

## Protein Folding

International Edition: DOI: 10.1002/anie.201914101  
German Edition: DOI: 10.1002/ange.201914101

## Triple-Helix-Stabilizing Effects in Collagen Model Peptides Containing PPII-Helix-Preorganized Diproline Modules

Andreas Maaßen, Jan M. Gebauer, Elena Theres Abraham, Isabelle Grimm, Jörg-Martin Neudörfl, Ronald Kühne, Ines Neundorf, Ulrich Baumann,\* and Hans-Günther Schmalz\*

**Abstract:** Collagen model peptides (CMPs) serve as tools for understanding stability and function of the collagen triple helix and have a potential for biomedical applications. In the past, interstrand cross-linking or conformational preconditioning of proline units through stereoelectronic effects have been utilized in the design of stabilized CMPs. To further study the effects determining collagen triple helix stability we investigated a series of CMPs containing synthetic diproline-mimicking modules (ProMs), which were preorganized in a PPII-helix-type conformation by a functionalizable intrastrand C<sub>2</sub> bridge. Results of CD-based denaturation studies were correlated with calculated (DFT) conformational preferences of the ProM units, revealing that the relative helix stability is mainly governed by an interplay of main-chain preorganization, ring-flip preference, adaptability, and steric effects. Triple helix integrity was proven by crystal structure analysis and binding to HSP47.

## Introduction

Collagen is the most abundant structural protein in animals, comprising a family of 28 known members differing in their composition and supramolecular assembly. By forming fibrils and networks, this main component of the extracellular matrix resides in skin, bone, and other tissues. Because of its unique biomechanical properties, collagen

guarantees the structural integrity of vertebrates.<sup>[1-3]</sup> In addition, collagen interacts with numerous proteins such as cell-surface receptors or matrix metalloproteinases and is involved in processes such as cell adhesion, proliferation, and extracellular matrix regulation.<sup>[4,5]</sup> The biocompatibility and bioactivity of collagen together with advances in synthetic substitutes<sup>[6,7]</sup> open the path for biomedical applications such as bone grafts,<sup>[8,9]</sup> wound dressings,<sup>[10,11]</sup> engineering of functional tissues,<sup>[12,13]</sup> tendon repair,<sup>[14]</sup> or inhibition of disease-related target proteins.<sup>[4,15]</sup> In the context of diseases, pathological conditions are coined by structural defects and impaired collagen stability (e.g., osteogenesis imperfecta).<sup>[16,17]</sup> A stable structure and correct folding are therefore of fundamental importance for molecular recognition and function of collagen. At the molecular level, collagen forms a right-handed triple helix consisting of three peptide strands, which adopt a left-handed polyproline II (PPII) type helix conformation and are held together by interstrand hydrogen bonds.<sup>[18]</sup> A repetitive unit of three amino acids [Xxx-Yyy-Gly]<sub>n</sub> with a conserved glycine (Gly, G) and all-*trans* amide bonds is observed. In the first two positions (Xxx and Yyy), proline (Pro, P) or (4*R*)-hydroxyproline (Hyp, O) are predominantly found.<sup>[1]</sup> Hyp can be introduced by prolyl-4-hydroxylase-mediated functionalization of proline residues.<sup>[3]</sup> Another member of the machinery involved in the complex biosynthesis of collagen is the chaperone HSP47. This essential heat shock protein transiently and exclusively binds to triple helical collagen to stabilize its conformation.<sup>[3,19]</sup> Upon binding, an arginine–aspartate salt bridge was observed in a model system.<sup>[20]</sup>

Collagen model peptides (CMPs) adhering to the [XxxYyyGly]<sub>n</sub> motif have been used to study the interactome, the structure, and structure–stability relationships of collagen.<sup>[1,5]</sup> Various CMPs with different amino acid substitutions have been prepared using preorganization as a guiding design principle.<sup>[2]</sup> Generally, the folding of three peptide chains into a triple helix is entropically unfavorable. However, if the tendency of the free peptide strands to adopt a PPII-helix-related conformation is increased, the entropic cost for folding is decreased (Figure 1A). Accordingly, the triple-helical state can be promoted by proper preorganization. In this context, the structural preferences (puckering) of the proline pyrrolidine rings deserve special attention (Figure 1B). Crystal structures of proline-rich CMPs like (PPG)<sub>10</sub> underlined that a *Cy-endo* conformation is found in the Xxx position and *Cy-exo* in the Yyy position.<sup>[2,21]</sup> As substituents influence the ring-flip preference and, accord-

[\*] A. Maaßen, I. Grimm, Dr. J.-M. Neudörfl, Prof. Dr. H.-G. Schmalz  
University of Cologne, Department of Chemistry  
Greinstraße 4, 50939 Cologne (Germany)  
E-mail: schmalz@uni-koeln.de

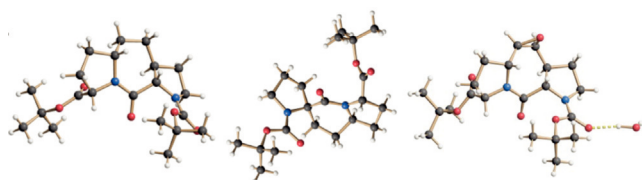
Dr. J. M. Gebauer, E. Theres Abraham, Prof. Dr. I. Neundorf,  
Prof. Dr. U. Baumann  
University of Cologne, Department of Chemistry  
Zùlpicher Straße 47a, 50674 Cologne (Germany)  
E-mail: ubaumann@uni-koeln.de

Dr. R. Kühne  
Leibniz-Institut für Molekulare Pharmakologie (FMP)  
Campus Berlin-Buch  
Robert-Rössle-Straße 10, 13125 Berlin (Germany)

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:  
<https://doi.org/10.1002/anie.201914101>.

