Inviting Trifluoromethylated Pseudoprolines into Collagen Model Peptides

Anaïs Terrien, Keyvan Rahgoshay, Emelyne Renaglia, Nathalie Lensen, Yves Jacquot, Rodrigue Marquant, Thierry Brigaud, Claire Loison, Grégory Chaume, Emeric Miclet

Inviting Trifluoromethylated Pseudoprolines into Collagen Model Peptides

Anaïs Terrien,[a] Keyvan Rahgoshay,[b] Emelyne Renaglia,[b] Nathalie Lensen,[b] Yves Jacquot,[c] Rodrigue Marquant,[a] Thierry Brigaud,[b] Claire Loison,[d] Grégory Chaume,*[a,b] Emeric Miclet*[a]

[a] Dr. A. Terrien, R. Marquant, Dr. G. Chaume, Dr. E. Miclet
Laboratoire des biomolécules, LBM, École normale supérieure, PSL University, Sorbonne Université, CNRS, Paris (France)
E-mail: emeric.miclet@sorbonne-universite.fr
[b] Dr. K. Rahgoshay, Dr. N. Lensen, Prof. T. Brigaud, Dr. G. Chaume
CY Cergy Paris Université, CNRS, BioCIS, Cergy-Pontoise (France)
E-mail: gregory.chaume@cyu.fr
[c] Prof. Y. Jacquot
Cibles Thérapeutiques et Conception de Médicaments (CiTCoM), CNRS, INSERM, Université Paris Descartes, 75270 Paris Cedex 06 (France)
[d] Dr. C. Loison
Institut Lumière Matière, Univ. Lyon, Université Claude Bernard Lyon1, CNRS, 69622 Villeurbanne (France)

Abstract: Numerous Collagen Model Peptides (CMPs) have been engineered using proline derivatives substituted at their C(3) and/or C(4) position in order to stabilize or to functionalize collagen triple helix mimics. However, no example has been reported so far with C(5) substitutions. Here, we introduce a fluorinated CMP incorporating trifluoromethyl groups at the C(5) position of pseudoproline residues. In tripeptide models, our NMR and Molecular Dynamics (MD) studies have shown that, when properly arranged, these residues meet the structural requirements for triple helix assembly. A host-guest CMP could be synthesized and its NMR analysis in solution confirmed the presence of structured homotrimers that we interpret as triple helices. MD calculations showed that the triple helix model remained stable throughout the simulation, with all six trifluoromethyl groups pointing outwards from the triple helix. Pseudoprolines substituted at the C(5) positions appeared as valuable tools for the design of new fluorinated collagen mimicking peptides.

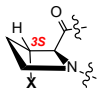
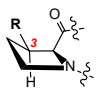
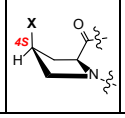
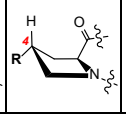
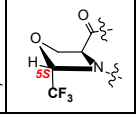
Introduction

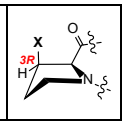
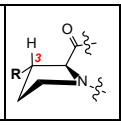
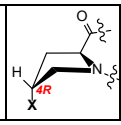
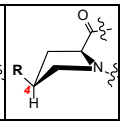
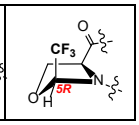
Collagen is the most abundant structural protein in mammals. It represents the major component of the extracellular matrix and is widespread all over the body (bones, skin, tendons, cartilages). Purified collagen can be engineered to serve as a valuable material for biomedical applications.[1] Alternatively, collagen model peptides (CMP) are synthesized to mimic the structural properties of natural collagen[2] or to give it specific functions.[3] At a molecular level, its tertiary structure consists of three individual left-handed polyproline II helices (PPII) folded into a right-handed triple helix which is characterized by a single-residue offset between the three peptide chains and is stabilized by inter-strand hydrogen bonds. Each strand comprises the repeat of the (Xaa-Yaa-Gly) primary sequence where the (2*S*)-proline (Pro) and the (2*S*,4*R*)-4-hydroxyproline (Hyp) are the most prevalent residues at Xaa and Yaa positions, respectively. One of the reasons for this prevalence is that amino acid rings can adopt two conformational states, the C(4)-endo and C(4)-exo ring puckers, which preorganize the suitable ϕ and ψ dihedral angles for the triple helix formation.[4] Crystal structures of short-length Collagen Model Peptides (CMPs) show a distinct preference for the conformation of the amino acid rings, depending on their position in the collagen chain. C(4)-endo and C(4)-exo puckering are preferred at the Xaa and Yaa positions, respectively.[5] This leads to dissimilar couples of dihedral angles with ($\phi = -73°$, $\psi = 163°$) on average for the Xaa site and ($\phi = -58°$, $\psi = 151°$) for the Yaa. Such a puckering alternation has been proposed as a key feature to stabilize the triple helix and accounts for the high occurrence of the non-proteinogenic Hyp residue at the Yaa position in collagen sequences.[6] While Pro residue is known to slightly favor the C(4)-endo conformation, Hyp residue exhibits a strong preference for the C(4)-exo ring pucker due to a gauche effect of the hydroxyl group. Therefore, replacing the Pro by a Hyp residue at the Yaa position allows a significant improvement of the triple helix thermal stability characterized by an increase of its melting temperature (T_m).[2]

