ANTIOXIDANTS & REDOX SIGNALING
Volume 10, Number 1, 2008
© Mary Ann Liebert, Inc.
DOI: 10.1089/ars.2007.1868

Forum Review

Natural and Artificial Cystine Knots for Assembly of Homo- and Heterotrimeric Collagen Models

CYRIL BOULÈGUE,[1] HANS–JÜRGEN MUSIOL,[1] MARION G. GÖTZ,[2] CHRISTIAN RENNER,[1,3] and LUIS MORODER[1]

ABSTRACT

Native collagens are molecules that are difficult to handle because of their high tendency towards aggregation and denaturation. It was discovered early on that synthetic collagenous peptides are more amenable to conformational characterization and thus can serve as useful models for structural and functional studies. Single-stranded collagenous peptides of high propensity to self-associate into triple-helical trimers were used for this purpose as well as interchain-crosslinked homotrimers assembled on synthetic scaffolds. With the growing knowledge of the biosynthetic pathways of natural collagens and the importance of their interchain disulfide crosslinks, which stabilize the triple-helical structure, native as well as *de novo* designed cystine knots have gained increasing attention in the assembly of triple-stranded collagen peptides. In addition, natural sequences of collagens were incorporated in order to biophysically characterize their functional epitopes. This review is focused on the methods developed over the years, and future perspectives for the production of collagen-mimicking synthetic and recombinant triple-helical homo- and heterotrimers. *Antioxid. Redox Signal.* 10, 113–125.

INTRODUCTION

Collagens are the most abundant protein components of the extracellular matrix. These proteins provide mechanical strength and structural integrity to connective tissues in the animal kingdom. Collagens guide cells to migrate, proliferate, and differentiate by interacting with numerous other molecules. All collagens consist of three polypeptide chains, which can either be assembled into homotrimers or heterotrimers, with the latter being composed of two or three distinguishable α-chains. The common structural signature of collagens is the rigid rod-shaped triple helix. It can either span the whole molecule, as in the case of fibrillar collagens, or it is interspersed by noncollagenous (NC) domains that render the proteins more flexible. The triple-helical sequence portions consist of the characteristic (GXY) tripeptide repeats. The presence of a glycine residue in every third position is essential, and the position X is often occupied by proline and position Y by hydroxyproline. Because of the fixed ψ angle and the restricted φ angle of these imino acids, the sequence composition favors an extended left-handed polyproline II-helix conformation of the single α-chains. The individual left-handed helices are further supercoiled in a right-handed manner around a common axis into a triple helix, with a stagger of one residue between adjacent chains. This triple helix is stabilized by a repetitive backbone hydrogen bonding pattern. However the contribution of the hydration network to the molecular stability is still subject to debate (2, 16, 33, 95).

So far, 27 types of human collagens are known (Table 1). They can be subdivided into different classes according to their supramolecular organization and the resulting specialized function in various diverse tissues (58, 85, 96). Collagens are se-

[1]Max-Planck-Institute of Biochemistry, Martinsried, Germany.
[2]Whitman College, Department of Chemistry, Walla Walla, Washington.
[3]Present address: Deutsche Forschungsgemeinschaft, Kennedyallee 40, 53175 Bonn/Bad-Godesberg, Germany.

TABLE 1. HUMAN COLLAGENS

- Fibrillar collagens (types I*, II, III, V*, XI*, XXIV, XXVI, XXVII)
- Basement membrane collagens (type IV*)
- Fibril-associated collagens with interrupted triple helices (FACIT collagens, types IX*, XII, XIV, XVI, XIX, XX, XXI, XXII)
- Short-chain collagens, fine fibrils (types VIII*, X)
- Anchoring fibril collagen (type VII)
- Multiplexins (types XV, XVIII)
- Membrane-associated collagens with interrupted triple helices (MACIT collagens, types XIII, XVII, XXIII, XXV)
- Beaded filaments (type VI*)

