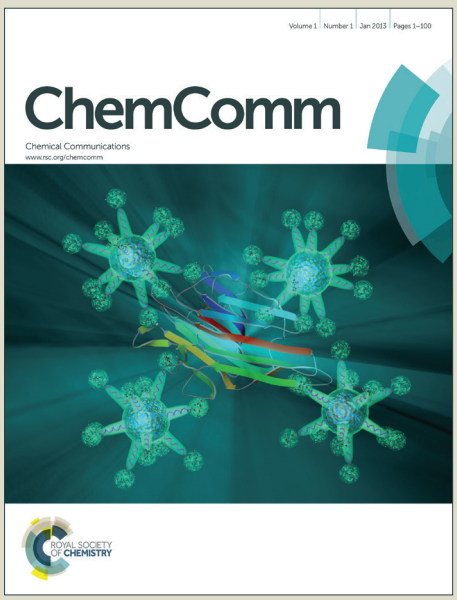
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

ChemComm

COMMUNICATION

Thioamides in the collagen triple helix†

Robert W. Newberry,‡[a] Brett VanVeller‡[ab] and Ronald T. Raines*[ac]

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

DOI: 10.1039/x0xx00000x

www.rsc.org/

To probe noncovalent interactions within the collagen triple helix, backbone amides were replaced with a thioamide isostere. This subtle substitution is the first in the collagen backbone that does not compromise thermostability. A triple helix with a thioamide as a hydrogen bond donor was found to be more stable than triple helices assembled from isomeric thiopeptides.

Collagen is the main structural protein in animals, forming a significant portion of the extracellular matrix and the dry weight of skin. At the core of collagen is a three amino–acid repeat, Xaa-Yaa-Gly,[1] that imparts a unique three-dimensional structure in which three individual polypeptide strands wind into a helix with a single-residue offset (Fig. 1).[2] Each individual strand adopts a polyproline type II conformation. This conformation is devoid of intramolecular hydrogen bonding within the peptide backbone, thereby allowing for intermolecular hydrogen bonds to form between strands.[3] Each of those hydrogen bonds contributes approximately 2 kcal/mol to stability.[4] The unique structure of collagen is enforced further by the high (2*S*)-proline and (2*S*,4*R*)-4-hydroxyproline content of collagen proteins. These amino acids serve not only to template the ideal backbone dihedral angles of the collagen triple helix, but also to effect attraction between adjacent backbone carbonyl groups through an $n\rightarrow\pi^*$ interaction.[5] The interplay of these steric and electronic contributions gives collagen its robust thermal and mechanical stability, and has thus received significant attention.[6]

Most previous studies on collagen stability have relied on the modification of protein side chains.[6-17] The core triple-helical structure of collagen is, however, enforced almost exclusively through backbone–backbone hydrogen bonds and $n\rightarrow\pi^*$ interactions (Fig. 1), which have been probed directly in only a few cases[4,18,19] and indirectly in others.[13,16,20] Those previous studies have employed either esters[4] or alkenes[4,18,19] as surrogates for the peptide bond, and examined the role of hydrogen bonding as well as *cis*/*trans* isomerization within the backbone as contributors to collagen stability. Though informative, the use of ester or alkene surrogates necessarily results in the deletion of donors or acceptors of hydrogen bonds, the alteration of $n\rightarrow\pi^*$ interactions, or the complete restriction of bond rotations.[22] Here we sought to use the most subtle of isosteres to interrogate the backbone of collagen—a thioamide.