Following this rationale, numerous CMPs have been engineered using proline derivatives bearing small substituents that stabilize a definite ring pucker through stereo-electronic and/or steric effects (Table 1). For electron-withdrawing substituents, the pseudoaxial orientation is preferred due to the gauche effect. Consequently, (4*S*) or (4*R*) configurations were respectively selected to stabilize the C(4)-endo at the Xaa position or the C(4)-exo at the Yaa position.[7] For the non-electron-withdrawing groups, the pseudoequatorial orientation is preferred due to steric repulsions.[8] Accordingly, a CMP series has been designed with (4*R*)- or (4*S*)-methylproline at the Xaa and Yaa positions, respectively.

Marked stabilizations of the triple helix were observed for CMP sequences incorporating exo inducer at the Yaa position. In contrast, stabilizing the endo pucker at the Xaa position only slightly improved the CMP thermal stability because this pucker also destabilized the *trans* conformation of the peptide bond which

Table 1. *Puckering preferences of C(3)-, C(4)- and C(5)-substituted proline derivatives.*

C(4)-pucker	*C(3)*-substitutions *ref: 2, 8, 11, 13*		*C(4)*-substitutions *ref: 2, 7, 8, 10, 12*		*C(5)*-substitutions ***this work***
C(4)-endo					

C(4)-exo					

X accounts for electron-withdrawing substituents as found in hydroxyproline, fluoroproline, chloroproline, methoxyproline, or azidoproline. R accounts for non-electron-withdrawing groups: alkyl, vinyl, or phenyl.

is mandatory in the triple helix assembly.[9] Some works of Wennemers et al. demonstrated an additional requirement for the stability of the triple helix when bulky C4-amidated prolines were engaged.[10] They showed that these large electron-withdrawing C4-substituents strongly destabilized the triple helix when grafted on a C4-endo puckered ring at the Xaa position or on a C4-exo puckered ring at the Yaa position. In these examples, steric repulsions were thus able to counteract the benefit of the dihedral angle preorganization achieved by the pyrrolidine puckerings.
Finally, the simultaneous incorporation of C(4)-substituted prolines at the Xaa and at the Yaa positions was almost always detrimental for the triple-helix formation. Steric clashes between the Xaa and Yaa substituents, or adverse dipolar interactions of small electron-withdrawing groups were responsible for the triple helix disruption.[11] The methylproline is one of the rare examples in which a double incorporation ((4*R*) and (4*S*) configurations at Xaa and Yaa positions, respectively) led to an increased T_m.[12]
It is therefore a real challenge to engineer collagen triple helix models in which sequential substituted prolines are present. Raines and coworkers have introduced 3S-fluoroproline (3S-flp) at the Xaa position in CMP to allow a tight control of the ring pucker without creating repulsive contacts between neighboring fluoro groups.[11] The (3*S*-Flp-4*R*-Flp-Gly)$_7$ CMP was then able to undergo triple helix formation whereas (4*S*-Flp-4*R*-Flp-Gly)$_7$ failed because of F-F interactions. However, the T_m value was deceptively low because 3*S*-flp weakens the hydrogen bond network within the triple helix, the electron-withdrawing fluoro group being close to the H-bond acceptor C=O. Thus, electronegative substituents at the 3-position are not good candidates to stabilize the triple helix.[11,13]

So far, no C5-substituted proline analogues have been incorporated in CMP sequences. We envisioned that they represent valuable tools for the design of triple helix models and could be placed sequentially in the CMP sequence. Our group has developed convenient methodologies for the synthesis of various trifluoromethylated pseudoprolines bearing a CF_3 group at the C(5) position and their incorporation into peptides.[14] Their marked stereoelectronic properties have been demonstrated and are very relevant for the design of CMP.[15] Nuclear Magnetic Resonance (NMR) studies as well as theoretical calculations showed that both (5*S*)-CF_3-ΨPro (tfmΨP) and (5*R*)-CF_3-ΨPro (TfmΨP) stabilize the PPII backbone conformation in water and efficiently constrain the O(4)-endo and O(4)-exo ring puckers, respectively (Table 1). Moreover, TfmΨP substantially decreases the *cis-trans* rotational barrier which appears as a molecular hinge for the CMP assembly. Therefore, we anticipated that tfmΨP and TfmΨP pseudoproline residues could be excellent mimics of proline and hydroxyproline, respectively. In this work, our Molecular Dynamics simulations (MD) and NMR studies showed that the trifluoromethylated pseudoprolines meet the structural requirements for the triple helix assembly, even in peptide sequences incorporating two adjacent proline derivatives.

