The New Science of Protein Mimetics

Murray Goodman, Weibo Cai, Garth A. Kinberger*

Department of Chemistry and Biochemistry, University of California - San Diego, La Jolla, California 92093, USA
Email: mgoodman@ucsd.edu

Summary: New chemistries have been developed for *de novo* protein design. Protein mimetics of different structural and functional properties such as synthetic peptide ligases and Dn symmetrical helical bundles have been reported. The Template-Assembled Synthetic Protein (TASP) method (as well as the "Molecular Kit" approach) has also been utilized to prepare protein-like molecules. Here we report the synthesis of single chain, scaffold (TRIS)- and dendrimer-assembled collagen mimetics composed of the Gly-Nleu-Pro sequence where Nleu denotes N-isobutyl glycine. From the CD spectra and the thermal denaturation studies it can be seen that the collagen mimetics prepared form stable triple helices except the single chain structure. Furthermore, the 162-residue collagen mimetic dendrimer exhibits enhanced triple helical stability compared to the equivalent scaffold-terminated structure by an increase in the melting temperature in both H_2O and 2:1 ethylene glycol/H_2O (4 °C and 12 °C respectively). The concentration dependence for the melting transition of the collagen mimetic dendrimer was measured from which it was determined that the stabilization effect arises from the intramolecular clustering of the triple helical arrays about the core structure. This ensemble excludes solvent from the interior portion of the array which stabilizes the triple helix cluster.

Keywords: collagen mimetics, circular dichroism, dendrimers, peptides; triple helix

Introduction

Proteins fold into specific three-dimensional structures to carry out their biological functions. Although there are vast numbers of possible conformations for each protein, the native state is achieved rapidly. Much research has been undertaken in *de novo* protein design to prepare protein mimetics and to understand the protein folding problem. In the past decade, many designed protein mimetics with specific functions have been reported.[1-3] Ghadiri and co-workers have synthesized a 33-residue synthetic peptide ligase.[1] Based on coiled-coil structural motif, the two short negatively charged electrophilic and nucleophilic peptide

 DOI: 10.1002/masy.200351125

substrates (17 and 16 residues respectively) are pre-organized on a longer complementary, positively charged catalyst, forming the ternary complex which facilitates the ligation reaction between a C-terminal thioester and an N-terminal cysteine (Figure 1).

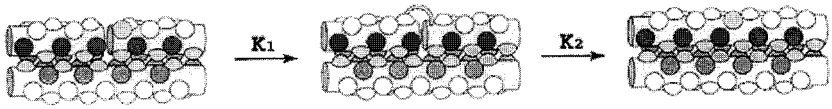


Fig. 1. Ghadiri's synthetic peptide ligase catalyzes the ligation of two short peptides.[1]

Anti-parallel helical bundles are found in a wide variety of proteins. DeGrado and co-workers synthesized Dn symmetrical helical bundles[2] where n equals 2, 3 and 6 (4, 6 and 12 helix bundles respectively), and the coiled helix backbone models generated of the synthesized compound superimpose with the experimental protein crystal structures (Figure 2).

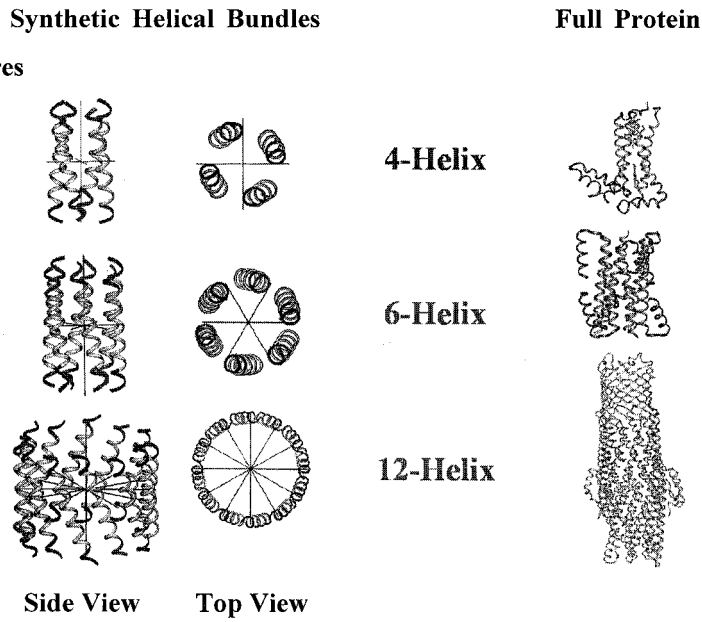


Fig. 2. DeGrado's Dn symmetrical anti-parallel synthetic helical bundles (the helices are represented as ribbons). Only the central helical portion of the proteins were synthesized.[2]

Mutter and co-workers extended his concept of the Template-Assembled Synthetic Protein (TASP) to a "Molecular Kit" approach (Figure 3).[3] Individual secondary structural elements such as α-helices, β-sheets, turns, and loops are covalently attached via both chain ends to appropriately functionalized templates. A series of protein-like molecules are formed exhibiting interesting structural and functional properties such as locked-in-4-helix bundle (α_4), β-sheet bundle (β_4), or more complex arrangements ($\beta_2\alpha$ and $\beta_3\alpha_2$).