© 2020 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

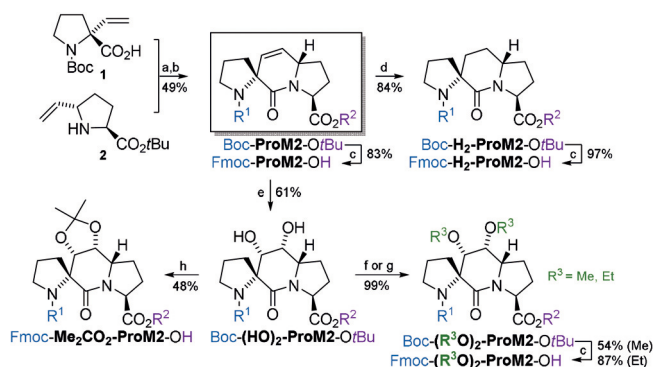




**Figure 3.** Structures of Boc-**H<sub>2</sub>-ProM1**-OtBu, Boc-**H<sub>2</sub>-ProM2**-OtBu, and Boc-**ep-ProM1**-OtBu in the crystalline state.

smoothly under the standard conditions. While other reagents, such as MCPBA, gave unsatisfactory results, epoxidation of the C=C double bond of Boc-**ProM1**-OtBu was achieved by dioxirane generated in situ from oxone<sup>®</sup> and trifluoroacetone at 15 °C.<sup>[31]</sup> Noteworthy, the  $\alpha$ -epoxide (Boc-**ep-ProM1**-OtBu) was obtained as a single diastereomer. The configuration was assigned through NOE NMR experiments and confirmed by crystal-structure analysis (Figure 3). Because of the sensitivity of the epoxide function, triisopropylsilane was added during TFA treatment in the preparation of Fmoc-**ep-ProM1**-OH. Hydroboration of Boc-**ProM1**-OtBu followed by oxidation led to an inseparable mixture of regioisomeric alcohols. However, after reacting this mixture with TBSOTf/NEt<sub>3</sub> at low temperature, the silyl ether Boc-**TBSO-ProM1**-OtBu was isolated as the sole isomer, and its structure was elucidated by NMR analysis (see the Supporting Information, Figure S2). It is noteworthy that the OTBS group remained untouched under the standard trans-protection conditions to give access to Fmoc-**TBSO-ProM1**-OH. First attempts to functionalize the C<sub>2</sub> bridge of Boc-**ProM1**-OtBu through 1,2-dihydroxylation employing Sharpless AD mix ( $\alpha$  or  $\beta$ )<sup>[32]</sup> only resulted in slow conversion, even when using increased amounts of K<sub>2</sub>O<sub>8</sub>(OH)<sub>4</sub>. In contrast, the CeCl<sub>3</sub>-improved method to in situ generate RuO<sub>4</sub> from RuCl<sub>3</sub>/NaIO<sub>4</sub> proved to be successful.<sup>[33]</sup> Upon increasing the RuCl<sub>3</sub> loading from 0.25 mol %<sup>[33]</sup> to 5 mol %, the dihydroxylation proceeded rapidly to diastereoselectively afford the  $\alpha$ -diol Boc-**(HO)<sub>2</sub>-ProM1**-OtBu in 62% yield (besides 11% of the  $\beta$ -isomer; see the Supporting Information). This compound was then smoothly converted into either Fmoc-**(MeO)<sub>2</sub>-ProM1**-OH or Fmoc-**(EtO)<sub>2</sub>-ProM1**-OH by double O-alkylation (NaH, alkyl iodide, DMF) and subsequent protecting group adjustment under the standard conditions. Finally, to avoid free hydroxy groups during SPPS, the diol unit of Boc-**(HO)<sub>2</sub>-ProM1**-OtBu was protected as an acetonide to afford Fmoc-**Me<sub>2</sub>CO<sub>2</sub>-ProM1**-OH after protecting-group exchange.

The synthesis of **ProM2**-derived building blocks is shown in Scheme 2. Boc-**ProM2**-OtBu was prepared from the known building blocks **1** and **2** in an improved two-step protocol through HATU-mediated peptide coupling (in MeCN as a superior solvent) and subsequent Grubbs II-catalyzed ring-closing metathesis. The syntheses of the (crystalline) dihydro derivative Boc-**H<sub>2</sub>-ProM2**-OtBu (see Figure 3), the dihydroxylated product Boc-**(HO)<sub>2</sub>-ProM2**-OtBu, and protected derivatives thereof were accomplished under the conditions elaborated before. Stereochemical assignments were again supported by NMR NOE experiments.



**Scheme 2.** Synthesis of **ProM2**-derived building blocks. Reagents and conditions: a) HATU, (*i*Pr)<sub>2</sub>NEt, MeCN, reflux, 22 h; b) Grubbs II catalyst, CuI, Et<sub>2</sub>O, reflux, 6 h; c) TFA; then FmocCl, NaHCO<sub>3</sub>, H<sub>2</sub>O/THF, 25 °C, 22 h; d) H<sub>2</sub>, Pd/C, MeOH, 25 °C, 10 min; e) RuCl<sub>3</sub>·3 H<sub>2</sub>O, CeCl<sub>3</sub>, NaIO<sub>4</sub>, EtOAc/MeCN/H<sub>2</sub>O, 0 °C, 10 min; f) NaH, MeI, DMF, 0 °C to 25 °C, 2 h; g) NaH, EtI, DMF, 0 °C to 25 °C, 15 h; h) TFA, H<sub>2</sub>O, TIPS; then FmocCl, NaHCO<sub>3</sub>, H<sub>2</sub>O/THF, 25 °C, 2 h; then 2,2-dimethoxypropane, *p*-TsOH, acetone/CH<sub>2</sub>Cl<sub>2</sub>, 25 °C, 6 h. HATU = 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium-3-oxide hexafluorophosphate.

It is noteworthy that epoxidation, hydroboration, and dihydroxylation of both Boc-**ProM1**-OtBu and Boc-**ProM2**-OtBu occurred predominantly from the bottom ( $\alpha$ ) face of the C=C double bond, probably because of steric effects.

### Synthesis of Collagen Model Peptides

Using common SPPS protocols (for details see the Supporting Information), the Fmoc-protected **ProM1** and **ProM2** derivatives were successfully incorporated into the [XY] position of the CMP sequence Ac-(PPG)<sub>5</sub>-PRG-PPG-[XY]G-(PPG)<sub>3</sub>-NH<sub>2</sub> (Table 1). All peptides were purified by HPLC and characterized by ESI-MS (see the Supporting Information for details). The acetal and TBS protecting