*Heterotrimers containing different α chains

cretory proteins and their biosynthesis involves a unique pathway that consists of extensive post-translational modifications followed by proper chain selection, correct chain registration, and triple helix formation, and additional enzymatic processing of the folded procollagens into mature proteins (57, 63, 88). The large body of data collected over the years on several collagen types provides useful insights into the chain selection mechanism. All pro-α-chains contain one or more collagenous (COL) domains that are flanked by two NC domains. For almost all classes of collagens, the folded C-terminal NC domains act as recognition modules. They select, bind, and register the three proper α-chains by shape complementarity, electrostatic charge distribution, and hydrophobicity to assemble the trimers. An exception are the transmembrane collagens, where the N-terminal NC domain dictates the trimerization process and thus folding pathway (1, 108). The process of homo- or heterotrimerization of the C-terminal propeptides of procollagen is accompanied by intra- and interchain disulfide bond formation with the assistance of protein disulfide isomerase. Although for fibrillar collagens interchain disulfide crosslinking precedes nucleation and propagation of the native triple helix in the COL domains of the molecules (5, 19, 20, 37, 67), it is not essential for the correct chain registration and triple helix formation *in vivo* (10, 23).

If purified collagens or procollagens are denatured *in vitro*, the presence of interchain disulfide bonds is crucial for efficient and fast refolding into the triple-helical structure (4, 18). In the absence of these interchain crosslinks or of the trimerization sites, refolding of the denatured molecules is very slow and incomplete, and mis-matched triple helices differing in length and stability are formed (11, 35, 52).

Following this observation, early studies of the kinetics and mechanisms of triple-helical folding have been performed with the COL1-3 fragment of procollagen type III and the mature collagen type III. The COL1-3 fragment comprises the entire N-terminal precursor-specific region of procollagen III and contains a cystine knot resulting from interchain crosslinking of two cysteine residues separated by a dipeptide sequence. Mature collagen type III, on the other hand, contains a C-terminal cystine knot derived from two vicinal cysteine residues located at the C-terminus of the triple helix (46); the sequences around these cystine knots are shown in Fig. 1. Using these natural forms of collagen, the C- to N-propagation of the triple helix in a zipper-like mode as well as the effects of the *cis*-to-*trans* isomerization of the aminoacyl-proline/hydroxyproline bonds on the folding kinetics were characterized (3, 4, 18).

Procollagens and mature collagens as well as related fragments are difficult to handle. Therefore, it was early on recognized that synthetic collagenous peptides are more amenable to structural characterization. Hence they served as useful models to clarify principles of triple helix stability, molecular details of the collagen structure by X-ray crystallography, nuclear magnetic resonance (NMR), and other biophysical techniques, as well as to mimic biological activities (2, 16, 17, 21, 33, 34, 41, 47, 53, 95, 105). Most peptides were studied as monomers that self-associate into trimers in a concentration-dependent manner. With the early findings that interchain disulfide bridges significantly stabilize collagen fragments (4, 18), homotrimeric assembly of collagenous peptides on synthetic scaffolds has found increasing attention. Using this approach, it became possible to markedly reduce the unfavorable entropy of self-association of single-stranded collagenous peptides and thus to stabilize the triple helix (47, 53, 82). N- or C-Terminal crosslinking of collagenous peptides was accomplished by various homotrifunctional scaffolds, which were designed with spacers that differ in length and flexibility to allow a staggered alignment of the three chains in the triple-helical conformation (36, 38–40, 48–51, 54, 56, 59, 69, 97–99, 110). Selected examples of scaffolded collagenous peptides are shown in Fig. 2. Similarly, N- and C-terminal metal ion chelators were exploited for a metal-assisted trimerization and triple helix stabilization of collagen models (24, 66).

