*-G^{HA})** chains, these derivatives exhibited poor water solubility due to the formation of β-sheet structures (see Figure S7). Contrary to this result, all thiol-ene derivatives of **G^{HA}** were found to possess good water solubility, except for **mEG$_3$-G^{HA}**, which was soluble in organic solvents.

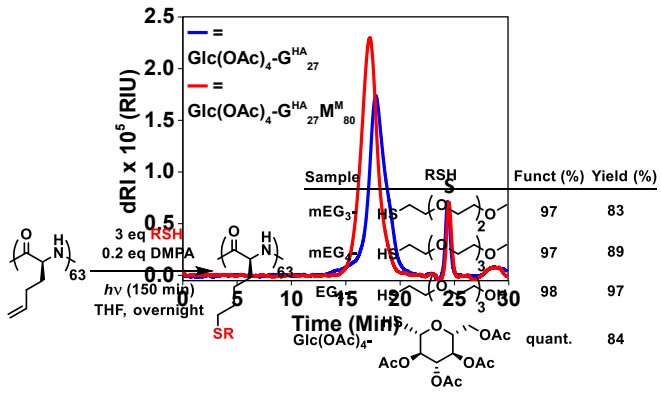
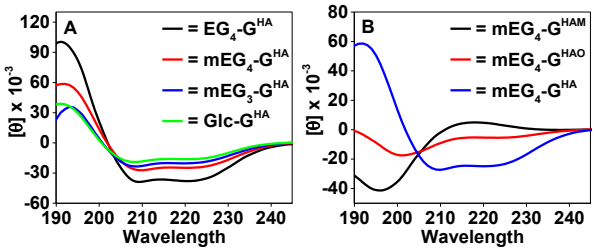


Figure 2. GPC Chromatograms of derivatized homo and diblock polypeptides **Glc(OAc)$_4$-G$^{HA}_{27}$** (blue) and **Glc(OAc)$_4$-G$^{HA}_{27}$M$^{M}_{80}$** (red) in HFIP containing 0.5 % (w/w) KTFA. S = solvent. RIU = arbitrary refractive index units.

Figure 3. Thiol-ene modification of **G^{HA}** (4 mg/mL) in THF with UV irradiation followed by overnight stirring at ambient temperature. Funct = percentage of side-chain modification. Yield = total isolated yield of purified polypeptide. quant. = quantitative

Aqueous solutions of **G^{HA}** derivatives analyzed by circular dichroism (CD) spectroscopy were found to primarily adopt α-helical conformations, similar to the parent **G^{HA}** (Figure 4A). α-Helical content was found to be greatest (*ca.* 100 % α-helix) for the **EG$_4$-G^{HA}** sample, which contained side-chains with greatest hydrophilicity. The methoxy terminated oligoethylene glycol and glycosylated samples (**mEG$_3$-G^{HA}**, **mEG$_4$-G^{HA}**, and **Glc-G^{HA}**) possessed diminished minima at 208 and 222 nm, yet retained partial helical content (49 to 71% α-helix). The addition of hydrophilic thiols to **G^{HA}** was found to be an efficient means to obtain water soluble, α-helical polypeptides with high degrees of functional modification.

Figure 4. Circular dichroism spectra of functionally modified **G$^{HA}_{63}$** samples. (A) Thiol-ene adducts **mEG$_3$-G^{HA}** (blue), **mEG$_4$-G^{HA}** (red), **EG$_4$-G^{HA}** (black), **Glc-G^{HA}** (green). All samples in DI water except **mEG$_3$-G^{HA}** in THF. (B) parent **mEG$_4$-G^{HA}** (blue, 71% α-helix) and its sulfonium (black, 0% α-helix) and sulfoxide (red, 22% α-helix) derivatives in DI water. All samples (0.5 mg/mL) analyzed at 20 °C. Molar ellipticity reported in deg·cm^2/dmol. Percent α-helix content for each sample was calculated

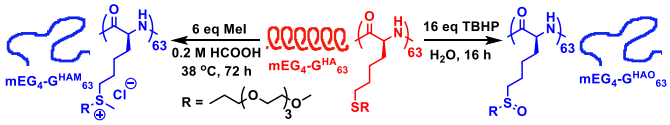


from its molar ellipticity at 222 nm using the formula % α-helix = 100·(-[θ]$_{222}$ + 3000)/39000).[16]