Results and Discussion

Structural Studies of tripeptide models. Preliminary structural studies were performed to evaluate the capabilities of trifluoromethylated pseudoprolines to fulfill the conformational features for CMP assemblies: *trans* amide bond, Xaa-endo/Yaa-exo puckering alternation, suitable ϕ and ψ dihedral angles (Figure 1). We started our investigation with three tripeptides incorporating one or two trifluoromethylated pseudoprolines. This series was based on the Xaa-Yaa-Gly triads with acetyl and carboxamide end groups, where Xaa = Pro or tfmΨP and Yaa = Hyp or TfmΨP. In addition, Ac-Pro-Pro-Gly-NH_2 (**P1**) and and Ac-Pro-Hyp-Gly-NH_2 (**P2**) have been studied as references. Peptides **P1**, **P2** and **P4** were obtained by solid phase peptide synthesis (SPPS) whereas the synthesis of peptides **P3** and **P5** was performed in solution since coupling reactions at the *N*-terminus of the trifluoromethylated pseudoproline TfmΨP requires the activation of the Fmoc-amino acids partners as acyl chlorides (see the Supporting Information for details).[14] To determine the structural propensity of peptides **P1-P5**, we undertook a detailed study combining solution NMR and MD simulations.

Peptide bond geometry. The *cis* or *trans* propensity of the two amide bonds of peptides **P1-P5** can be readily deduced from NMR data, providing the populations of the four possible conformers (noted *cc, tc, cc, tt,* see Table 2).[15] For Yaa = Pro or Hyp (**P1**, **P2** and **P4**), the Xaa-Yaa amide bond was almost entirely in the *trans* conformation whereas for Yaa = TfmΨP, the *cis* amide bond, unsuitable for triple helix assembly, represented

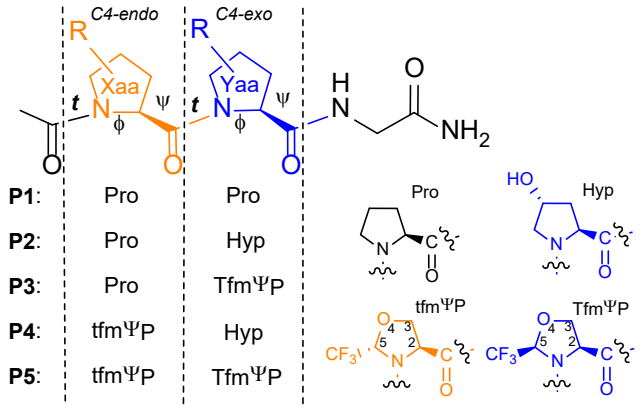


Figure 1. Chemical structures of peptide triads corresponding to established (**P1-P2**) or prospective (**P3-P5**) repeated units of CMPs. Conformational features found in triple helix structures are specified: *trans* peptide bond, C4-endo, ϕ = –73° and ψ = 163° for Xaa residue and *trans* peptide bond (t), C4-exo, ϕ = –58° and ψ = 151° for Yaa residue. Numbering of the atoms is indicated for the trifluoromethylated pseudoprolines residues.

Table 2. *Cis-trans* populations (%) for each peptide **P1-P5**

Peptide Ac-Xaa-Yaa-Gly		*cis* Xaa-Yaa		*trans* Xaa-Yaa	
		cis Ac-Xaa	*trans* Ac-Xaa	*cis* Ac-Xaa	*Trans* Ac-Xaa
conformers		*cc*	*tc*	*ct*	*tt*
P1	Ac-Pro-Pro-Gly	0	0	20	**80**

P2	Ac-Pro-Hyp-Gly	0	0	20	**80**
P3	Ac-Pro-TfmΨP-Gly	5	40	5	**50**
P4	Ac-tfmΨP-Hyp-Gly	2	3	30	**65**
P5	Ac-tfmΨP-TfmΨP-Gly	5	20	10	**65**

Measurements have been performed by integration of ^{1}H signals (and ^{19}F when available) in water at 25°C. Four conformers are considered for each Ac-Xaa-Yaa-Gly-NH_2 peptide since two amide bonds are subject to *cis-trans* isomerism.

as much as 45% for **P3** or 25% for **P5** (*cc* plus *tc* populations). These *cis* amide bond populations contrast with previous observations on small homochiral Fmoc-Pro-TfmΨP-OMe dipeptides which showed no *cis* conformation.[15a] However, similar increased *cis* populations were found for longer peptides incorporating Xaa-TfmΨP moieties.[15b] This did not reflect any propensity of the TfmΨP residue to favor *cis* peptide bonds but was a consequence of CO_i…HN_{i+3} intramolecular H-bonds stabilizing type VI β-turns. Such secondary structure elements were also found in this work for the *tc* conformers in **P3** and **P5**: the chemical shift dependence of their Gly amide proton towards the temperature confirmed the presence of hydrogen bonds (see the Supporting Information, Table S1). However, in longer CMP peptides, these glycine amide protons are expected to be involved in the inter-strand H-bonds of the triple helix and thus will no longer be available to form such type VI β-turns around the

Table 3. *Analysis of ring puckers by NMR and MD simulations for the all trans conformation (tt) of peptides* **P1-P5**.