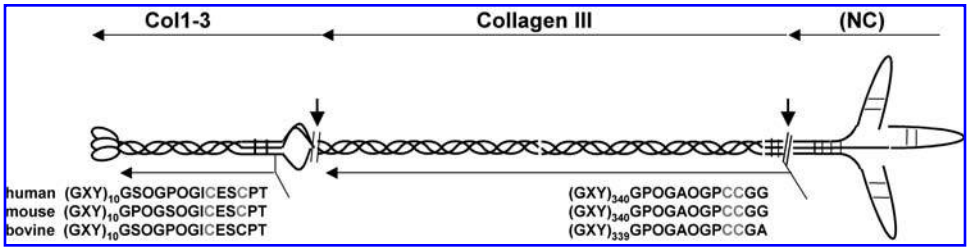

FIG. 1. Schematic representation of procollagen type III with the C-terminal NC domains and the N-terminal propeptide which are enzymatically processed during maturation leading to native collagen type III with the interchain C-terminal cystine knot. The amino acid sequences of this cystine knot and the cysteine motif at the junction between the triple helix and the NC domains of the propeptide COL1-3 of human, mouse, and bovine collagen type III are inserted. The cysteine residues are highlighted in *gray*.

It was found during these studies that the structure of the scaffold and the nature of the linker significantly affect the stability of the triple helix due to the difference in steric hindrance imposed on the three chains for the correct adjustment of the one-residue register shift. Moreover, synthetic challenges arose in the assembly of the peptides. Highly selective and sophisticated chemistry was required to obtain satisfactory yields. Crosslinking of different chains to produce the desired heterotrimers was expected to increase the synthetic difficulties in exponential manner. The obvious solution was to attempt the use of the interchain cystine knots of native collagens or related mimetic structures in order to facilitate an assembly of homo- and heterotrimeric collagen models. Advances and perspectives of this more recent approach are the subject of the present review.

HOMOTRIMERIC MODEL PEPTIDES WITH THE COLLAGEN TYPE III C-TERMINAL CYSTINE KNOT

The interchain cystine knot of the N-terminal propeptide COL1-3 of collagen type III (Fig. 1) has so far not been applied to the oxidative assembly of synthetic collagen trimers. However, the use of the C-terminal cystine knot of collagen type III for covalent homotrimeric assembly of synthetic model peptides has been successfully pioneered by Bächinger and associates (74, 80). Collagenous peptides that contain $(GXY)_n$ repeats and are C-terminally extended by the bis-cysteinyl-sequence GPCCGGG were exposed to air or GSH/GSSG oxidation at lower temperature under alkaline conditions and the disulfide-crosslinked homotrimers were obtained in satisfactory yields. Thermal denaturation confirmed the significantly enhanced stabilities of the triple-helical structures resulting from the cystine knot. This approach was then successfully applied by Engel and coworkers even for recombinant constructs containing collagen peptides such as $GS(GPP)_{10}GPPGPCCGGG$ (13) and $GSYGPPGPCCSGPP(GPP)_{10}$ (43). Upon enzymatic recovery of these peptides, oxidation was performed in the presence of GSH/GSSG (9:1) at 20°C [*i.e.*, at a temperature slightly below the triple helix melting point of $(GPP)_{10}$ (T_m = 25°C) or more appropriately of the reference peptide Ac-$(PPG)_{10}$-NH_2 (T_m = 45°C)] (22). The trimers were formed in good yields and both showed a significant increase of the thermal stability, with the C-terminal knot being more efficient (T_m = 82°C) than the N-terminal disulfide crosslink (T_m = 67°C). The lower stability induced by the N-terminal crosslink may at least partly derive from electrostatic repulsions at the C-terminus of the triple helix. Since the refolding rates of both peptides (0.00012 sec^{-1} for the C-terminal and 0.00037 sec^{-1} for the N-terminal cystine knot peptide) are very similar, triple helix formation can be nucleated at both ends. The rate-limiting step is still represented by the *cis*-to-*trans* isomerization as supported by the very similar activation energy which is >50 kJ/mol for both trimers (13, 43). These results confirmed that N-terminal triple helix nucleation with N- to C-propagation can occur as efficiently as from the C-terminus. This fact is in full agreement with previous findings from collagen models, which were N-terminally crosslinked with homotrifunctional templates (Fig. 2). It also explains the efficient *in vivo* folding of membrane-associated collagens where nucleation and triple-helix folding starts at the N-termini (1, 79, 107, 108).