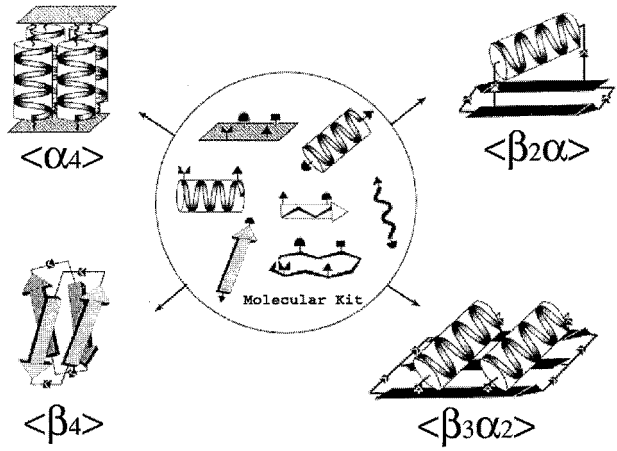


Fig. 3. Mutter's "Molecular Kit" and his synthetic protein mimetics.[3]

Collagen

Collagen is the most abundant extracellular protein in vertebrates. It has a unique triple helical motif in which three polypeptide chains form left-handed helices and are supercoiled into a right-handed triple helix. The primary sequence of triple helical collagen is composed of Gly-Xaa-Yaa trimer repeats. The imino acid proline (Pro) is usually found in the Xaa and Yaa positions whereas 4-hydroxyproline (Hyp) is normally located in the Yaa position.[4] There are two structural models for collagen. One is the Rich & Crick model where three polypeptide chains fold into a 10/3 helix with an axial 28.6 Å repeat.[5] The other model was proposed by Okuyama in which three strands form a 7/2 helix with an axial repeat of 20.0 Å.[6]

Collagen is an important biomaterial because of its low immunogenicity and high tensile strength. However, it is difficult to purify natural collagen without degrading its structural

integrity. Furthermore, natural collagen is subjected to enzymatic degradation under biological conditions. Synthetic collagen structures including unnatural amino acid residues offer an alternative to natural collagens and have been shown to exhibit better enzymatic stability.[7]

Triple-helical conformations of collagen can be detected by a variety of experimental techniques such as X-ray diffraction,[5,6] electron microscopy, circular dichroism (CD) spectroscopy,[8] NMR spectroscopy and molecular modeling.[9] The most frequently used technique to study triple helical conformations is CD spectroscopy. The natural collagen triple helix has a characteristic CD spectrum with a small positive peak around 220 nm, a crossover at about 213 nm and a trough near 197 nm.[8]

Early studies of the preparation of synthetic collagens were conducted using polymerization methods. However, this approach lacks sequence and molecular weight control. Solid phase peptide synthesis (SPPS) has facilitated the synthesis of peptides with specific lengths and sequences. Sakakibara and co-workers have successfully synthesized $(Pro\text{-}Pro\text{-}Gly)_n$ (n = 10, 15, 20) and $(Pro\text{-}Hyp\text{-}Gly)_n$ (n = 5, 10) using the SPPS method.[10-11] $(Pro\text{-}Pro\text{-}Gly)_{20}$ and $(Pro\text{-}Hyp\text{-}Gly)_{10}$ formed stable triple helices in water. Many other peptides with different trimer sequences have also been prepared using the SPPS method.

Pioneering research by Prockop and co-workers revealed that Hyp greatly increases the thermal stability of collagen.[12] Later they also found that only 4-(R)-Hyp stabilized the triple helix while 4-(S)-Hyp destabilized triple helix.[13]