	Peptide Ac-Xaa-Yaa-Gly	% endo [a] of Xaa		% exo [b] of Yaa	
		NMR [c]	*MD* [d]	*NMR* [c]	*MD* [d]
P1	Ac-Pro-Pro-Gly	52	64	59	29
P2	Ac-Pro-Hyp-Gly	55	63	100	86
P3	Ac-Pro-TfmΨP-Gly	30	67	100	98
P4	Ac-tfmΨP-Hyp-Gly	97	99	100	85
P5	Ac-tfmΨP-TfmΨP-Gly	95	99	100	98

[a] % of C(4)-endo ring pucker for **P1-P3** or O(4)-endo ring pucker for **P4-P5**. [b] % of C(4)-exo ring pucker for **P1**, **P2**, **P4** or O(4)-endo ring pucker for **P3** and **P5**. [c] For each amino acid ring, populations (%) have been calculated using a two-sites equilibrium model (C4)-endo/(C4)-exo and the corresponding optimized geometries (see also the Supporting Information, Tables S2 and S3).[14c-d] [d] Populations have been determined from 1 μs simulations in explicit water based on the amber99sb-inlp force field, completed by Generalized Amber Force Field (GAFF) for a few of missing parameters. The puckering propensities were calculated through the distributions of the dihedral angle χ_1 N-C(2)-C(3)-X(4) (X = C or O) using a sum of two gaussians, corresponding to the exo (χ_1<-15°) or the endo (χ_1>15°) puckerings.

Xaa-TfmP moieties. Finally, we observed that the *ct* population was significant in **P4**. This is not surprising since the presence of the CF_3 group at the C5 position in tfmΨP was shown to slightly destabilize the *trans* geometry of Ac-tfmΨP peptide bond[15c] but the *cis* population was shown to progressively decrease with increasing size of the preceding amino acid.[15a] Therefore, a higher *trans* population content is expected for the tfmΨP residue when it is centrally located in a CMP sequence.

Ring puckering preferences. Measurements of the $^{3}J_{H\alpha H\beta 2}$ and $^{3}J_{H\alpha H\beta 3}$ scalar couplings have allowed the calculation of puckering preferences for the main *trans-trans* conformations (Tables 3 and S2-3). Based on a two-conformations X(4)-endo/X(4)-exo jump model (X = C or O), our NMR measurements were in good agreement with the MD simulations and confirmed that the Pro residue is highly flexible within the tripeptide models both at the Xaa and the Yaa positions. The stabilization of the X(4)-exo conformation was also demonstrated for the Hyp and TfmΨP residues at the Yaa positions, whereas the tfmΨP residue considerably enhanced the endo conformation at the Xaa position over the unsubstituted proline residue. Thus, based on the assumption that the triple helix is stabilized by the amino acids puckering alternation,[6] **P4** and **P5** should be good candidates for the design of new CMP sequences.

Backbone conformation. Next, we were interested in the distribution of the ϕ and ψ dihedral angles for these five tripeptides in the time course of 1 μs-long MD simulations. Ramachandran plots of Xaa and Yaa are presented in Figure 2 for **P1-P5** and compared with the triple helix canonical geometry. For **P2**, the ϕ and ψ dihedral angles were very close to those measured in the triple helix. This is not surprising since it is composed of the same amino acids as in the triple helix model. As shown on Figure 2a, the replacement of the Pro residue at the Xaa position of by the tfmΨP residue (**P2**→**P4** and **P3**→**P5**) induced an increase of the ϕ-value by ca. 10° without affecting the ψ-value. On the contrary, the replacement of the Hyp by the TfmΨP residue at the Yaa position (**P2**→**P3** and **P4**→**P5**) decreased the Xaa ψ-value by ca. 20°, Xaa ϕ-value being unchanged. Concerning the Yaa residue, small angle shifts were observed for these Hyp→TfmΨP substitutions (–10° for ϕ and +10° for ψ, see Figure 2b). Overall, these small structural deviations compensate within each tripeptide units and could accommodate the triple helix structure.