Since functionalized **G^{HA}** contain thioether linkages, there is potential for additional secondary modification of the polypeptide side-chains via selective alkylation or oxidation reactions.[10] To examine the feasibility of such modifications and test their effects on polymer properties, **mEG$_4$-G^{HA}** was reacted separately with either iodomethane or *tert*butylhydroperoxide (TBHP) to obtain the methylated derivative, **mEG$_4$-G^{HAM}**, or oxidized derivative, **mEG$_4$-G^{HAO}**, respectively (Scheme 4). These reactions gave high yields of the fully modified polypeptides, which retained the water solubility of the precursor **mEG$_4$-G^{HA}**. CD analysis of **mEG$_4$-G^{HAM}** and **mEG$_4$-G^{HAO}** in water revealed that both modifications destabilized the α-helical conformation of the parent **mEG$_4$-G^{HA}** (Figure 4B), similar to results obtained for alkylation and oxidation of thioether containing **M** chains even though the thioether groups in **mEG$_4$-G^{HAM}** are two bonds further removed from the peptide backbone compared to Met residues.[10] The degree of conformational disruption was greater for **mEG$_4$-G^{HAM}**, likely due to the introduction of charged groups as compared to the non-ionic sulfoxides in **mEG$_4$-G^{HAO}**. This ability to switch between α-helical and disordered conformations in **mEG$_4$-G^{HA}** polypeptides is a desirable feature that has not been demonstrated in thiol-ene derivatives of other alkene containing polypeptides.

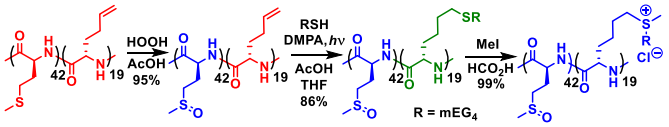
To illustrate how **G^{HA}** segments can be used in conjunction with other polypeptide segments to obtain chains with discrete modified thioether domains, we sought to prepare diblock copolypeptides containing both sulfoxide and sulfonium functionality in separate segments (Scheme 5). Independent control over placement of bio-inert segments, i.e. sulfoxide,[17] and segments that may promote cell uptake, i.e. sulfonium,[18] is

needed for continued development of multifunctional biomaterials. While both sulfoxide and sulfonium groups can be introduced into **M** homopolymers, there is no means to control placement of these groups as they will be statistically distributed along the chains. In our experience, due to limited solubility of **M** in suitable reaction media, precise control over partial oxidation or partial alkylation of **M** chains is challenging. Hence, methodology for facile installation of sulfoxide and sulfonium functionality in discrete segments within copolypeptide sequences would be valuable.



Scheme 4. Conformational changes induced by thioether alkylation or oxidation of **mEG$_4$-G$^{HA}_{63}$**.

To demonstrate the feasibility of such modifications, a block copolymer of Met and Hag, **M$_{42}$G$^{HA}_{19}$** prepared as described above, was subjected to a sequence of selective reactions (Scheme 5). Hydrophobic, α-helical **M$_{42}$G$^{HA}_{19}$** was first oxidized at Met residues to give the amphiphilic copolymer **M$^{O}_{42}$G$^{HA}_{19}$** containing disordered hydrophilic poly(L-methionine sulfoxide), **M^{O}**, segments.[17] The thiol mEG$_4$SH was then selectively added to the Hag residues via radical coupling in acidic media, which is beneficial for thiol-ene conjugation and also prohibits undesirable reduction of sulfoxides by thiols. The resulting copolymer, **M$^{O}_{42}$mEG4-G$^{HA}_{19}$**, now became fully hydrophilic, but retained α-helical conformations in the **mEG4-G^{HA}** domains. The thioether groups in this copolymer were then selectively alkylated using iodomethane, taking advantage of the resistance of **M^{O}** residues toward alkylation under these conditions.[19] The resulting sulfoxide-sulfonium diblock copolypeptide, **M$^{O}_{42}$mEG4-G$^{HAM}_{19}$**, was water soluble and both segments were now conformationally disordered in water. In addition to successful selective functional modification of each copolypeptide domain, the respective thioether modifications also allowed independent conformational switching of each segment (see Figure S8).



Scheme 5. Synthesis of diblock copolypeptide **M$^{O}_{42}$mEG$_4$-G$^{HAM}_{19}$** that contains discrete sulfoxide and sulfonium domains. Percent yields are total isolated yields of purified copolypeptides.

The efficient polymerization of Hag NCA, good solubility of **G^{HA}** allowing preparation of high molecular weight homo- and copolymers, facile modification of Hag residues with thiols, and ability to further modify the thioether products provide a number of attractive features supporting utilization of Hag residues in peptidic materials. Beyond what has been achieved in previous alkene containing polypeptides, the example process in Scheme 5 shows how incorporation of Hag residues into polypeptides can be used to differentially modify discrete segments in an orthogonal manner and also modulate polypeptide chain conformations.