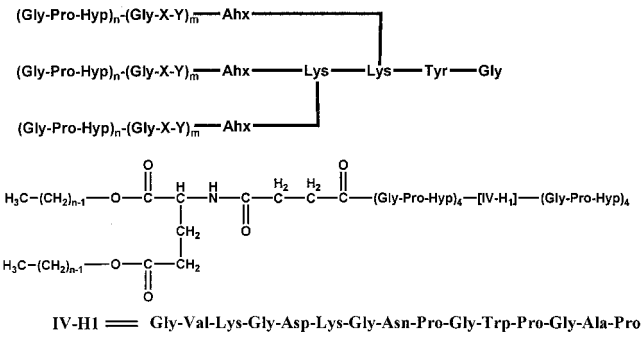
Raines and co-workers synthesized collagen mimetic peptides incorporating 4(R)-fluoro-L-proline (Flp) and found that $(Pro\text{-}Flp\text{-}Gly)_{10}$ forms a highly stabilized triple helix which melts at 91 °C.[14] They concluded that Flp (similar to Hyp) in the Yaa position stabilizes collagen-like triple helical structures by a combination of two effects: stereo-electronic (the gauche-effect) which fixes the pyrrolidine ring pucker and an increase of the trans/cis ratio of the peptide bond. Through these effects Flp as well as Hyp in the Yaa position preorganizes all three main chain torsion angles: ω, φ and ψ which stabilizes the collagen triple helix.

Scaffold-Assembled Collagen Mimetics

The folding of peptides into their secondary or tertiary structures is the essential requirement to induce the proper biological response or activity. To facilitate protein folding, Mutter and co-

workers introduced the TASP approach for the design of a four α-helix bundle.[3]

The concept of scaffold-assembled structures has been applied to the design of collagen-like triple helices. The terms template and scaffold are often used interchangeably. We prefer to define the structures which hold the three peptide chains together as scaffolds. Fields and co-workers have used a lysine-lysine dimer as a scaffold to assemble peptide chains into collagen mimetics.[15] Later they also employed aliphatic fatty acid chains as a scaffold to self-assemble peptide chains into collagen mimetics (Figure 4).[16]



Fatty Acid Scaffold

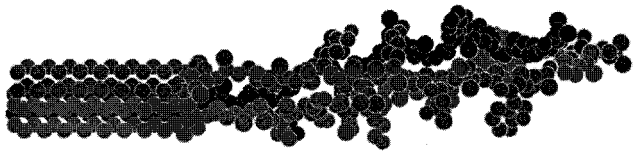


Fig. 4. The scaffold (Lys-Lys dimer) assembled collagen mimetics[15] and self-assembling peptide amphiphiles[16] used by Fields and co-workers to study triple helicity.

Tanaka and co-workers also used the lysine-lysine dimer derived scaffolds to assemble collagen-like peptide chains. They extended this approach by synthesizing collagen-like peptides with both ends tethered by scaffolds and found enhanced thermal stability of the triple helix (Figure 5).[17]

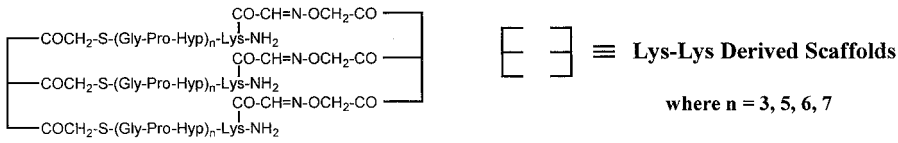


Fig. 5. Tanaka's di-scaffold-assembled collagen mimetics.[17]

Moroder and co-workers synthesized heterotrimeric bioactive collagen mimetic peptides using a "Cystine knot" strategy (Figure 6).[18] They incorporated a collagenase cleavage site of collagen type I of the two α1-chains and one α2-chain which was stabilized by the N-terminal extension of $(Pro\text{-}Hyp\text{-}Gly)_5$ repeats. The three chains were tethered by a "Cystine knot". A tryptophan and a danzyl group were incorporated into the peptide for fluorescence resonance energy transfer. They discovered that cleavage of the peptide by collagenase led to a 6-fold increase in the Trp fluorescence intensity, thus making it a useful fluorogenic substrate for interstitial collagenases.

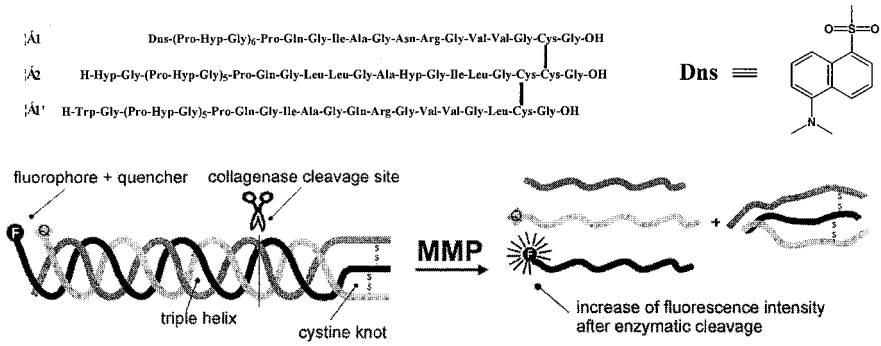


Fig. 6. Moroder's heterotrimeric bioactive collagen mimetic. Dns = dansyl group.[18]