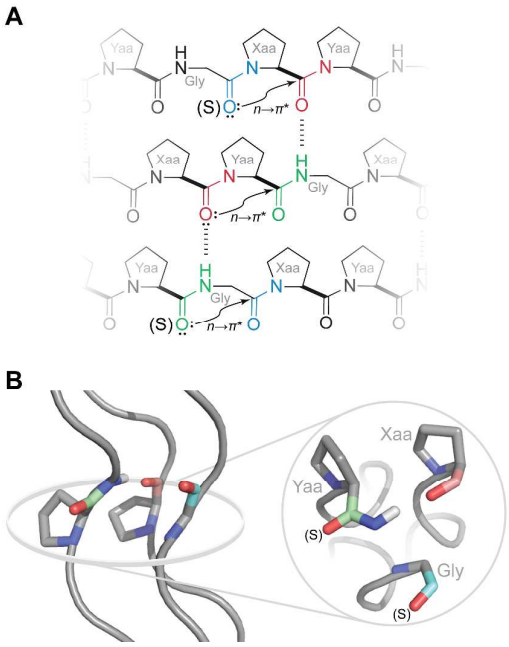


Fig. 1 (A) Schematic diagram and (B) molecular model of the collagen triple helix. $n\rightarrow\pi^*$ Interactions across an Xaa, Yaa, and Gly residue and backbone amides serving as hydrogen bond acceptors (red), hydrogen bond donors (green), or neither (blue) are indicated explicitly. The image in panel B was derived from PDB entry 1v4f.[21] Thioamides at the peptide bonds indicated with an "(S)" were examined in this work.

Replacing the oxygen of a peptide bond with sulfur has several consequences.[22] First, the thiocarbonyl bond is longer than that of a

[a]*Department of Chemistry, University of Wisconsin–Madison, 1101 University Avenue, Madison, WI 53706-1322, USA. Email: rtraines@wisc.edu*
[b]*Current address: Department of Chemistry, Iowa State University, Ames, IA 50011-3111, USA*
[c]*Department of Biochemistry, University of Wisconsin–Madison, 433 Babcock Drive, Madison, WI 53706-1544, USA*
†Electronic Supplementary Information (ESI) available: Experimental procedures and spectral data for novel compounds. See DOI: 10.1039/x0xx00000x
‡These authors contributed equally to this work.

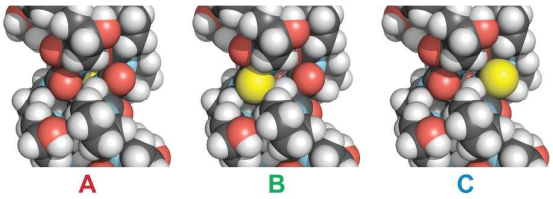


Fig. 2 Space-filling models of collagen triple helices in which a thioamide replaces an (A) Xaa-Yaa, (B) Yaa-Gly, or (C) Gly-Xaa peptide bond. The thioamide sulfur is rendered in yellow in the CPK coloring scheme. Models were made by replacing an oxygen with sulfur in PDB entry 1v4f.[21] Thioamides at the (B) Yaa-Gly and (C) Gly-Xaa peptide bonds were examined in this work.

carbonyl group, which could disturb close atomic packing within the triple helix. Second, the thioamide has a higher rotational barrier than does an oxoamide, owing to greater C–N double bond character.[23-26] Third, thioamides are weaker hydrogen bond acceptors than are oxoamides.[27] Fourth, a thioamide is a stronger hydrogen bond donor than is an oxoamide.[27] Finally, thioamides are known to perturb backbone $n\rightarrow\pi^*$ interactions.[28,29] Thioamides therefore provide a strategy for subtly modulating a variety of effects on collagen stability.

To deconvolute these effects, we have pursued site-specific thioamide-substitution within a collagen mimetic peptide (CMP): $(Pro\text{-}Pro\text{-}Gly)_{10}$ (**O** in Table 1). Collagen has only three types of peptide bonds: Xaa-Yaa, Yaa-Gly, and Gly-Xaa (Fig. 1). The carbonyl of the Xaa-Yaa peptide bond (red in Fig. 1) is directed toward the center of the collagen triple helix, where it accepts a hydrogen bond from another strand. Accordingly, this position is likely to be especially sensitive to the steric effects of thioamide-incorporation. Indeed, our molecular modeling suggested significant steric clashes of the thioamide with the backbone of adjacent strands (Fig. 2). Accordingly, a thioamide is unlikely to be tolerated in this position. On the other hand, the Yaa-Gly peptide bond (green in Fig. 1) serves as a hydrogen bond donor, and its carbonyl group points away from the center of the triple helix (Fig. 2). This position is likely to be relatively insensitive to the steric effects of thioamide-incorporation while being sensitive to electronic effects on hydrogen bonding and the $n\rightarrow\pi^*$ interaction. The Gly-Xaa peptide bond (blue in Fig. 1) is not engaged in a hydrogen bond. Thus, this position should be insensitive to the effect of thioamide-incorporation on hydrogen bonding. Moreover, the carbonyl group of the Gly-Xaa peptide bond is oriented away from the center of the triple helix, which should reduce the effect of thioamide sterics on collagen stability (Fig. 2). Consequently, the Gly-Xaa peptide bond appears to isolate the effect of the thioamide on the $n\rightarrow\pi^*$ interaction.