**Table 1:** Synthesized CMPs and their experimental transition temperatures.

Entry	Collagen model peptide sequence <sup>[a]</sup>	T <sub>m</sub> [°C]
1	Ac-(PPG) <sub>5</sub> -PRG-PPG- <b>ProPro</b> G-(PPG) <sub>3</sub> -NH <sub>2</sub>	46.5
2	Ac-(PPG) <sub>5</sub> -PRG-PPG- <b>ProHyp</b> G-(PPG) <sub>3</sub> -NH <sub>2</sub> <sup>[b]</sup>	49.8
3	Ac-(PPG) <sub>5</sub> -PRG-PPG- <b>ProM1</b> G-(PPG) <sub>3</sub> -NH <sub>2</sub>	39.3
4	Ac-(PPG) <sub>5</sub> -PRG-PPG- <b>H<sub>2</sub>-ProM1</b> G-(PPG) <sub>3</sub> -NH <sub>2</sub>	39.9
5	Ac-(PPG) <sub>5</sub> -PRG-PPG- <b>HO-ProM1</b> G-(PPG) <sub>3</sub> -NH <sub>2</sub>	43.0
6	Ac-(PPG) <sub>5</sub> -PRG-PPG- <b>(HO)<sub>2</sub>-ProM1</b> G-(PPG) <sub>3</sub> -NH <sub>2</sub>	44.1
7	Ac-(PPG) <sub>5</sub> -PRG-PPG- <b>(MeO)<sub>2</sub>-ProM1</b> G-(PPG) <sub>3</sub> -NH <sub>2</sub>	43.6
8	Ac-(PPG) <sub>5</sub> -PRG-PPG- <b>(EtO)<sub>2</sub>-ProM1</b> G-(PPG) <sub>3</sub> -NH <sub>2</sub>	44.5
9	Ac-(PPG) <sub>5</sub> -PRG-PPG- <b>ProM2</b> G-(PPG) <sub>3</sub> -NH <sub>2</sub>	43.8
10	Ac-(PPG) <sub>5</sub> -PRG-PPG- <b>H<sub>2</sub>-ProM2</b> G-(PPG) <sub>3</sub> -NH <sub>2</sub>	45.1
11	Ac-(PPG) <sub>5</sub> -PRG-PPG- <b>(HO)<sub>2</sub>-ProM2</b> G-(PPG) <sub>3</sub> -NH <sub>2</sub>	37.9
12	Ac-(PPG) <sub>5</sub> -PRG-PPG- <b>(MeO)<sub>2</sub>-ProM2</b> G-(PPG) <sub>3</sub> -NH <sub>2</sub>	40.2
13	Ac-(PPG) <sub>5</sub> -PRG-PPG- <b>(EtO)<sub>2</sub>-ProM2</b> G-(PPG) <sub>3</sub> -NH <sub>2</sub>	38.9
14	Ac-(PPG) <sub>3</sub> - <b>ProM1</b> G-PPG-PRG-(PPG) <sub>5</sub> -NH <sub>2</sub>	40.5
15	Ac-(PPG) <sub>3</sub> - <b>H<sub>2</sub>-ProM2</b> G-PPG-PRG-(PPG) <sub>5</sub> -NH <sub>2</sub>	44.6

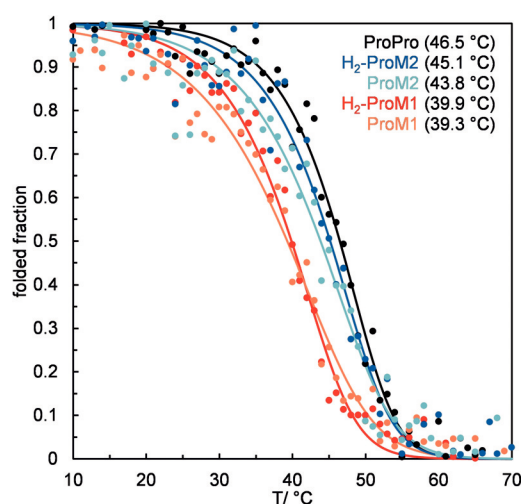
[a] See the Supporting Information for biotinylated or R-free peptides.

[b] Hyp = (2*S*,4*R*)-hydroxyproline.

groups were lost under the acidic conditions (TFA) used to cleave off the peptide from the resin. The MS data of the **ep-ProM1**-derived model peptide indicated that the epoxide function had not survived the cleavage conditions. Thus, the obtained undefined mixture of peptides (containing a *trans*-diol unit) was excluded from further studies. In contrast, the ProM units of all other CMPs remained unaffected according to the analytical data.

### Determination of Triple-Helix Stability

For all prepared model peptides, CD spectra were recorded in phosphate-buffered saline (after incubation for 24 h at 4 °C). Characteristic ellipticity curves and a maximum at  $\lambda = 225$  nm (see Figure S7) indicated a PPII-type helix conformation within a collagen triple-helical architecture in all cases.<sup>[34]</sup> The peptides were then subjected to thermal denaturation studies (see the Supporting Information for details) monitoring the decrease of ellipticity. The curves were fitted to assess triple-helical stability by the transition temperature ( $T_m$ ,  $\pm 1$  °C error) in a standardized fashion (see Table 1 and the Supporting Information). Selected denaturation curves are depicted in Figure 4.

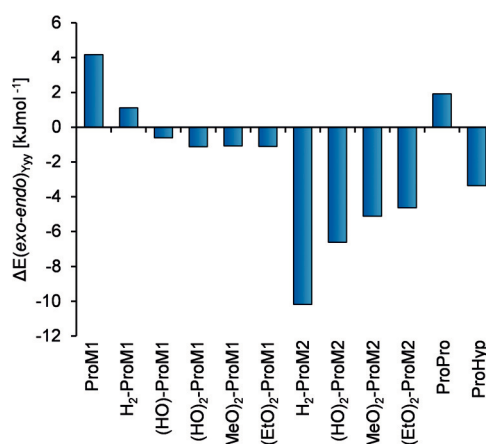


**Figure 4.** Thermal triple-helix denaturation monitored by the decrease in CD ellipticity at 225 nm for ProM-modified peptides Ac-(PPG)<sub>5</sub>-PRG-PPG-[XY]G-(PPG)<sub>3</sub>-NH<sub>2</sub> and corresponding  $T_m$  values (heating rate: 12 °C h<sup>-1</sup>; see the color code for incorporated ProM at [XY] and see the Supporting Information for experimental conditions).

### Computer-Assisted Conformational Analysis

As a precondition for an in-depth interpretation of the relative triple-helical stabilities of the CMPs, we computationally assessed the relevant structural space of all ProM units using conformational searches and subsequent DFT geometry optimizations (B3LYP-D3/6-31G\*, H<sub>2</sub>O polarized continuum model; for details see the Supporting Information). In a minimalistic approach, only the N-acetylated dipeptide methyl esters (Ac-[ProM]-OMe) were evaluated to save computational costs. In addition to this, only collagen-