In our design of scaffold-assembled collagen mimetics, we have incorporated both cis-1,3,5-trimethyl cyclohexane-1,3,5-tricarboxylic acid (also known as the Kemp triacid, KTA) and tris(2-aminoethyl)amine (TREN) into collagen mimetic peptides. Spacers were used for both scaffolds, glycine for KTA and succinic acid for TREN respectively (Figure 7).[19,20]

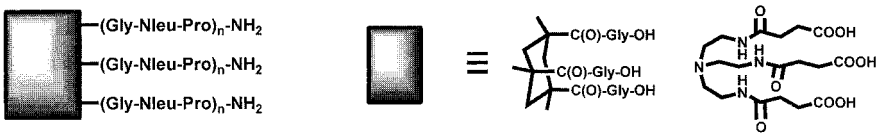


Fig. 7. The KTA-(Gly-OH)$_3$ and the TREN-(suc-OH)$_3$ scaffolds.[19,20]

We have also incorporated an unnatural amino acid N-isobutyl glycine (Nleu) as a proline surrogate (Figure 8).[21] Ring opening between the α and β carbons leads to N-propyl glycine which was synthesized and found to be too hydrophilic. Since hydrophobic interactions are important in triple-helix stabilization, a methyl group was added to achieve the desired hydrophobicity. Nleu was incorporated into both the Gly-Nleu-Pro and the Gly-Pro-Nleu sequences. It was found that scaffold-assembled collagen mimetic peptides composed of both sequences form stable triple helices when the peptide chains attain certain lengths (5 or 6 trimer repeats).

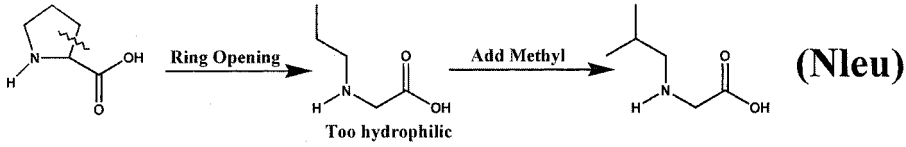


Fig. 8. Conceptual approach to the design of Nleu as a proline mimetic.[21]

Some of the results from biophysical studies of the KTA and TREN assembled collagen mimetic peptides composed of both sequences are shown below (Figure 9,10).[22,20]