Structural Studies of fluorinated CMP. To further investigate the behavior of the C(5)-substituted pseudoprolines within a collagen-mimicking triple-helix, CMP were modeled using molecular dynamics simulations. The stable non-fluorinated triple helix of $(Pro\text{-}Hyp\text{-}Gly)_7$ was chosen as a reference structure and was compared to the homotrimer triple helix of the host-guest peptide containing the triad found in **P5**, *i.e.* [(Pro-Hyp-Gly)$_3$-**tfmΨP-TfmΨP-Gly**-(Pro-Hyp-Gly)$_3$], denoted as **CMP$_{(P5)}$**. The simulated trimeric structure [**CMP$_{(P5)}$**]$_3$ is displayed in Figure 3. The triple-helical structure was well conserved in the time course of the simulation as attested by the low root mean square displacements (RMSD). Relative to the reference crystalline structure (PDB 1K6F),[5b] the RMSD along the 500 ns-long trajectories was about 1.2 Å for the backbone atoms of the triple helix both in the absence and in the presence of trifluoromethyl pseudoproline residues (see the Supporting Information Figure S9).

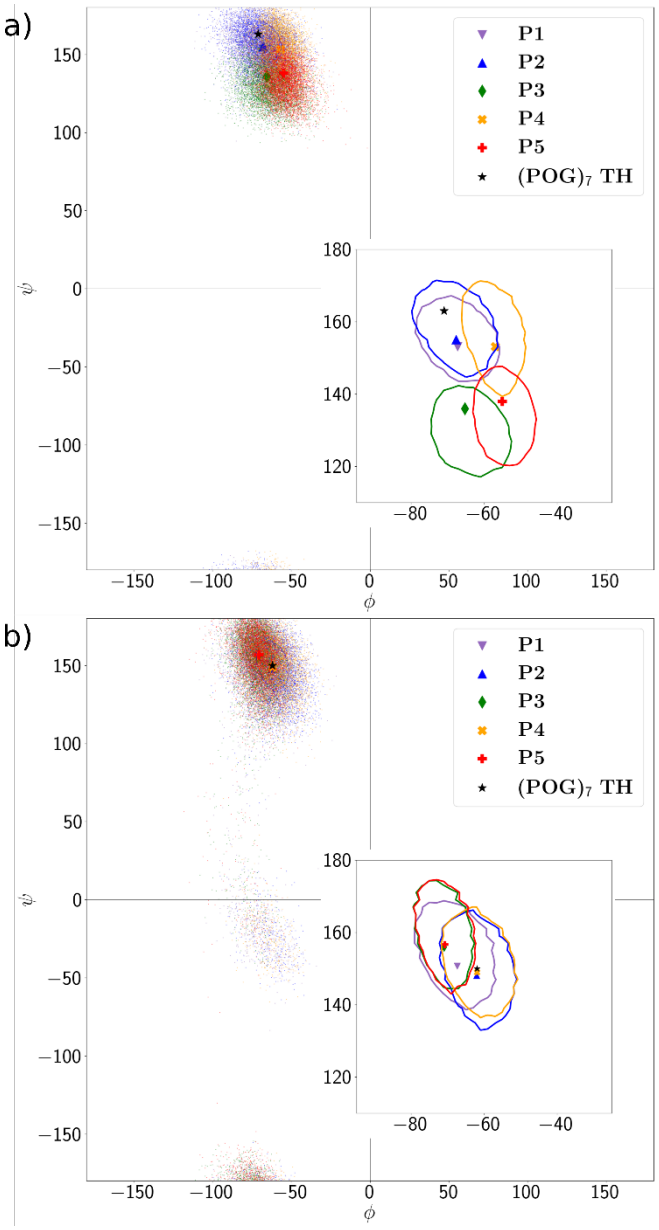


Figure 2. Distribution of the ϕ and ψ dihedral angles obtained by MD simulations for Xaa (a) and Yaa (b) amino acids in peptides **P1-P5** with a carboxamide *C*-terminal group (see Table S4 for detailed values). For each peptide, 500 000 structures have been extracted from a 1μs MD trajectory. The corresponding ϕ and ψ dihedral angles have been plotted on Ramachandran diagrams. The large colored symbols correspond to the average values. On the inserts, the contour lines delineate the 1σ variation. The star (★) corresponds to the average dihedral angles obtained on a 500 ns-long MD simulation for the central triplet within the (Pro-Hyp-Gly)$_7$ triple helix (POG$_7$ TH in the legend).

In the simulation, the six trifluoromethyl groups rotated around the C(5)-CF$_3$ bond, indicating that they were not entangled in a bulky environment, and that the structure around them was relaxed. The analysis of the ring puckering confirmed that the tfmΨP residue at the Xaa position greatly stabilized the O(4)-endo conformation within the triple helix (χ_1 = 18±6°), whereas the Pro residue displayed a high flexibility in the reference triple helix, the endo and exo pucker being almost equally populated (Figure 4a). At the Yaa position, the stabilization of the C(4)-exo by the Hyp residue was clearly demonstrated (χ_1 = -30±8°) but we observed that the

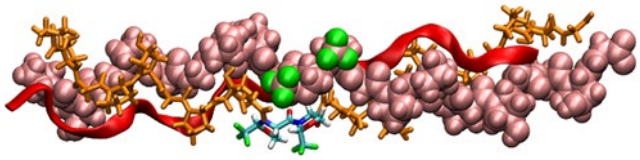


Figure 3. Conformation of the homo-trimer helix of **CMP$_{(P5)}$** after a 500 ns simulation in explicit solvent. Initial structure was derived from the 1K6F PDB structure.[5b] The three backbone chains are represented in three different modes: CPK (pink), lines (orange) and ribbons (red). The fluorine atoms are depicted in green. Usual atomic color code is used for the central host peptide in the line representation. Picture was generated using VMD and povray rendering [15].