relevant conformers were selected and then further optimized. For each ProM a set of up to four collagen-relevant conformers was obtained (all-*trans* amides,  $\psi_{Yyy}$  close to 180°, *endo/endo*, *endo/exo*, *exo/endo*, or *exo/exo* puckered rings). Generally, increased thermal stability was anticipated if this set of DFT-optimized structures contained a low-energy conformer that fulfills the structural requirements of the collagen triple helix (preorganization, see above). In (PPG)<sub>10</sub>,  $\varphi_{Xxx} = -75^\circ$ ,  $\psi_{Xxx} = 164^\circ$ ,  $\varphi_{Yyy} = -60^\circ$  and  $\psi_{Yyy} = 152^\circ$  were found as mean values around which the main-chain torsional angles fluctuate.<sup>[21]</sup> Consideration of the main-chain torsional angles of the DFT-optimized, lowest-energy structure of Ac-**ProM1**-OMe ( $\varphi_{Xxx} = -67^\circ$ ,  $\psi_{Xxx} = 170^\circ$ ,  $\varphi_{Yyy} = -74^\circ$ ,  $\psi_{Yyy} = 160^\circ$ ) underlined the preorganization of the main chain. In Table S3, structural parameters of all calculated structures are given, and the deviation from (PPG)<sub>10</sub> average values is depicted in Figure S10. Because of the overall small deviation (< 15°) for each torsion angle, the concept of preorganization in ProMs could be confirmed. As an important aspect, the intrinsic *endo* or *exo* ring-flip preferences were analyzed for all ProMs using the set of DFT-calculated collagen-relevant conformers. As the ring pucker at the Yyy position was found to be crucial for triple-helix stability,<sup>[1]</sup> energy differences only for the Yyy position were determined in the following (see Figure 5).

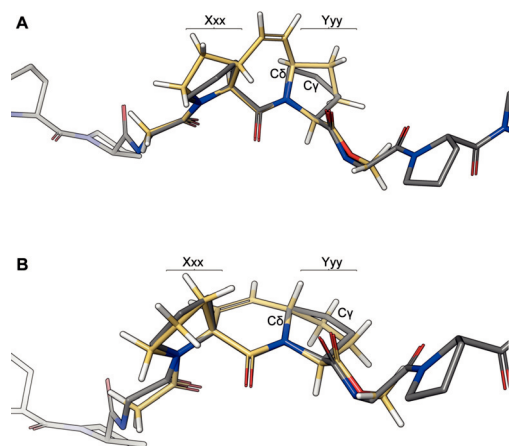


**Figure 5.** Relative DFT energies for the ring flip at the Yyy position (see the Supporting Information for details). Negative values indicate *exo* preference, positive ones *endo* preference.

### Analysis of Structure–Stability Relationships

In Figure 4, thermal denaturation curves of ProM-modified peptides are presented in comparison to Ac-(PPG)<sub>5</sub>-PRG-(PPG)<sub>5</sub>-NH<sub>2</sub> as a reference peptide. Unexpectedly, the incorporation of **H<sub>2</sub>-ProM1** or **ProM1** led to significant destabilization of the triple helix ( $\Delta T_m = -6$  °C). Regarding the reference peptide, the stability was slightly decreased for the **ProM2**-CMP or maintained in the case of the **H<sub>2</sub>-ProM2**-CMP. To address the question why the **ProM1** derivatives significantly impaired triple-helical stability while the **ProM2** derivatives did not, we analyzed the conformational preferences of the two systems with regard to five-ring puckering

(compare Figures 1 and 5). Remarkably, Ac-**ProM1**-OMe and Ac-**H<sub>2</sub>-ProM1**-OMe showed an *endo* preference whereas Ac-**H<sub>2</sub>-ProM2**-OMe exhibited a strong preference for the *exo* state. In the case of Ac-**ProM2**-OMe, no conformer with an *endo* ring flip at Yyy was found in the initial conformational search (30 kJ mol<sup>-1</sup> cutoff), demonstrating the general *exo* propensity of **ProM2** derivatives. Thus, considering the DFT-optimized structures, the different ring-flip preferences can be explained by the configuration at C $\delta$  in the proline ring at the Yyy position. C $\delta$  connects the C $\gamma$  atom with the rigid C<sub>2</sub> bridge. A pseudo-equatorial orientation of the rigid ethylene linker relative to the Yyy ring is sterically favored and thus controls the C $\gamma$  ring flip depending on the configuration at C $\delta$  (*endo* for Ac-**ProM1**-OMe, *exo* for Ac-**ProM2**-OMe; see Figure 6). Comparable steric effects have already been reported for 4-methylproline.<sup>[25,35]</sup>



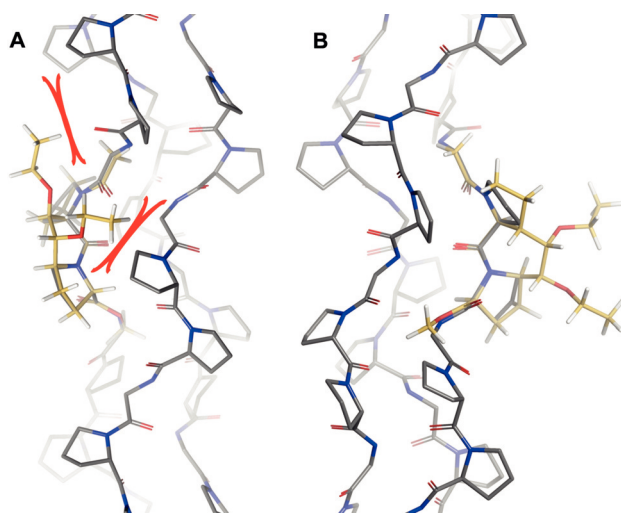
**Figure 6.** Main-chain atom alignment of *exo/endo*-Ac-**ProM1**-OMe (A, yellow) or *endo/exo*-Ac-**ProM2**-OMe (B, yellow) with (PPG)<sub>10</sub> (gray, PDB: 1k6f).<sup>[21]</sup>

Alignment of the calculated lowest-energy structures of Ac-**ProM1**-OMe and Ac-**ProM2**-OMe with the collagen model peptide (PPG)<sub>10</sub> again confirmed the preorganization of the main chain,<sup>[21]</sup> but did not explain the different stabilities of the **ProM1**- and the **ProM2**-CMP. Because of their desired *endo/exo* flip the proline rings of Ac-**ProM2**-OMe were positioned in space that is also occupied by proline rings in a collagen environment (see Figure 6). In fact, this *endolexo* flip could be found in a crystal structure of the **ProM2**-CMP facilitating a close packing of proline residues (see below). However, the *exolendo* preference observed for Ac-**ProM1**-OMe probably led to significant distortion in the corresponding CMP. This distortion could perturb the inter-strand packing (and the van der Waals interactions) of unpolar proline residues at the triple helix periphery, resulting in reduced triple-helix stability.<sup>[36]</sup>