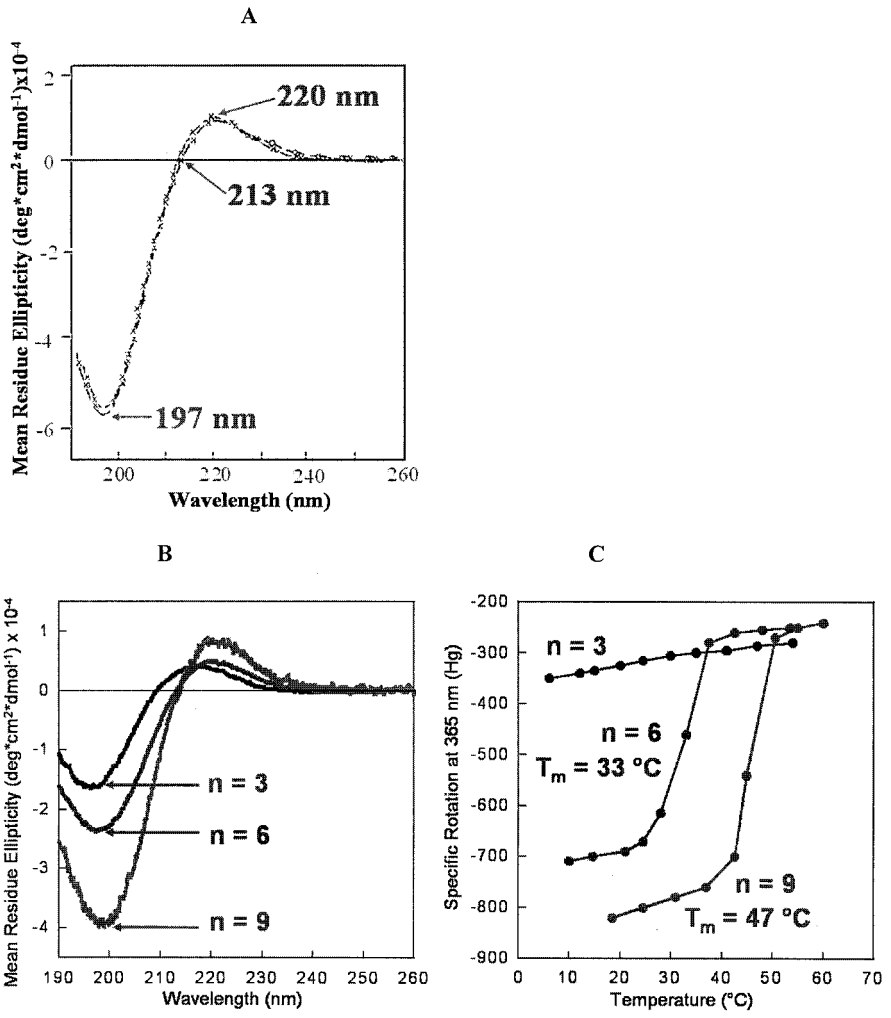


Fig. 9. (A) CD spectra of natural collagen. (B) CD spectra of KTA-[Gly-(Gly-Pro-Nleu)$_n$-NH_2]$_3$ (n = 3, 6 and 9). (C) Melting transitions of KTA-[Gly-(Gly-Pro-Nleu)$_n$-NH_2]$_3$ (n = 3, 6 and 9).[22]

Collagen Mimetic Dendrimers

Dendrimers are highly branched globular macromolecules with functional groups located at the periphery of the globules. We have incorporated another scaffold N-(t-butyloxycarbonyl)-β-alanyl-tris(carboxyethoxymethyl) aminomethane (Boc-β-Ala-TRIS[OH]$_3$) into our collagen mimetic research (Figure 11).[23] The TRIS scaffold not only has three carboxylic acid groups for attachment of peptide chains, it also has an amino group where it can be attached to a core structure to form collagen mimetic dendrimers.

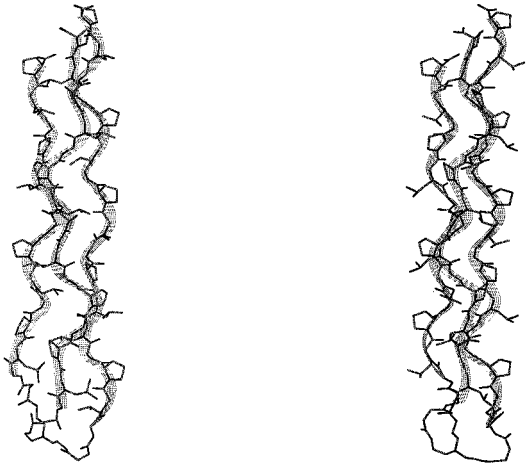


Fig. 10. Lowest energy structures of TREN-[suc-(Gly-Nleu-Pro)$_5$-NH_2]$_3$ from each of the two families of clusters generated from molecular modeling. The two families are based on the orientation of the isobutyl side chains of the Nleu residue: $\chi^1 < 0$ (pointing in) and $\chi^1 > 0$ (pointing out).[20]

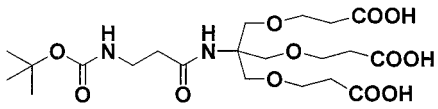


Fig. 11. The Boc-β-Ala-TRIS[OH]$_3$ scaffold for collagen mimetic assembly.[23]

Recently we reported the synthesis and characterization of collagen-like dendrimers.[23] We have synthesized the single chain structure Boc-(Gly-Nleu-Pro)$_6$-OMe, scaffold-assembled