TfmΨP residue was even able to restrict further the dynamic of the 5-membered ring (χ_1 = -24±4°, Figure 4b).
Under close scrutiny, some other differences could be detected between the reference and the fluorinated triple-helix. The pseudoprolines presence shifted the average backbone dihedral angles relative to those of (Pro-Hyp-Gly)$_7$ (see the Supporting Information, Table S5), as already noticed in model tripeptides (Figure 2). These modifications of the triple helix structure locally distorded the Gly-NH→Pro-CO hydrogen bond network between the three peptide chains. Indeed, the hydrogen-bond donating N of Gly and the accepting O of Pro or tfmΨP was

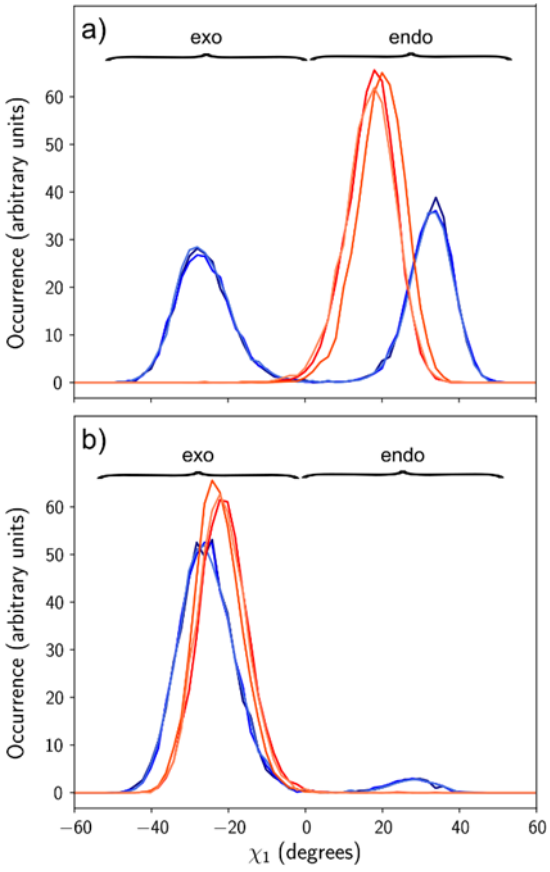


Figure 4. Probability distribution of the dihedral angle χ_1 (N-C(2)-C(3)-C(4)) for the (a) Xaa and (b) Yaa amino acids of the central triplets in the reference triple helix (Pro-Hyp-Gly)$_7$ (blue curves) and in the modelized homo-trimer helix of **CMP$_{(P5)}$** (orange curves). For both triple helices, distributions are given for the three chains along 500-ns long trajectories.

3.0±0.2 Å), but the hydrogen bond orientations were less favorable in the central region of the fluorinated triple helix (see Figure S10). These distortions only locally decreased the hydrogen-bond probability, the tfmΨP-TfmΨP motifs present in the triple helix did not destabilize the overall PPII structure of the three peptide chains during the time course of the MD simulation (see the Supporting Information, Figure S11).

In order to experimentally demonstrate the presence of a triple helix assembly in solution, we synthetized the **CMP$_{(P5)}$** peptide and undertook a liquid state NMR study. First, Fmoc-Gly-tfmΨP-TfmΨP building blocks were synthetized and SPPS was used to obtain the desired **CMP$_{(P5)}$** with an overall yield of 20% (see the Supporting Information). To our knowledge, this is the first CMP model incorporating C(5)-substituted proline analogues. Since the peptide chain contains two adjacent trifluoromethylated pseudoprolines, ^{19}F NMR spectra were recorded to get a clear picture of its structural features.

Interestingly, a first set of six resonances was observed between –76.0 and –76.6 ppm, away from a large signal centered at –77.9 ppm. On the ^{19}F-^{13}C HSQC spectrum, the latter split into two ^{13}C signals at 117.30 and 117.65 ppm (Figure 5b). To better understand these observations, a ^{19}F DOSY spectrum was recorded which indicated that the six downfield resonances shared a same diffusion coefficient (7.53×10^{-11} m^2.s^{-1}), whereas the signal at –77.9 ppm corresponded to a significantly larger diffusion coefficient (1.14×10^{-10} m^2.s^{-1}). Thus, the DOSY measurements allowed the calculation of a hydrodynamic radius ratio of 1.51 between these two species which was associated to a volume ratio of 3.4. These data strongly suggested that the six upfield resonances correspond to the triple helix whereas the two signals at –77.9 ppm accounts for the two CF_3 groups of the tfmΨP-TfmΨP motif in the isolated **CMP$_{(P5)}$** peptide strand. Within the triple helix, the one-residue offset between the three chains leads to different environments for each of the six trifluoromethyl groups and lifts the chemical shift degeneracies (Figure 5).