It was also attempted to oxidatively form the cystine knot in a recombinant peptide that corresponds to the native C-terminal sequence of collagen type III, but with the hydroxyproline residues replaced by proline [*i.e.*, $GS(GXY)_{13}GPCCGGG$]. The peptide was obtained by thrombin cleavage of an expressed thioredoxin construct. This attempt unfortunately failed because of the low triple helix propensity of the natural sequence (14). However, using the obligatory trimeric molecule minifibritin as a fusion protein, which dictates trimerization of this collagen fragment, formation of the cystine knot was readily achieved at low temperatures in yields over 50%. Since at room temperature mainly crosslinked dimers and monomers were obtained, the necessity of the triple helix conformation with properly aligned chains prior to the oxidation of the cysteine residues was well assessed. This observation fully agrees with the results obtained in our laboratory by analysing the synthetic collagen molecules Ac-$(POG)_n$PCCGGG-NH_2 where *n* was var-

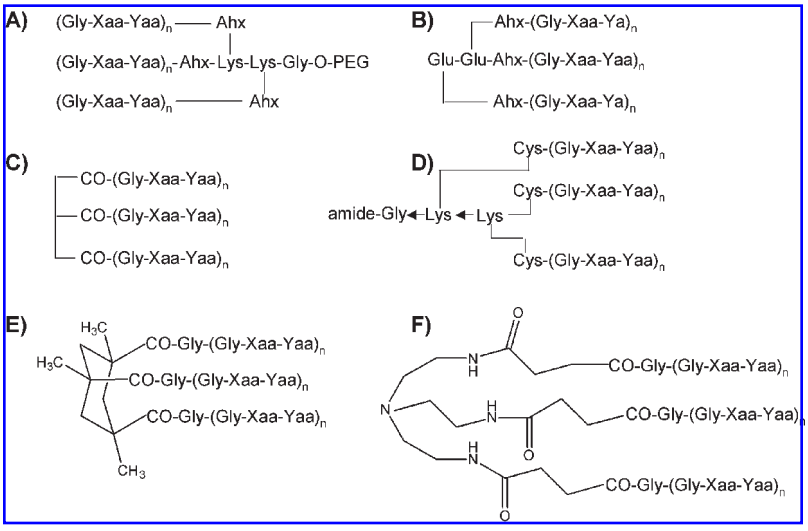

FIG. 2. Examples of homotrimeric collagenous peptides assembled on scaffolds: (**A**) di-lysine (97); (**B**) di-glutamate (54); (**C**) 1,2,3-propane tricarboxylic acid (51); (**D**) N-terminal di-lysine (110); (**E**) Kemp's triacid [cis,cis-1,3,5-trimethylcyclohexane-1,3,5-tricarboxylic acid (48)], and (**F**) tris(2-aminoethylamine) (69).

ied from 3 to 7 (Fig. 3) (7). With $n = 3$ self-association into a triple-helical structure was not achieved according to the CD spectra, even after incubation at low temperatures and at 1 m*M* concentration in order to entropically control the process. As a consequence, oxidation was found to generate a mixture of products, in which the desired trimer could not be detected by LC–MS. On the other hand, pre-equilibration of the peptide with $n = 5$ at low temperature (7–8°C) and at 1 m*M* concentration yielded a CD spectrum that reflects a high content of triple helix ($T_m = 20.3$°C). Performing the oxidation experiments at this concentration, at pH 8 and at a temperature far below the T_m value (*i.e.*, at 7–8°C), air oxidation was found to generate a product distribution consisting on the average of ~70% trimer. The remaining product was essentially oxidized monomer. The product distribution was not affected by the pH value in the range of 7 to 9, although at higher pH values, as expected, oxidation rates were enhanced. Oxidation with GSH/GSSG (peptide/GSH/GSSG ratio of 1:10:1) to possibly reshuffle unproductive intermediates was found to occur at significantly lower rates, but again without appreciable effect on the product distribution. An increase in yields of trimer formation was expected for the peptide with $n = 7$ since the triple helix stability of the self-associated trimer ($T_m = 41.6$°C) was significantly increased; but the resulting product distribution was found to be very similar with ~70% trimer.