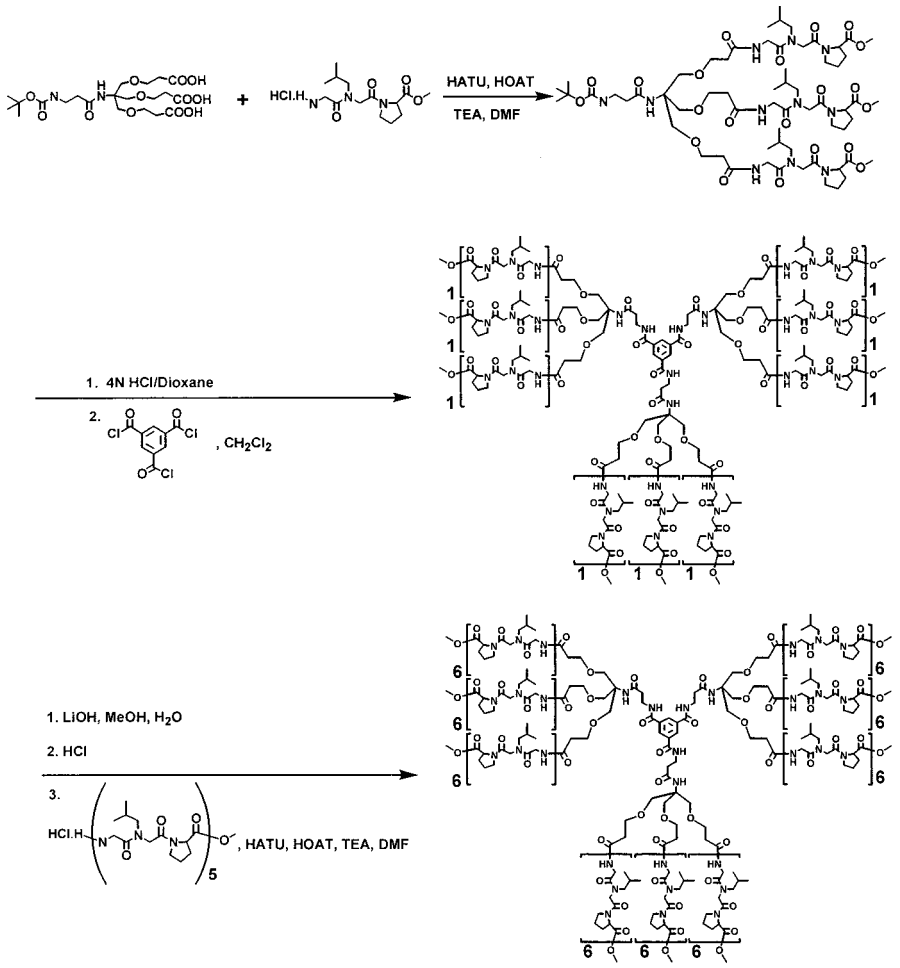


Fig. 12. The synthesis of the dendrimer TMA[β-Ala-TRIS-[(Gly-Nleu-Pro)$_6$-OMe]$_3$]$_3$.

structure Boc-β-Ala-TRIS-[(Gly-Nleu-Pro)$_6$-OMe]$_3$ and the collagen mimetic dendrimer TMA[β-Ala-TRIS-[(Gly-Nleu-Pro)$_6$-OMe]$_3$]$_3$ where TMA denotes the trimesic acid core. Isomeric structures composed of the Gly-Pro-Nleu sequence were also prepared and characterized in similar fashion. In this paper, only the results of the collagen mimetics composed of the Gly-Nleu-Pro sequence are discussed. The single chain structure and the

scaffold-assembled structure were synthesized by a series of stepwise peptide condensations. A combination of divergent and convergent approaches was employed to synthesize the collagen mimetic dendrimer (Figure 12).

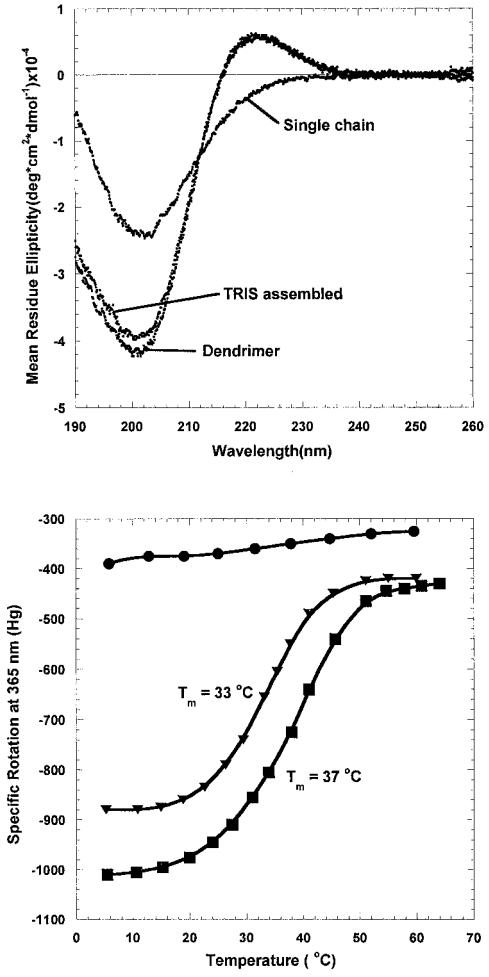


Fig. 13. CD spectra (carried out at 8 °C) and thermal denaturations measured by changes in optical rotations at 365 nm carried out in H_2O (0.2 mg/ml): Single chain structure (Boc-(Gly-Nleu-Pro)$_6$-OMe, ●), TRIS assembled structure (Boc-β-Ala-TRIS-[(Gly-Nleu-Pro)$_6$-OMe]$_3$, ▼), and Dendrimer (TMA[β-Ala-TRIS-[(Gly-Nleu-Pro)$_6$-OMe]$_3$]$_3$, ■).[23]