The synthesis of thiopeptides presents numerous technical challenges. Two routes to thiopeptides are pursued commonly: thioacylation of free amines[30-34] and direct thionation of oxoamides.[35-38] Thioacylation offers advantages in regiospecificity while suffering from lower yields and the difficulty of synthesizing activated thioacylating reagents. Thionation, while easier to implement and often affording higher yields, suffers from idiosyncratic dependencies on steric effects near the thionation site, complicating site-selective incorporation. Ultimately, we combined these synthetic approaches to access our targets.

Table 1 Sequence of CMPs and thermostability of their triple helices

CMP	Sequence	$T_m \pm$ SE (°C)[a]
O	$(PPG)_{10}$	30.5 ± 1.0
Y	$(PPG)_4\text{-}PP^SGPPG\text{-}(PPG)_4$	31.4 ± 1.1
G	$(PPG)_4\text{-}PPG^SPPG\text{-}(PPG)_4$	24.3 ± 0.7
G-N	$PPG\text{-}PPG^SPPG\text{-}(PPG)_7$	24.7 ± 1.9
G-C	$(PPG)_7\text{-}PPG^SPPG\text{-}PPG$	29.5 ± 0.8

[a]Values were obtained in 50 mM sodium phosphate buffer, pH 7.0, as described in Fig. 5.

Motivated by our predictions (*vide supra*), we established four thiopeptides as targets (Table 1). Peptides **Y** and **G** have thioamides replacing the Yaa-Gly and Gly-Xaa peptide bonds, respectively. Peptides **G-N** and **G-C** are controls (*vide infra*). We constructed these four target peptides by first introducing the thioamide into protected amino-acid trimers and then condensing those trimers on a solid phase using Fmoc-based chemistry. We discovered early-on that successful thioamide-incorporation was predicated on placing the thioamide distal to the coupling site of the trimer. Peptide synthesis of trimers containing internal thioamides failed, likely due to attack of the thiocarbonyl sulfur on the activated ester generated during peptide coupling; similar results were observed for the incorporation of thioamides into proteins using native chemical ligation.[39-41] Thus, to realize our CMPs, we pursued the synthesis of tripeptides **1** and **2** as building blocks for peptide synthesis (Fig. 3).

To access tripeptide **1**, we first synthesized the thioacylated 6-nitrobenzotriazole derivative of Fmoc-ProOH.[32] Test reactions with amine nucleophiles indicated, however, the presence of both thioamide and oxoamide products. In contrast, when we performed the same reaction with a thioacetylated triazole prepared from Boc-ProOH, we observed quantitative yield of the thioamide.[42] Hence, we used this route to access to trimer **1** (Fig. 3).

To access tripeptide **2**, we used the known preference of Lawesson's reagent for amides over esters or carbamates,[43] along with previous observations on the inefficiency of thionation of proline residues.[44,45] Accordingly, we employed direct thionation, treating Fmoc-Gly-Pro-Pro-O*t*Bu with excess Lawesson's reagent at elevated temperatures (Fig. 3). We observed only a single, monothionated product. We confirmed its regiochemistry unambiguously with NMR spectroscopy, which demonstrated coupling between the thiocarbonyl carbon and the glycine

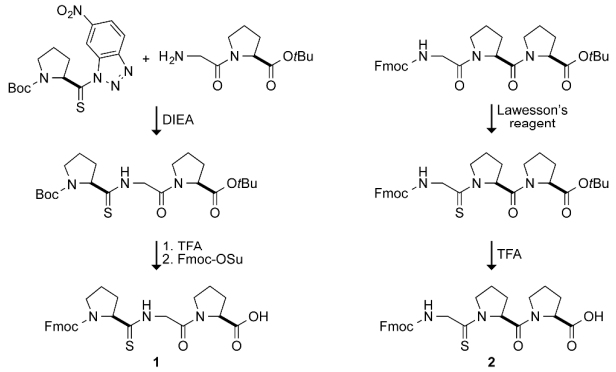


Fig. 3 Synthetic route to tripeptide units **1** and **2**, which were used in the synthesis of thiopeptides **Y** and **G**, respectively.