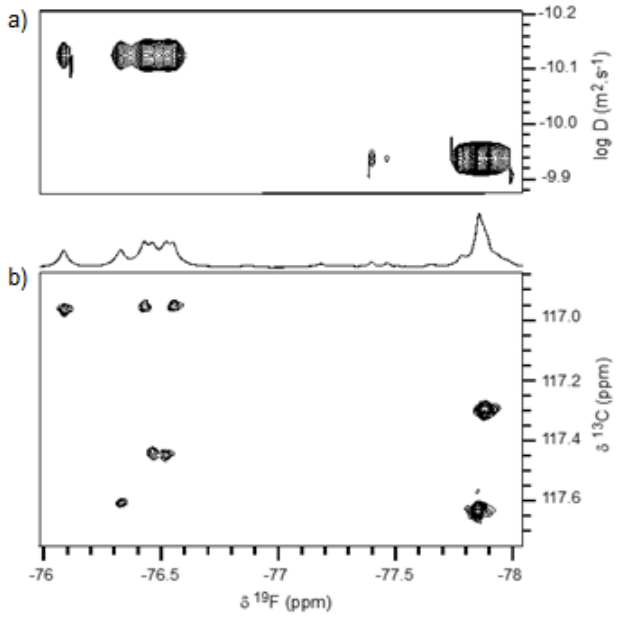


Figure 5. a) ^{19}F DOSY spectrum of the **CMP$_{(P5)}$** peptide. b) 2D ^{19}F-^{13}C HSQC spectrum together with a ^{19}F 1D projection. Peptide concentration 10mM, pH = 6.8, 40°C, H_2O:D_2O 90:10.

Therefore, our NMR experiments demonstrated that the **CMP$_{(P5)}$** peptide was in a slow exchange regime between a triple helix and a monomeric state. At 40°C, the integration of the ^{19}F NMR signals indicated a population of 25% for the triple helix composed of **CMP$_{(P5)}$** peptide incorporating two adjacent trifluoromethylated proline analogues with unusual C5-substitutions.

Conclusion

In this work, we have demonstrated that trifluoromethylated pseudoprolines, bearing a CF_3 group at the C(5) position, represent valuable tools for the design of triple helix models. Our NMR and MD studies showed that peptides **P3-P5**, incorporating tfmΨP and/or TfmΨP residues, meet the structural requirements for the triple helix assembly. In particular, the incorporation of two adjacent tfmΨP and TfmΨP residues preserved a high population of the all-*trans* conformation (*tt*) which is required for triple helix assembly. Analysis of ring puckers for the *tt* conformers showed that tfmΨP and TfmΨP residues displayed the strongest propensities to stabilize the endo and exo conformation, respectively. This feature provides a strong rationale for selecting trifluoromethylated pseudoprolines as proline analogues in CMP sequences.

MD simulations showed that tfmΨP and TfmΨP imposed small deviations relative to the canonical ϕ and ψ angles found in collagen triple helices, which slightly affects the hydrogen bond network in the close neighborhood of the substitutions. Overall, these deviations compensate along the peptide chain and the triple helix remains stable in the MD simulations. ^{19}F NMR was used to demonstrate the existence of the fluorinated triple helix in solution. The analysis of diffusion coefficients was in good agreement with the formation of a trimer. In addition, NMR spectra showed distinct signals for the six trifluoromethyl groups embedded in the triple helix.

Our results pave the way for the development of new fluorinated collagen model peptides. Our work is now focused on the in-depth analysis of the role of trifluoromethylated pseudoprolines on the kinetic and thermodynamic properties of the triple helix and their relevance for biomaterial applications.

Acknowledgements

E. M. thanks the Agence Nationale de la Recherche for financial support (No. ANR-12-JS08-0005-01). G. C. thanks the French ministry of research for awarding a research fellowship to K. R. C.L. thanks GENCI-TGCC (grant x2018087662) and the Pôle Scientifique de Modélisation Numérique (PSMN, ENS Lyon, France) for the generous computing time allocation. We thank Ewen Lescop and Helena Kovacs for their contribution to the NMR acquisitions.

Keywords: peptide • collagen model peptide • fluorine • NMR • Molecular dynamics

[1] L. Gu, T. Shan, Y. X. Ma, R. R. Tay, L. Niu. *Trends Biotechnol.* **2019**, *37*, 464-491.