Hydroxy functionalization of proline in general can change conformational properties and can lead to triple-helix stabilization (see above). We therefore used the opportunity to test the influence of linker hydroxylation instead of ring hydroxylation. CD experiments revealed that  $\alpha$ -monohydroxylation or  $\alpha$ -dihydroxylation of the **ProM1**-CMP caused

significant triple-helix stabilization with regard to the non-hydroxylated **ProM1**-CMP (see Figure S8). The  $T_m$  value of the (**HO**)<sub>2</sub>-**ProM1**-modified peptide is close to the one of the reference peptide (XY = PP). To probe whether hydrophilic interactions<sup>[18,37]</sup> between the hydroxylated linker and water were responsible for this stabilizing effect, the (**MeO**)<sub>2</sub>- and (**EtO**)<sub>2</sub>-**ProM1** peptides were compared with the (**HO**)<sub>2</sub>-**ProM1**-CMP. The transition temperature remained unchanged at 44°C. Therefore, the water interaction hypothesis<sup>[18,37]</sup> had to be discarded as alkylation of the hydroxy groups would have been expected to weaken a possible hydrogen-bond interaction. In the literature, an analogous observation was made for (4*R*)-hydroxyproline and (4*R*)-methoxyproline.<sup>[38]</sup> It was concluded that the stabilization in such substitutions rather originates from a stereoelectronic *gauche* effect, which preorganizes the proline ring and requires stereochemically defined substitution with electro-negative groups (-OH, -OMe, -F).<sup>[1,38]</sup> In the case of Ac-**ProM1**-OMe, calculations underlined that the relative position of proline rings to each other was altered upon oxy functionalization of the C<sub>2</sub> bridge (see the torsion angles in Table S3 and Figure S10). Furthermore, the ring pucker preference was inverted from *endo* to *exo* at the Yyy position (see Figure 5). Possibly, oxy functionalization of the C<sub>2</sub> bridge affected the dipeptide conformation and overall triple-helical stability via a *gauche* effect on the ring-linker-ring torsion.

For the **ProM2**-based CMPs, different results were obtained when the interconnecting C<sub>2</sub> linker was HO-, MeO-, or EtO-functionalized. The (**HO**)<sub>2</sub>-**ProM2**-modified peptide showed significant ( $\Delta T_m = -7^\circ\text{C}$ ) destabilization with regard to the non-functionalized system. This was also observed to a slightly lower extent ( $\Delta T_m = -5$  or  $-6^\circ\text{C}$ ) for the (**MeO**)<sub>2</sub>-**ProM2**- and the (**EtO**)<sub>2</sub>-**ProM2**-CMP. In a previous work, an intrastrand H-bond to the main chain was reported to destabilize the triple helix.<sup>[39,40]</sup> DFT calculations in this work demonstrated that  $\alpha$ -dihydroxylation of Ac-**ProM2**-OMe introduced at least one intramolecular H-bond (see Figure S11). Hence, the diproline structure was distorted and notably deviated from the required values for the main chain of a collagen triple helix (see Figures S10), resulting in reduced stability. Although alkylation of the HO groups in the (**HO**)<sub>2</sub>-**ProM2**-CMP was expected to switch off intramolecular H-bonding, the corresponding peptides were still destabilized. As steric repulsion within the triple helix could explain this result,<sup>[39]</sup> the DFT-optimized structures of Ac-(**EtO**)<sub>2</sub>-**ProM1**-OMe and Ac-(**EtO**)<sub>2</sub>-**ProM2**-OMe were aligned with (PPG)<sub>10</sub> to describe the steric situation in the corresponding collagen model peptides. The alignments (Figure 7) showed that the ethoxy groups in Ac-(**EtO**)<sub>2</sub>-**ProM2**-OMe indeed clash with a preceding carbonyl group of the same strand and a proline ring of a neighboring strand. However, the ethoxy groups in Ac-(**EtO**)<sub>2</sub>-**ProM1**-OMe were found to be radially oriented, pointing away from the triple helix axis. In addition to existing methods for triple-helix functionalization,<sup>[41-43]</sup> the (**HO**)<sub>2</sub>-**ProM1** scaffold might be used for double functionalization by O-alkylation. In conclusion, the steric demand of the linker group was determined to be crucial for the thermal stability of ProM-modified CMPs.



**Figure 7.** Main-chain alignment of A) Ac-(EtO)<sub>2</sub>-ProM2-OMe (steric clashes in red) and B) Ac-(EtO)<sub>2</sub>-ProM1-OMe with reference peptide (PPG)<sub>10</sub>.<sup>[21]</sup>

A still remaining question was why most of the ProM-modified collagen peptides (except for **H<sub>2</sub>-ProM2-CMP**) showed a significantly lower  $T_m$  value than the reference peptide (XY = PP) despite their entropic advantage resulting from preorganization in a PPII-helix-type conformation (as confirmed by DFT calculations and structural alignments). The observed trend remained valid upon substituting a different pair of prolines in the model peptide sequence by a ProM unit or upon replacing an arginine by a proline (see Figure S9). Hence, a surprising destabilizing effect seemed to affect all ProM-modified peptides. In diproline, ring movement can rather occur individually when compared to ProMs in which the two pyrrolidine rings are fused to a central ring. Possibly, the two proline rings in ProMs might fail to individually adapt to thermal fluctuations<sup>[44,45]</sup> of the main chain without producing strain in the tricyclic system. Adaptability was therefore proposed to be a relevant factor for triple-helix stability and was assumed to be improvable for the thus far tested ProMs. Generally, the importance of adaptability beside preorganization was highlighted in previous work on linker optimization in fragment-based drug design.<sup>[46]</sup>

In the previous sections, four key aspects (main-chain and ring-flip preorganization, adaptability, steric effects) were identified to be crucial for the triple-helix stability of ProM-modified CMPs. DFT calculations on dipeptide systems and crystal structure alignments were performed to investigate ProMs regarding these key aspects. Guided by literature and DFT-results, the ProMs were evaluated in terms of triple-helix suitability. Significant differences in triple-helix suitability were observed and now allowed us to rank the ProMs for each above-mentioned key aspect (for details, see the Supporting Information). To consider all rankings for all key aspects simultaneously, a scoring system was introduced and optimized. As a result, the theoretical findings could be correlated with experimental data (see Table 2 and the Supporting Information). More investigations will be necessary to enhance the predictive applicability of this simplified,