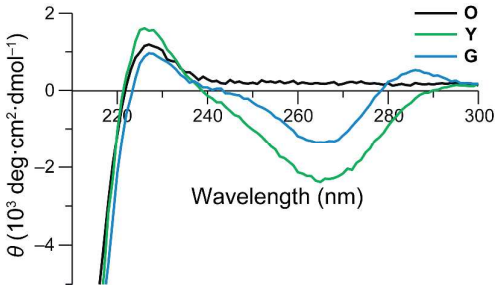


Fig. 4 Representative far-UV circular dichroism spectra of $(Pro\text{-}Pro\text{-}Gly)_{10}$ peptides containing only oxoamides (**O**) or with a thioamide replacing a central Yaa-Gly (**Y**) or Gly-Xaa (**G**) peptide bond. Spectra were recorded at 4 °C in 50 mM sodium phosphate buffer, pH 7.0.

methylene protons in heteronuclear multiple-bond correlation experiments.

With tripeptides **1** and **2** in hand, we synthesized 30-mer thiopeptides **Y** and **G**, respectively, each of which features a single thioamide near its center. We allowed thiopeptides **Y** and **G** to self-assemble overnight in neutral phosphate buffer before examining thermostability with circular dichroism spectroscopy. For comparison, we performed similar analyses on peptide **O**, which does not contain a thioamide. The spectra of all three peptides show the characteristic CD spectra of PPII helices and collagen, featuring a maximum at 226 nm (Fig. 4). In addition, those peptides containing thioamides had additional features in the 250–300 nm range, corresponding to the thioamide electronic absorption. We note three interesting characteristics of these spectra. First, despite outnumbering the thioamide nearly thirty to one, the absorption band produced by the oxoamides is only marginally more intense than that of the thioamide. Second, the thioamide region of the CD spectra of **Y** and **G** differ qualitatively in the folded state, with the spectrum of **G** displaying a maximum at 287 nm in addition to the minimum that both spectra share at 265 nm. Third, no qualitative differences are observed in the spectra of **Y** and **G** in the unfolded state (Fig. S1 in the Supplementary Information). These observations suggest that the CD spectrum of the peptide in the thioamide region might report on local structural changes around the chromophore, as proposed by others.[46,47]

Upon thermal denaturation, we observed that self-assembled thiopeptides **Y** and **G** exhibit cooperative denaturation, a hallmark of triple helices that distinguishes them from other PPII structures. Moreover, the transition temperatures of these peptides were consistent regardless of which spectral feature we tracked (Fig. S1 and S2 in the Supplementary Information). In accord with our predictions, triple-helical **Y** denatured at a temperature 7 °C higher than that of triple-helical **G** (Fig. 5A and Table 1). Again, the thioamide is a stronger hydrogen bond donor than is the oxoamide, so replacing the hydrogen bond donor at the Yaa–Gly position should increase stability of CMPs relative to replacement at the Gly-Xaa position, which does not participate in interstrand hydrogen bonding. Moreover, the thermostabilty of triple-helical **Y** was greater than that of triple-helical **O** (though not significantly so), making a thioamide the first backbone modification of a CMP that does not compromise the stability of a collagen triple helix. This tolerance is meaningful given the extreme constraints imposed by

