This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication. Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal’s standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.
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,† that imparts a unique three-dimensional structure in which three individual polypeptide strands wind into a helix with a single-residue offset (Fig. 1). 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. Each of those hydrogen bonds contributes approximately 2 kcal/mol to stability. The unique structure of collagen is enforced further by the high (2S)-proline and (2S,4R)-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. The interplay of these steric and electronic contributions gives collagen its robust thermal and mechanical stability, and has thus received significant attention.

Most previous studies on collagen stability have relied on the modification of protein side chains. 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 and indirectly in others. Those previous studies have employed either esters or alkenes 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. Here we sought to use the most subtle of isosteres to interrogate the backbone of collagen—a thioamide.

Replacing the oxygen of a peptide bond with sulfur has several consequences. First, the thio carbonyl bond is longer than that of a peptide bond, thereby extending the hydrogen bond distance and altering the hydrogen bond donor and acceptor properties. Thioamides, however, are isosteric with amides in that they are planar and have the same number of valence electrons, which allows for the retention of hydrogen bond properties. In addition, thioamides are significantly weaker hydrogen bond donors than amides and are stronger hydrogen bond acceptors.

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. Thioamides at the peptide bonds indicated with an "(S)" were examined in this work.

†Electronic Supplementary Information (ESI) available: Experimental procedures and spectral data for novel compounds. See DOI: 10.1039/x0x00000x
‡These authors contributed equally to this work.
To deconvolute these effects, we have pursued site-specific thioamide-substitution within a collagen mimetic peptide (CMP): (Pro-Pro-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→π* 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). Accordingly, we employed direct thionation, treating Fmoc-Gly-Pro-Pro-OH 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 thioaryl carbon and the glycine.

Motivated by our predictions (vide supra), we established four thiopptides 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. 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-Pro-OH. 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-Pro-OH, we observed quantitative yield of the thioamide. 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, along with previous observations on the inefficiency of thionation of proline residues. Accordingly, we employed direct thionation, treating Fmoc-Gly-Pro-Pro-OtBu 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 thioaryl carbon and the glycine.

Table 1: Sequence of CMPs and thermostability of their triple helices

<table>
<thead>
<tr>
<th>CMP</th>
<th>Sequence</th>
<th>(T_m) ± SE (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>(PPG)(_{10})</td>
<td>30.5 ± 1.0</td>
</tr>
<tr>
<td>Y</td>
<td>(PPG)(<em>{10})−PPG−PPG−(PPG)(</em>{4})</td>
<td>31.4 ± 1.1</td>
</tr>
<tr>
<td>G</td>
<td>(PPG)(<em>{10})−PPG−PPG−(PPG)(</em>{4})</td>
<td>24.3 ± 0.7</td>
</tr>
<tr>
<td>G-N</td>
<td>PPG−PPG−PPG−(PPG)(_{4})</td>
<td>24.7 ± 1.9</td>
</tr>
<tr>
<td>G-C</td>
<td>(PPG)−PPG−PPG−PPG</td>
<td>29.5 ± 0.8</td>
</tr>
</tbody>
</table>

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

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.

Fig. 3 Synthetic route to tripeptide units 1 and 2, which were used in the synthesis of thiopptides Y and G, respectively.
tolerance is meaningful given the extreme constraints imposed by does not compromise the stability of a collagen triple helix. This making a thioamide the first backbone modification of a CMP that differences are observed in the spectra of minimum that both spectra share at 265 nm. Third, no qualitative band produced by the oxoamides is only marginally more intense than that of the thioamide. Second, the thioamide region of the CD note three interesting characteristics of these spectra. First, despite 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 bond 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 tripeptides 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 thermotability 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 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 tripeptide G-C (Table 1). Upon thermal denaturation of this peptide, we found that, triple-helical G-C had a Tm 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 tripeptide G-N (Table 1), led to no measurable change in thermotability (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

Fig. 4 Representative far-UV circular dichroism spectra of (Pro-Pro-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.

Fig. 5. Effect of thioamide-substitution in the backbone of (Pro-Pro-Gly)10 peptides on the thermotability of their triple helices. (A) Thermal denaturation data of (Pro-Pro-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-Pro-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 Tm (± SE) listed in Table 1. For comparison, the data for peptide O (blue) are depicted in both panels A and B.
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.\textsuperscript{4,5,51} 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.\textsuperscript{52} Finally, thioamides are protease-resistant,\textsuperscript{33} making thioamide-incorporation a potential strategy for producing robust collagen-based biomaterials.\textsuperscript{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.\textsuperscript{4,5,56}

This work was supported by grant RO1 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 was supported by NIH Biotechnology Training Grant T32 GM103399 (NIH). This work was supported by NIH grant R01 AR044276 (NIH). R.W.N. was supported by NIH Biotechnology Training Grant T32 GM103399 (NIH).

Notes and references