**Table 2:** Ranking results based on DFT calculations and correlation with  $T_m$  values of the corresponding CMPs.

Ac-[ProM]-OMe	Main chain <sup>[a]</sup>	Ring flip <sup>[b]</sup>	Adaptability <sup>[c]</sup>	Sterics <sup>[d]</sup>	S <sup>[e]</sup>	$T_m$ [°C]
ProHyp	o	+	+	o	++	49.8
ProPro	+	-	+	o	+	46.5
H <sub>2</sub> -ProM2	o	+	o	o	+	45.1
(EtO) <sub>2</sub> -ProM1	+	o	o	o	+	44.5
(HO) <sub>2</sub> -ProM1	+	o	o	o	+	44.1
ProM2	o	+	o	o	+	43.8
(MeO) <sub>2</sub> -ProM1	+	o	o	o	+	43.6
(HO)-ProM1	+	o	o	o	+	43.0
(MeO) <sub>2</sub> -ProM2	o	+	o	-	-	40.2
H <sub>2</sub> -ProM1	o	-	o	o	-	39.9
ProM1	o	-	o	o	-	39.3
(EtO) <sub>2</sub> -ProM2	o	+	o	-	-	38.9
(HO) <sub>2</sub> -ProM2	-	+	o	o	-	37.9

[a] Preorganization of the torsional angles  $\varphi_{X_{xxx}}$ ,  $\psi_{X_{xxx}}$  and  $\varphi_{Y_{yyy}}$ . [b] The *exo* ring flip preference in Yyy position. [c] Adaptability to small torsion-angle fluctuations. [d] Steric toleration in the collagen triple helix. [e] Sum of the results from the four ranking categories. “+” above average; “o” average; “-” below average in terms of triple helix suitability. See the main text and the Supporting Information for more details.

semi-empirical model. However, the model sums up and explains the structure–stability relationships of ProM-modified CMPs described in this work.

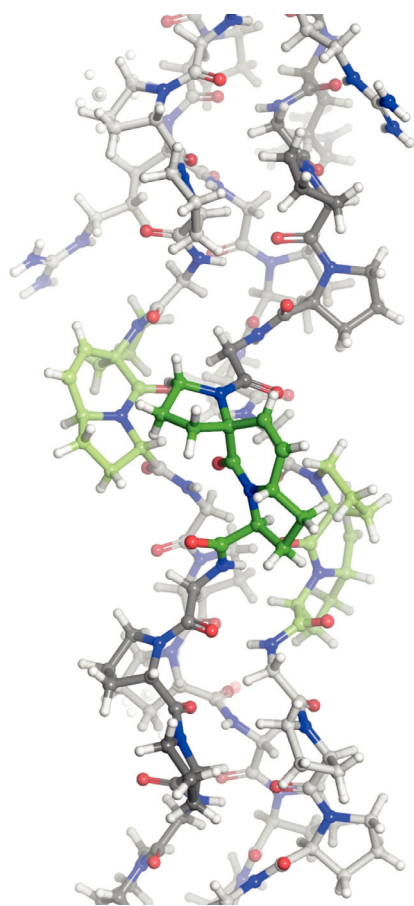
#### Crystal-Structural Analysis of the ProM2-CMP

In addition to this model, more structural insights could be obtained. To our delight, the Ac-(PPG)<sub>5</sub>-PRG-PPG-[ProM2]G-(PPG)<sub>3</sub>-NH<sub>2</sub> peptide (**ProM2-CMP**) crystallized, giving access to an X-ray structure of high resolution (PDB: 6SYJ; see Figure 8 and Tables 3 and S6 for parameters). This is the first crystal structure of an intrastrand-linked collagen model peptide in which two subsequent proline residues are connected.

Regarding the overall structure, the three peptide strands adopted the typical triple-helical conformation, and no distortion could be detected, which was underlined by comparison of key parameters (see Table 3) and alignment with (PPG)<sub>10</sub> (see Figure S15).<sup>[21]</sup> Considering the modified positions, **ProM2** units were assembled in a PPII-type helical shape (like all the other Pro rings) and were well-embedded into the overall structure. In combination with the fact that all **ProM2** units adopted an *endo,exo* ring puckering (as predicted by DFT calculations), a close packing of Pro rings within the triple helix was observed. The crystal structure not only confirmed the synthetic success of introducing covalent linkages in collagen models, but also demonstrated that the triple helical structure was maintained upon this manipulation.

#### Binding of ProM-Modified CMPs to HSP47

To evaluate structural and functional consequences of ProM modifications in collagen models, the interaction with



**Figure 8.** Section of the X-ray crystal structure of Ac-(PPG)<sub>5</sub>-PRG-PPG-[ProM2]G-(PPG)<sub>3</sub>-NH<sub>2</sub> centered on the ProM2 substitution sites (PDB: 6SYJ).

**Table 3:** Comparison of triple helix parameters from crystal structures.

Torsion angle [°]	ProM2-CMP <sup>[a]</sup>	(PPG) <sub>10</sub> <sup>[b]</sup>
$\Phi_{X_{XX}}$	$-73.8 \pm 5.0$	$-74.5 \pm 2.9$
$\Psi_{X_{XX}}$	$163.6 \pm 4.7$	$164.3 \pm 4.1$
$\omega_{X_{XX}}$	$174.2 \pm 3.9$	$176.0 \pm 2.5$
$\Phi_{Y_{YY}}$	$-59.9 \pm 3.5$	$-60.1 \pm 3.6$
$\Psi_{Y_{YY}}$	$151.2 \pm 3.8$	$152.4 \pm 2.6$
$\omega_{Y_{YY}}$	$172.8 \pm 4.1$	$175.4 \pm 3.4$
$\Phi_{Gly}$	$-68.4 \pm 3.9$	$-71.7 \pm 3.7$
$\Psi_{Gly}$	$173.8 \pm 4.2$	$175.9 \pm 3.1$
$\omega_{Gly}$	$178.8 \pm 2.9$	$179.7 \pm 2.0$

[a] Peptide sequence: Ac-(PPG)<sub>5</sub>-PRG-PPG-[ProM2]G-(PPG)<sub>3</sub>-NH<sub>2</sub>; torsion angles were averaged for the central nine amino acid triplets. [b] See Ref. [21].