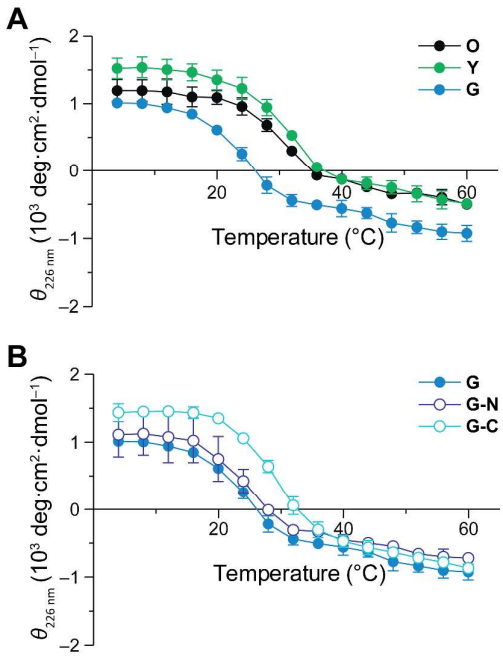


Fig. 5. Effect of thioamide-substitution in the backbone of $(Pro\text{-}Pro\text{-}Gly)_{10}$ peptides on the thermostability of their triple helices. (A) Thermal denaturation data of $(Pro\text{-}Pro\text{-}Gly)_{10}$ containing only oxoamides (**O**) or with a thioamide replacing a central Yaa-Gly (**Y**) or Gly-Xaa (**G**) peptide bond. (B) Thermal denaturation of $(Pro\text{-}Pro\text{-}Gly)_{10}$ with a thioamide replacing a central (**G**), N-terminal (**G-N**), or C-terminal (**G-C**) Gly-Xaa peptide bond. Denaturation was done in 50 mM sodium phosphate buffer, pH 7.0. Values of θ at 226 nm were recorded every 4 °C from 4–60 °C following a 5-min equilibration. Experiments were performed in triplicate. Data were fitted to a two-state model to obtain the values of T_m (± SE) listed in Table 1. For comparison, the data for peptide **G** (blue) are depicted in both panels A and B.

the tight packing of backbone atoms in the center of the triple helix (Fig. 1).

Why does a thioamide at the Gly-Xaa position diminish triple-helical stability despite the absence of a direct effect on interstrand hydrogen bonding? We reasoned that this thioamide could impart a steric clash. To test this hypothesis, we moved the thioamide from the center of the peptide toward the C terminus, resulting in thiopeptide **G-C** (Table 1). Upon thermal denaturation of this peptide, we found that, triple-helical **G-C** had a T_m value close to that of **O** (Fig. 5B and Table 1). Other backbone modifications have been found to be less destabilizing at the C-terminus than elsewhere,[48] consistent with our results. In contrast, moving the thioamide from the center toward the N terminus, as in thiopeptide **G-N** (Table 1), led to no measurable change in thermostability (Fig. 5B). CMPs are thought to fold in a C- to N-terminal direction.[49] We suspect that a disruption in the backbone near the N terminus might reduce the extent of folding, consistent with the lower values of θ observed for **G-N** relative to **G-C** (Fig. 5B). Regardless, these findings motivate additional study of end-effects on helix stability.

In conclusion, our results demonstrate that thioamide-substitution is a conservative approach to the rational modification of the backbone of CMPs. The effect of thioamide-incorporation at different positions within the triplet amino-acid repeat is well

predicted from known hydrogen-bonding properties of thioamides. Importantly, thioamide-substitution has yielded the first backbone-modified collagen triple helix that does not suffer a loss in thermostability. These thiopeptides inspire and enable a variety of future studies. For example, the *cis*/*trans* prolyl peptide bond isomerization during collagen folding could be modulated by exploiting the tendency of thioamides to isomerize upon UV radiation.[24,50,51] In addition, thioamide-containing CMPs could be useful in high-throughput assays of collagen binding by accessing the ability of thioamides to quench pendant fluorophores.[52] Finally, thioamides are protease-resistant,[53] making thioamide-incorporation a potential strategy for producing robust collagen-based biomaterials.[54] Importantly, backbone modification of a CMP allows for the full suite of side-chain modifications, which are well-known to be effective in tuning various properties of collagens.[6,55,56]

This work was supported by grant R01 AR044276 (NIH). R.W.N. was supported by NIH Biotechnology Training Grant T32 GM008349, an ACS Division of Organic Chemistry Graduate Fellowship, and an Eastman Summer Research Award. B.V. was supported by postdoctoral fellowship 289613 (CIHR). This study made use of the National Magnetic Resonance Facility at Madison, which is supported by NIH grant P41 GM103399 (NIH). This work also makes use of the Biophysics Instrumentation Facility at UW–Madison, which was established by grants BIR-9512577 (NSF) and S10 RR13790 (NIH).