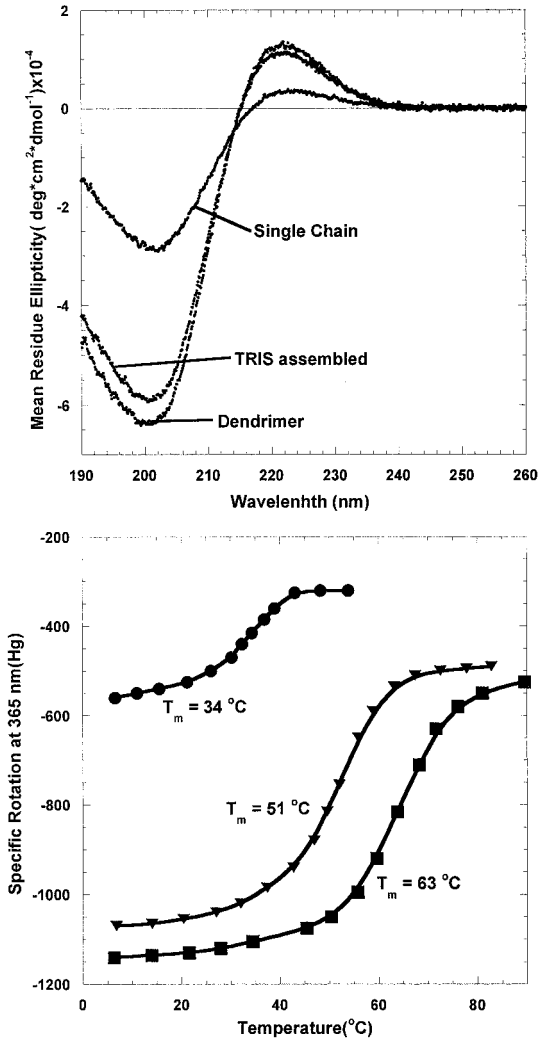


Fig. 14. CD spectra (carried out at 8 °C) and thermal denaturations measured by changes in optical rotations at 365 nm carried out in ethylene glycol/H_2O (2:1, v/v) (0.2 mg/ml): Single chain structure (Boc-(Gly-Nleu-Pro)$_6$-OMe, ●), TRIS assembled structure (Boc-β-Ala-TRIS-[(Gly-Nleu-Pro)$_6$-OMe]$_3$, ▼), and Dendrimer (TMA[β-Ala-TRIS-[(Gly-Nleu-Pro)$_6$-OMe]$_3$]$_3$, ■).[23]

The triple helicity of all structures were determined by CD measurements and thermal denaturation studies monitored by optical rotation (Figure 13, 14).[23] These studies were carried out in both H_2O and the triple helicity-enhancing solvent ethylene glycol:H_2O (EG:H_2O, 2/1, v/v).

The optical rotation data shown in Figure 13 and 14 indicates that the molecules studied exhibit cooperative melting transitions except the single-chain compound in water shown in Figure 13. A broad shallow transition for the single chain molecule can be seen in 2:1 EG/H_2O (Figure 14). This indicates that the single chain structure may form triple helices to a small extent in 2:1 EG/H_2O. The CD data shown in Figure 13 and 14 are also indicative of triple helical conformations for the scaffold-assembled and dendritic Gly-Nleu-Pro containing molecules at low temperature.

It is clear from Figure 13 and 14 that the dendritic collagen mimetic forms more thermally stable triple helices than the corresponding scaffold-assembled structure, the melting temperature (T_m) of the dendrimer is 4 °C (in H_2O) and 12 °C (in 2:1 EG/H_2O) higher than those of the scafold-assembled structure. In order to determine whether the stabilizing effect of the dendrimer arises from intermolecular or intramolecular interactions, we measured the melting transition of the collagen mimetic dendrimer in H_2O at different concentrations (0.05 mg/ml, 0.2 mg/ml and 2.0 mg/ml). No significant change in the T_m was observed over the concentration range.[23] We therefore believe that the stabilizing effect arises from an intramolecular clustering of the triple helical arrays about the core structure. Figure 15 shows a possible schematic representation for such a cluster. This cluster excludes solvent from the interior portion of the array which leads to stabilization of the triple helix bundle.