This work was supported by the NSF under MSN 1412367. Mass Spectrometry Instrumentation was made available through the support of Dr. Gregory Khitrov at the University of California, Los Angeles Molecular Instrumentation Center – Mass Spectrometry Facility in the Department of Chemistry. We thank Emma Pelegri-O'Day for assistance in setting up thiol-ene reactions, and Brian Shao and Professor Hosea Nelson (UCLA) for assistance with chiral HPLC studies on Hag amino acids.

Notes and references

1 (a) J. N. deGruyter, L. R. Malins and P. S. Baran, *Biochemistry* 2017, **56**, 3863−3873. (b) T. J. Deming, *Chem. Rev.* 2016, **116**, 786–808. (c) K. Lang and J. W. Chin, *Chem. Rev.* 2014, **114**, 4764-4806.
2 A. Dondoni, *Angew. Chem. Int. Ed.* 2008, **47**, 8995-8997.
3 S. M. Brosnan and H. Schlaad, *Polymer* 2014, **55**, 391.
4 (a) K. Schlögl and H. Fabitschowitz, *Monatsh. Chem.* 1954, **85**, 1060–1076. (b) R. M. Guinn, A. O. Margot, J. R. Taylor, M. Schumacher, D. S. Clark and H. W. Blanch, *Biopolymers*, 1995, **35**, 503–512.
5 (a) J. Sun and H. Schlaad, *Macromolecules*, 2010, **43**, 4445–4448. (b) K.-S. Krannig and H. Schlaad, *J. Amer. Chem. Soc.* 2012, **134**, 18542-18545.
6 H. Tang, L. Yin, H. Lu and J. Cheng, *Biomacromolecules* 2012, **13**, 2609–2615.
7 J. Zhou, P. Chen, C. Deng, F. Meng, R. Cheng and Z. Zhong, *Macromolecules* 2013, **46**, 6723–6730.
8 (a) D. S. Poche, S. J. Thibodeaux, V. C. Rucker, I. M. Warner and W. H. Daly, *Macromolecules* 1997, **30**, 8081–8084. (b) H. Tang and D. Zhang, *Polym. Chem.* 2011, **2**, 1542–1551. (c) Y. Zhang, H. Lu, Y. Lin and J. Cheng, *Macromolecules* 2011, **44**, 6641–6644.
9 W. Wang and P. T. Hammond, *Polym. Chem.* 2018, **9**, 346–351.
10 T. J. Deming, *Bioconjugate Chem.* 2017, **28**, 691−700.
11 T. Hayakawa, Y. Kondo and N. Kobayashi, *Polym. J.* 1975, **7**, 538–543.
12 (a) J. C. M. van Hest and D. A. Tirrell, *FEBS Lett.* 1998, **428**, 68-70. (b) N. Floyd, B. Vijaykrishnan, J. R. Koeppe and B. G. Davis, *Angew. Chem.* 2009, **121**, 7398-7942.
13 (a) S. C. G. Biagini, S. E. Gibson née Thomas and S. P. Keen. *J. Chem. Soc., Perkin Trans. 1* 1998, **0**, 2485-2500. (b) H. K. Chenault, J. Dahmer and G. M. Whitesides, *J. Amer. Chem. Soc.* 1989, **111**, 6354-6364. (c) M. J. O'Donnell and K. Wojciechowski, *Synthesis* 1984, **4**, 313-315.
14 J. R. Kramer and T. J. Deming, *Biomacromolecules* 2010, **11**, 3668 - 3672.
15 T. J. Deming, *Macromolecules* 1999, **32**, 4500-4502.
16 J. A. Morrow, M. L. Segal, S. Lund-Katz, M. C. Philips, M. Knapp, B. Rupp and K. H. Weigraber, *Biochemistry* 2000, **39**, 11657-11666.
17 (a) A. R. Rodriguez, J. R. Kramer and T. J. Deming, *Biomacromolecules* 2013, **14**, 3610-3614. (b) A. L. Wollenberg, T. M. O'Shea, J. H. Kim, A. Czechanski, L. G. Reinholdt, M. V. Sofroniew and T. J. Deming, *Biomaterials* 2018, DOI: 10.1016/j.biomaterials.2018.03.057
18 J. R. Kramer, N. W. Schmidt, K. M. Mayle, D. T. Kamei, G. C. L. Wong and T. J. Deming, *ACS Central Sci.* 2015, **1**, 83-88.
19 J. R. Kramer and T. J. Deming, *Biomacromolecules* 2012, **13**, 1719-1723.