Notes and references

1. J. A. M. Ramshaw, N. K. Shah and B. Brodsky, *J. Struct. Biol.*, 1998, 122, 86-91.
2. A. Rich and F. H. C. Crick, *J. Mol. Biol.*, 1961, 3, 483-506.
3. J. Bella, M. Eaton, B. Brodsky and H. M. Berman, *Science*, 1994, 266, 75-81.
4. C. L. Jenkins, M. M. Vasbinder, S. J. Miller and R. T. Raines, *Org. Lett.*, 2005, 7, 2619-2622.
5. L. E. Bretscher, C. L. Jenkins, K. M. Taylor, M. L. DeRider and R. T. Raines, *J. Am. Chem. Soc.*, 2001, 123, 777-778.
6. M. D. Shoulders and R. T. Raines, *Annu. Rev. Biochem.*, 2009, 78, 929-958.
7. S. K. Holmgren, K. M. Taylor, L. E. Bretscher and R. T. Raines, *Nature*, 1998, 392, 666-667.
8. J. A. Hodges and R. T. Raines, *J. Am. Chem. Soc.*, 2003, 125, 9262-9263.
9. M. D. Shoulders, J. A. Hodges and R. T. Raines, *J. Am. Chem. Soc.*, 2006, 128, 8112-8113.
10. M. D. Shoulders, I. A. Guzei and R. T. Raines, *Biopolymers*, 2008, 89, 443-454.
11. F. W. Kotch, I. A. Guzei and R. T. Raines, *J. Am. Chem. Soc.*, 2008, 130, 2952-2953.
12. M. D. Shoulders, K. A. Satyshur, K. T. Forest and R. T. Raines, *Proc. Natl. Acad. Sci. USA*, 2010, 107, 559-564.
13. R. S. Erdmann and H. Wennemers, *Angew. Chem. Int. Ed.*, 2011, 50, 6835-6838.
14. M. D. Shoulders and R. T. Raines, *J. Biol. Chem.*, 2011, 286, 22905-22912.
15. R. S. Erdmann and H. Wennemers, *Org. Biomol. Chem.*, 2012, 10, 1982-1986.
16. R. S. Erdmann and H. Wennemers, *J. Am. Chem. Soc.*, 2012, 134, 17117-17124.
17. C. Siebler, R. S. Erdmann and H. Wennemers, *Angew. Chem. Int. Ed.*, 2014, 53, 10340-10344.
18. N. Dai, X. J. Wang and F. A. Etzkorn, *J. Am. Chem. Soc.*, 2008, 130, 5396-5397.
19. N. Dai and F. A. Etzkorn, *J. Am. Chem. Soc.*, 2009, 131, 13728-13732.
20. M. D. Shoulders, F. W. Kotch, A. Choudhary, I. A. Guzei and R. T. Raines, *J. Am. Chem. Soc.*, 2010, 132, 10857-10865.
21. K. Okuyama, C. Hongo, R. Kukushima, G. Wu, H. Narita, K. Noguchi, Y. Tanaka and N. Nishino, *Biopolymers*, 2004, 76, 367-377.
22. A. Choudhary and R. T. Raines, *ChemBioChem*, 2011, 12, 1801-1807.
23. K. B. Wiberg and P. R. Rablen, *J. Am. Chem. Soc.*, 1995, 117, 2201-2209.
24. R. Frank, M. Jakob, F. Thunecke, G. Fischer and M. Schutkowski, *Angew. Chem. Int. Ed.*, 2000, 39, 1120-1122.
25. K. B. Wiberg and D. J. Rush, *J. Am. Chem. Soc.*, 2001, 123, 2038-2046.
26. K. B. Wiberg and D. J. Rush, *J. Org. Chem.*, 2002, 67, 826-830.
27. H.-J. Lee, Y.-S. Choi, K.-B. Lee, J. Park and C.-J. Yoon, *J. Phys. Chem. A*, 2002, 106, 7010-7017.
28. A. Choudhary, D. Gandla, G. R. Krow and R. T. Raines, *J. Am. Chem. Soc.*, 2009, 131, 7244-7246.