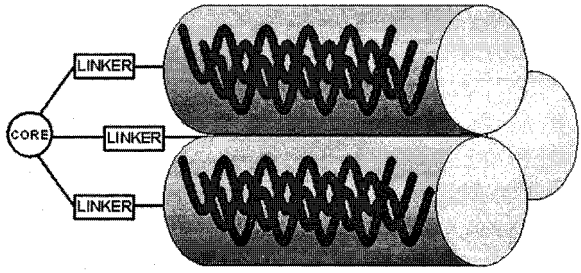


Fig. 15. A schematic representation of the triple helix cluster.[23]

Conclusion

New chemistries have been developed for the preparation of protein mimetics. As part of this new science, we have designed and synthesized single chain, scaffold (TRIS)- and dendrimer-assembled collagen mimetics composed of the Gly-Nleu-Pro sequence. From the CD spectra and the thermal denaturation studies it can be seen that the 162-residue collagen mimetic dendrimer exhibits enhanced triple helical stability compared to the corresponding scaffold-terminated structure by an increase in the melting temperature in both H_2O and 2:1 EG/H_2O (4 °C and 12 °C respectively). The concentration dependence for the melting transition of the collagen mimetic dendrimer was measured from which it was determined that the stabilization effect arises from the intramolecular clustering of the triple helical arrays about the core structure. This ensemble excludes solvent from the interior portion of the array which leads to the stabilization of the bundle of triple helices.

It is our intention to create other collagen-like dendrimers with functional groups attached including metal binding sites, integrin sequences, drugs and others. These ensembles will represent a novel class of structures for biomaterial applications.

[1] K. Severin, D. H. Lee, A. J. Kennan, M. R. Ghadiri, *Nature* **1997**, *389*, 706.
[2] B. North, C. M. Summa, G. Ghirlanda, W. F. DeGrado, *J. Mol. Biol.* **2001**, *311*, 1081.
[3] G. Tuchscherer, D. Grell, M. Mathieu, M. Mutter, *J. Peptide Res.* **1999**, *54*, 185.
[4] F. R. Brown, J. P. Carver, E. R. Blout, *J. Mol. Biol.* **1969**, *39*, 307.
[5] A. Rich, F. H. C. Crick, *J. Mol. Biol.* **1961**, *3*, 483.
[6] K. Okuyama, K. Okuyama, S. Arnott, M. Takayanagi, M. Kakudo, *J. Mol. Biol.* **1981**, *152*, 427.
[7] R. J. Simon, R. S. Kania, R. N. Zuckermann, V. D. Huebner, D. A. Jewell, S. Banville, S. Ng, L. Wang, S. Rosenberg, C. K. Marlowe, D. C. Spellmeyer, R. Tan, A. D. Frankel, D. V. Santi, F. E. Cohen, P. A. Bartlett, *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9367.
[8] F. R. III Brown, A. DiCorato, G. P. Lorenzi, E. R. Blout, *J. Mol. Biol.* **1972**, *63*, 85.
[9] M. Li, P. Fan, B. Brodsky, J. Baum, *Biochemistry* **1993**, *32*, 7377.
[10] S. Sakakibara, Y. Kishida, Y. Kikuchi, R. Sakai, K. Kakiuchi, *Bull. Chem. Soc. Jpn.* **1968**, *41*, 1273.
[11] S. Sakakibara, K. Nouye, K. Shudo, Y. Kishida, D. J. Prockop, *Biochim. Biophys. Acta* **1973**, *303*, 198.
[12] R. A. Berg, D. J. Prockop, *Biochem. Biophys. Res. Commun.* **1973**, *52*, 115.
[13] K. Inouye, S. Sakakibara, D. J. Prockop, *Biochimica. et Biophys. Acta* **1976**, *420*, 133.
[14] C. L. Jenkins, R. T. Raines, *Nat. Prod. Rep.* **2002**, *19*, 49.
[15] C. G. Fields, B. Grab, J. L. Lauer, G. B. Fields, *Anal. Biochem.* **1995**, *231*, 57.
[16] G. B. Fields, *Bioorg. Med. Chem.* **1999**, *7*, 75.
[17] Y. Tanaka, K. Suzuki, T. Tanaka, *J. Peptide.Res.* **1998**, *51*, 413.
[18] J. C. D. Müller, J. Ottl, L. Moroder, *Biochemistry* **2000**, *39*, 5111.
[19] M. Goodman, Y. Feng, G. Melacini, J. P. Taulane, *J. Am. Chem. Soc.* **1996**, *118*, 5156.
[20] J. Kwak, A. De Capua, E. Locardi, M. Goodman, *J. Am. Chem. Soc.* **2002**, *124*, 14085.
[21] M. Goodman, G. Melacini, Y. Feng, *J. Am. Chem. Soc.* **1996**, *118*, 10928.
[22] Y. Feng, G. Melacini, J. P. Taulane, M. Goodman, *Biopolymers* **1996**, *39*, 859.
[23] G. A. Kinberger, W. Cai, M. Goodman, *J. Am. Chem. Soc.* **2002**, *124*, 15162.