[2] a) M. D. Shoulders, R. T. Raines, *Annu. Rev. Biochem.* **2009**, *78*, 929-958; b) J. Bella, *Biochem. J*. **2016**, *473*, 1001-1025; c) V. Kubyshkin V. *Org Biomol Chem*. **2019**, *17*, 8031-8047; d) C. Tanrikulu, W. M. Westler, A. J. Ellison, J. L. Markley, R. T. Raines, *J. Am. Chem. Soc.* **2020**, *142*, 1137-1141.

[3] a) M. R. Aronoff, J. Egli, M. Menichelli, H. Wennemers, *Angew. Chem. Int. Ed.* **2019**, *58*, 3143 –3146; b) C. M. Peterson, M. R. Helterbrand, J. D. Hartgerink, *Biomacromolecules* **2022**, *23*, 2396-2403; c) Y. Ren, H Zhang, W. Qin, B. Du, L. Liu, J. Yang, *Materials Science & Engineering C* **2020**, *108*, 110276; d) S. Chattopadhyay,L. B. C. Teixeira, L. L. Kiessling, J. F. McAnulty, R. T. Raines, *ACS Chem Biol. **2022**, 17,* 314-321.

[4] L. Vitagliano, R. Berisio, A. Mastrangelo, L. Mazzarella, A. Zagari, *Protein Sci.* **2001**, *10*, 2627–2632.

[5] a) J. Bella, M. Eaton, B. Brodsky, H. M. Berman, *Science* **1994**, *266*, 75-81; b) R. Berisio, L. Vitagliano, L. Mazzarella, A. Zagari, *Protein Sci.* **2002**, *11*, 262–270; c) K. Okuyama, K. Miyama, K. Mizuno, H. P. Bächinger, *Biopolymers* **2012**, *97*, 607-616.

[6] a) S. K. Holmgren, K. M. Taylor, L. E. Bretscher, R. T. Raines. *Nature* **1998**, *392*, 666-667; b) L. Vitagliano, R. Berisio, L. Mazzarella, A. Zagari, *Biopolymers* **2001**, *58*, 459–464.

[7] a) J. A. Hodges, R. T. Raines, *J. Am. Chem. Soc.* **2003**, *125*, 9262-9263; b) M. D. Shoulders, I. A. Guzei, R. T. Raines, *Biopolymers* **2008**, *89*, 443–454; c) F. W. Kotch, I. A. Guzei, R. T. Raines, *J. Am. Chem. Soc.* **2008**, *130*, 2952-2953; d) L. S. Sonntag, S. Schweizer, C. Ochsenfeld, H. Wennemers, *J. Am. Chem. Soc.* **2006**, *128*, 14697-14703.

[8] H. K. Ganguly, G. Basu, *Biophys. Rev.* **2020**, *12*, 25-39.

[9] M. P. Hinderaker, R. T. Raines, *Protein Sci.* **2003**, *6*, 1188-1194.

[10] a) R. S. Erdmann, H. Wennemers, *J. Am. Chem. Soc.* **2010**, *132*, 13957-13959; b) R. S. Erdmann, H. Wennemers, *Angew. Chem., Int. Ed. Engl.* **2011**, *50*, 6835-6838; c) R. S. Erdmann, H. Wennemers, *J. Am. Chem. Soc.* **2012**, *134*, 17117-17124; d) J. Egli, C. Siebler, M. Köhler, R. Zenobi, H. Wennemers, *J. Am. Chem. Soc.* **2019**, *141*, 5607-5611.

[11] J. A. Hodges, R. T. Raines, *J. Am. Chem. Soc.* **2005**, *127*, 15923-15932.

[12] M. D. Shoulders, J. A. Hodges, R. T. Raines, *J. Am. Chem. Soc.* **2006**, *128*, 8112–8113.

[13] C. L. Jenkins, L. E. Bretscher, I. A. Guzei, R. T. Raines, *J. Am. Chem. Soc.* **2003**, *125*, 6422-6427.

[14] a) G. Chaume, O. Barbeau, P. Lesot, T. Brigaud, *J. Org. Chem.* **2010**, *75*, 4135-4145; b) G. Chaume, J. Simon, C. Caupène, N. Lensen, E. Miclet, T. Brigaud, *J. Org. Chem.* **2013**, *78*, 10144-10153.

[15] a) G. Chaume, J. Simon, N. Lensen, J. Pytkowicz, T. Brigaud, E. Miclet. *J. Org. Chem.* **2017**, *82*, 13602-13608; b) G. Chaume, D. Feytens, G. Chassaing, S. Lavielle, T. Brigaud, E. Miclet, *New J. Chem.* **2013**, *37*, 1336−1342 ; c) D. Feytens, G. Chaume, G. Chassaing, S. Lavielle, T. Brigaud, B. J. Byun, Y. K. Kang, E. Miclet, *J. Phys. Chem. B* **2012**, *116*, 4069−4079; d) H. S. Park, Y. K. Kang, *New. J. Chem.* **2019**, *43*, 17159--17173.

[16] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graphics* **1996**, *14*, 33-