29. R. W. Newberry, B. VanVeller, I. A. Guzei and R. T. Raines, *J. Am. Chem. Soc.*, 2013, 135, 7843-7846.
30. B. Zacharie, G. Sauve and C. Penney, *Tetrahedron*, 1993, 49, 10489-10500.
31. T. Hoeg-Jensen, C. E. Olsen and A. Holm, *J. Org. Chem.*, 1994, 59, 1257-1263.
32. M. A. Shalaby, C. W. Grote and H. Rapoport, *J. Org. Chem.*, 1996, 61, 9045-9048.
33. C. T. Brain, A. Hallett and S. Y. Ko, *J. Org. Chem.*, 1997, 62, 3808-3809.
34. A. R. Katritzky, R. M. Witek, V. Rodriguez-Garcia, P. P. Mohapatra, J. W. Rogers, J. Cusido, A. A. A. Abdel-Fattah and P. J. Steel, *J. Org. Chem.*, 2005, 70, 7866-7881.
35. D. W. Brown, M. M. Campbell and C. V. Walker, *Tetrahedron*, 1983, 39, 1075-1083.
36. J. Jurayj and M. Cushman, *Tetrahedron*, 1992, 48, 8601-8614.
37. T. Sifferlen, M. Rueping, K. Gademann, B. Jaun and D. Seebach, *Helv. Chim. Acta*, 1999, 82, 2067-2093.
38. F. Formaggio, M. Crisma, C. Toniolo and C. Peggion, *Eur. J. Org. Chem.*, 2013, 2013, 3455-3463.
39. S. Batjargal, Y. J. Wang, J. M. Goldberg, R. F. Wissner and E. J. Petersson, *J. Am. Chem. Soc.*, 2012, 134, 9172-9182.
40. R. F. Wissner, S. Batjargal, C. M. Fadzen and E. J. Petersson, *J. Am. Chem. Soc.*, 2013, 135, 6529-6540.
41. S. Batjargal, Y. Huang, Y. J. Wang and E. J. Petersson, *J. Pept. Sci.*, 2014, 20, 87-91.
42. We speculate that stacking interactions between the electron-rich Fmoc group and electron-deficient triazole slowed acylation, allowing for loss of sulfur by hydrolysis.
43. T. Ozturk, R. Ertas and O. Mert, *Chem. Rev.*, 2007, 107, 5210-5278.
44. M. Thorsen, B. Yde, U. Pederson, K. Clausen and S.-O. Lawesson, *Tetrahedron*, 1983, 39, 3429-3435.
45. O. E. Jensen and S.-O. Lawesson, *Tetrahedron*, 1985, 41, 5595-5606.
46. M. Hollosi, E. Kollat, J. Kajtar, M. Kajtar and G. D. Fasman, *Biopolymers*, 1990, 30, 1061-1072.
47. M. De Poli and J. Clayden, *Org. Biomol. Chem.*, 2014, 12, 836-843.
48. Y. S. Chen, C. C. Chen and J. C. Horng, *Biopolymers*, 2011, 96, 60-68.
49. C. M. Stultz, *Protein Sci.*, 2006, 15, 2166-2177.
50. Y. Huang, Z. Cong, L. Yang and S. Dong, *J. Pept. Sci.*, 2008, 14, 1062-1068.
51. H. Bregy, H. Heimgartner and J. Helbing, *J. Phys. Chem. B*, 2009, 113, 1756-1762.
52. J. M. Goldberg, S. Batjargal and E. J. Petersson, *J. Am. Chem. Soc.*, 2010, 132, 14718-14720.
53. S. Yao, R. Zutshi and J. Chmielewski, *Bioorg. Med. Chem. Lett.*, 1998, 8, 699-704.
54. S. Chattopadhyay and R. T. Raines, *Biopolymers*, 2014, 101, 821-833.
55. R. S. Erdmann and H. Wennemers, *J. Am. Chem. Soc.*, 2010, 132, 13957-13959.
56. C. Siebler, R. S. Erdmann and H. Wennemers, *Chimia*, 2013, 67, 891-895.