HSP47 was investigated using biolayer interferometry (BLI). Biotinylated collagen model peptides containing the PRG-binding motif were prepared (see the Supporting Information) and immobilized on streptavidin biosensors. HSP47 binding kinetics were then analyzed to give  $K_D$  values for these peptides (see Tables 4 and S7). First of all, HSP47 was detected to bind to all triple-helical peptides, which is consistent with previous observations.<sup>[47]</sup> Although the (HO)<sub>2</sub>-ProM1 peptide showed a weaker interaction, the  $K_D$  values for all other ProM-modified peptides were very similar

**Table 4:** HSP47 dissociation constants of model peptides.

Peptide, <sup>[a]</sup> XY =	$K_D$ [nM] <sup>[b]</sup>
PP	370.3
ProM2	361.7
H <sub>2</sub> -ProM2	353.0
H <sub>2</sub> -ProM1	268.9
(HO) <sub>2</sub> -ProM1	502.3

[a] Biotin-Ebes-(PPG)<sub>5</sub>-PRG-PPG-[XY]G-(PPG)<sub>3</sub>-NH<sub>2</sub> with XY indicating the (modified) diproline unit. [b] Determined by fit of biolayer interferometry parameters  $k_{on}$ ,  $k_{off}$  and using  $K_D = k_{off}/k_{on}$  (see the Supporting Information). Ebes-OH = N-(8-amino-3,6-dioxaoctyl)succinamic acid.

to that for the reference peptide; they even hinted at a tighter binding. As a result, an HSP47-compatible position for an intrastrand linkage could be identified in this work. These modified peptides were recognized by HSP47 and therefore all formed a regular triple helix.

## Conclusion

In summary, novel ProM derivatives have been synthesized that differ in the linker motif connecting two consecutive proline rings. The stereochemistry of the linker oxy functionalization was dominated by substrate control. By using Fmoc/*t*Bu-based SPPS, the ProMs could be incorporated into collagen model peptides, which folded into triple helices as confirmed by CD. As a result, a method for intrastrand diproline linkage could be developed, which complements literature-known options for structural modifications in collagen. Previous studies were mainly focused on 4-substituted proline derivatives. Relations between structural modification and effects on collagen triple-helix stability have been elucidated. Theoretical and experimental results were successfully correlated to explain such relations in the context of ProM-modified CMPs by considering the interplay of four key aspects (main-chain and ring-flip preorganization, adaptability, and steric effects). Although the adaptability of ProM-modified CMPs remains improvable, the H<sub>2</sub>-ProM2-modified peptide was as stable as the reference system. An X-ray crystal structure and an HSP47 binding assay confirmed the structural and functional integrity of ProM-modified CMPs. Our strategy of using intrastrand diproline linkages might be attractive for future studies on proteolytic stability, folding kinetics, or CMP functionalization. Generally, our structure–stability model can support the design of proline-rich collagen-based biomaterials.<sup>[48]</sup>

## Acknowledgements

This work was financially supported by the Fonds der Chemischen Industrie (doctoral fellowship to A.M.) and the University of Cologne. E.T.A. has been funded by the German Research Foundation (DFG), CRC 829 “Molecular Mechanisms Regulating Skin Homeostasis”, project B11. We thank the staff of the beamline P13 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg,

Germany) for their assistance during data collection, and iNEXT for funding (HORIZON 2020 program of the EU, pid 1859,7162). Crystals were grown in the Cologne Crystallisation facility (<http://C2f.uni-koeln.de>, supported by the German Research Foundation (DFG) grant INST 216/682-1 FUGG). CD and BLITZ measurements were performed at the Protein Interaction Platform Cologne (<http://PIPC.uni-koeln.de>). We also thank Evonik for a gift of proline and Prof. Jordi Bella for fruitful discussions.

### Conflict of interest

The authors declare no conflict of interest.

**Keywords:** collagen · HSP47 · peptidomimetics · protein folding · triple helix stability

**How to cite:** *Angew. Chem. Int. Ed.* **2020**, *59*, 5747–5755  
*Angew. Chem.* **2020**, *132*, 5796–5804

- [1] M. D. Shoulders, R. T. Raines, *Annu. Rev. Biochem.* **2009**, *78*, 929–958.
- [2] J. Bella, *Biochem. J.* **2016**, *473*, 1001–1025.
- [3] J. Brinckmann, H. Notbohm, P. K. Müller, *Collagen. Primer in structure, processing, and assembly*, Springer, Berlin, New York, **2005**.
- [4] B. An, Y.-S. Lin, B. Brodsky, *Adv. Drug Delivery Rev.* **2016**, *97*, 69–84.
- [5] G. B. Fields, *Org. Biomol. Chem.* **2010**, *8*, 1237–1258.
- [6] I. C. Tanrikulu, A. Forticaux, S. Jin, R. T. Raines, *Nat. Chem.* **2016**, *8*, 1008–1014.
- [7] L. E. R. O'Leary, J. A. Fallas, E. L. Bakota, M. K. Kang, J. D. Hartgerink, *Nat. Chem.* **2011**, *3*, 821–828.
- [8] C. F. Marques, G. S. Diogo, S. Pina, J. M. Oliveira, T. H. Silva, R. L. Reis, *J. Mater. Sci. Mater. Med.* **2019**, *30*, p. 32.
- [9] J. A. Inzana, D. Olvera, S. M. Fuller, J. P. Kelly, O. A. Graeve, E. M. Schwarz, S. L. Kates, H. A. Awad, *Biomaterials* **2014**, *35*, 4026–4034.
- [10] G. D. Mogoşanu, A. M. Grumezescu, *Int. J. Pharm.* **2014**, *463*, 127–136.
- [11] Y. Wang, J. Beekman, J. Hew, S. Jackson, A. C. Issler-Fisher, R. Parungao, S. S. Lajevardi, Z. Li, P. K. M. Maitz, *Adv. Drug Delivery Rev.* **2018**, *123*, 3–17.
- [12] B. Sarkar, P. K. Nguyen, W. Gao, A. Dondapati, Z. Siddiqui, V. A. Kumar, *Biomacromolecules* **2018**, *19*, 3597–3611.
- [13] T. R. Chan, P. J. Stahl, S. M. Yu, *Adv. Funct. Mater.* **2011**, *21*, 4252–4262.
- [14] A. J. Lomas, C. N. M. Ryan, A. Soroushanova, N. Shologu, A. I. Sideri, V. Tsioli, G. C. Fthenakis, A. Tzora, I. Skoufos, L. R. Quinlan, G. O'Laighin, A. M. Mullen, J. L. Kelly, S. Kearns, M. Biggs, A. Pandit, D. I. Zeugolis, *Adv. Drug Delivery Rev.* **2015**, *84*, 257–277.
- [15] M. Bhowmick, D. Tokmina-Roszyk, L. Onwuha-Ekpete, K. Harmon, T. Robichaud, R. Fuerst, R. Stawikowska, B. Steffensen, W. Roush, H. R. Wong, G. B. Fields, *J. Med. Chem.* **2017**, *60*, 3814–3827.
- [16] A. Forlino, W. A. Cabral, A. M. Barnes, J. C. Marini, *Nat. Rev. Endocrinol.* **2011**, *7*, 540.
- [17] M. A. Bryan, H. Cheng, B. Brodsky, *Biopolymers* **2011**, *96*, 4–13.
- [18] J. Bella, M. Eaton, B. Brodsky, H. Berman, *Science* **1994**, *266*, 75–81.
- [19] S. Ito, K. Nagata, *Semin. Cell Dev. Biol.* **2017**, *62*, 142–151.
- [20] C. Widmer, J. M. Gebauer, E. Brunstein, S. Rosenbaum, F. Zaucke, C. Drögemüller, T. Leeb, U. Baumann, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 13243–13247.
- [21] R. Berisio, L. Vitagliano, L. Mazzarella, A. Zagari, *Protein Sci.* **2002**, *11*, 262–270.
- [22] B. K. Ho, E. A. Coutasias, C. Seok, K. A. Dill, *Protein Sci.* **2005**, *14*, 1011–1018.
- [23] L. Vitagliano, R. Berisio, L. Mazzarella, A. Zagari, *Biopolymers* **2001**, *58*, 459–464.
- [24] S. K. Holmgren, L. E. Bretscher, K. M. Taylor, R. T. Raines, *Chem. Biol.* **1999**, *6*, 63–70.
- [25] M. D. Shoulders, K. A. Satyshur, K. T. Forest, R. T. Raines, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 559–564.
- [26] N. B. Hentzen, L. E. J. Smeenk, J. Witek, S. Riniker, H. Wennemers, *J. Am. Chem. Soc.* **2017**, *139*, 12815–12820.
- [27] J. Zaminer, C. Brockmann, P. Huy, R. Opitz, C. Reuter, M. Beyermann, C. Freund, M. Müller, H. Oschkinat, R. Kühne, H.-G. Schmalz, *Angew. Chem. Int. Ed.* **2010**, *49*, 7111–7115; *Angew. Chem.* **2010**, *122*, 7265–7269.
- [28] C. Reuter, P. Huy, J.-M. Neudörfl, R. Kühne, H.-G. Schmalz, *Chem. Eur. J.* **2011**, *17*, 12037–12044.
- [29] C. Reuter, R. Opitz, A. Soicke, S. Dohmen, M. Barone, S. Chiha, M. T. Klein, J.-M. Neudörfl, R. Kühne, H.-G. Schmalz, *Chem. Eur. J.* **2015**, *21*, 8464–8470.
- [30] R. Opitz, M. Müller, C. Reuter, M. Barone, A. Soicke, Y. Roske, K. Piotukh, P. Huy, M. Beerbaum, B. Wiesner, M. Beyermann, P. Schmieder, C. Freund, R. Volkmer, H. Oschkinat, H.-G. Schmalz, R. Kühne, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 5011–5016.
- [31] J. A. Díez, J. A. Gálvez, M. D. Díaz-de-Villegas, R. Badorrey, B. Bartholomew, R. J. Nash, *Org. Biomol. Chem.* **2012**, *10*, 9278–9286.
- [32] K. B. Sharpless, W. Amberg, Y. L. Bennani, G. A. Crispino, J. Hartung, K. S. Jeong, H. L. Kwong, K. Morikawa, Z. M. Wang, *J. Org. Chem.* **1992**, *57*, 2768–2771.
- [33] B. Plietker, M. Niggemann, *J. Org. Chem.* **2005**, *70*, 2402–2405.
- [34] J. L. S. Lopes, A. J. Miles, L. Whitmore, B. A. Wallace, *Protein Sci.* **2014**, *23*, 1765–1772.
- [35] M. D. Shoulders, J. A. Hodges, R. T. Raines, *J. Am. Chem. Soc.* **2006**, *128*, 8112–8113.
- [36] R. Improta, F. Mele, O. Crescenzi, C. Benzi, V. Barone, *J. Am. Chem. Soc.* **2002**, *124*, 7857–7865.
- [37] K. Kawahara, Y. Nishi, S. Nakamura, S. Uchiyama, Y. Nishiuchi, T. Nakazawa, T. Ohkubo, Y. Kobayashi, *Biochemistry* **2005**, *44*, 15812–15822.
- [38] F. W. Kotch, I. A. Guzei, R. T. Raines, *J. Am. Chem. Soc.* **2008**, *130*, 2952–2953.
- [39] R. S. Erdmann, H. Wennemers, *J. Am. Chem. Soc.* **2012**, *134*, 17117–17124.
- [40] C. Siebler, R. S. Erdmann, H. Wennemers, *Angew. Chem.* **2014**, *126*, 10508–10512.
- [41] R. S. Erdmann, H. Wennemers, *J. Am. Chem. Soc.* **2010**, *132*, 13957–13959.
- [42] M. R. Aronoff, J. Egli, M. Menichelli, H. Wennemers, *Angew. Chem. Int. Ed.* **2019**, *58*, 3143–3146; *Angew. Chem.* **2019**, *131*, 3175–3178.
- [43] J. Egli, C. Siebler, M. Köhler, R. Zenobi, H. Wennemers, *J. Am. Chem. Soc.* **2019**, *141*, 5607–5611.
- [44] W. Y. Chow, C. J. Forman, D. Bihan, A. M. Puszkarska, R. Rajan, D. G. Reid, D. A. Slatter, L. J. Colwell, D. J. Wales, R. W. Farndale, M. J. Duer, *Sci. Rep.* **2018**, *8*, 13809.
- [45] M. W. H. Kirkness, K. Lehmann, N. R. Forde, arXiv **2019**, arXiv:1903.07276 [q-bio.BM].
- [46] S. Chung, J. B. Parker, M. Bianchet, L. M. Amzel, J. T. Stivers, *Nat. Chem. Biol.* **2009**, *5*, 407–413.

- [47] S. Oecal, E. Socher, M. Uthoff, C. Ernst, F. Zaucke, H. Sticht, U. Baumann, J. M. Gebauer, *J. Biol. Chem.* **2016**, *291*, 12612–12626.
- [48] X-ray structures were deposited at the Cambridge Crystallographic Data Centre and at the Worldwide Protein Data Bank (wwPDB). CCDC 1963764, 1963766, and 1963765 (Boc-**H<sub>2</sub>**-**ProM1**-OtBu, Boc-**H<sub>2</sub>**-**ProM2**-OtBu, and Boc-**ep-ProM1**-OtBu, respectively) contain the supplementary crystallographic data for this paper. These data can be obtained free of charge from

The Cambridge Crystallographic Data Centre. The structure of the **ProM2**-CMP-based triple helix can be found in the PDB protein data bank under the number 6SYJ.

Manuscript received: November 5, 2019  
Accepted manuscript online: January 16, 2020  
Version of record online: February 3, 2020