Since LC–MS analysis of the product mixture revealed the presence of a single trimer and even the NMR spectra were consistent with a single set of disulfide connectivities, prefolding of the collagenous monomeric peptide, followed by oxidation, leads to the formation of a well-defined cystine knot isomer. However, we were unable to derive the exact cysteine pairings from NMR structural analysis (7), as it has been impossible to resolve the structure of the cystine knot of collagen type XIV by NMR conformational analysis (*vide infra*) (71). Molecular dynamic (MD) simulations performed on the trimer with the two most plausible cystine knots, shown in Fig. 3, would suggest that neither of the two is well compatible with a triple-helical conformation, but both types of disulfide connectivities allow a triple-helical conformation of the (POG) triplet directly adjacent to the knot. Whereas the type Y cystine knot (Fig. 3) seems more compatible with the triple-helical conformation in terms of dihedral angles, type X represents the intuitively simplest way of forming the three disulfide bonds (7).

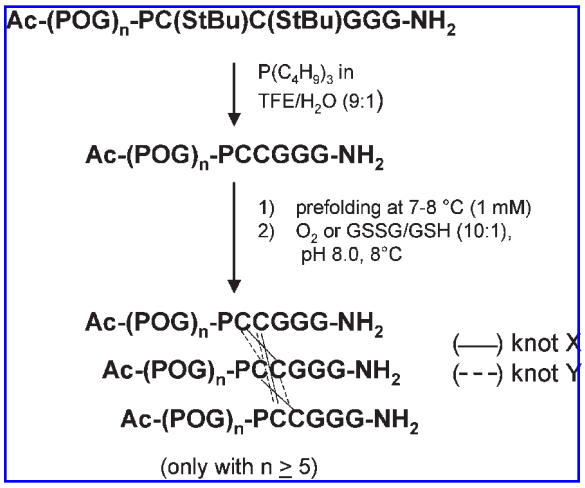

FIG. 3. Oxidative folding of collagen model peptides containing exclusively (POG) repeats and the C-terminal biscysteinyl motif of collagen type III (7).

Although for the Ac-$(POG)_5$CCGGG-NH_2 construct crystals were successfully grown (Fig. 4), their diffraction pattern was unfortunately insufficient to resolve the structure of the cystine knot. Most interestingly, the Ac-$(POG)_7$CCGGG-NH_2 trimer formed fibrils suggesting that this triple helix size may possibly represent the threshold-length for self-association of triple helices into ordered supramolecular aggregates (Fig. 4).

The formation of the intramolecular disulfide bridge of the monomer as the main side product of the oxidative assembly of homotrimers was completely unexpected because such disulfide bridges between adjacent cysteine residues are known to be conformationally very disfavored (26, 29, 32, 92). These are rarely found in folded proteins (27, 112, 115), but their formation is observed in folding pathways of cysteine-rich peptides where vicinal disulfides are found in productive intermediates that readily undergo disulfide reshuffling because of the thermodynamically disfavored structure (15, 27, 28).

Two case-limiting mechanisms for oxidative folding that allow for structural interpretation (*i.e.*, the folded-precursor and the quasi-stochastic mechanism) have emerged from extensive studies on cysteine-rich peptides and proteins (113). In the folded precursor mechanism, local regions of the polypeptide chain adopt native-like structures which are locked in by disulfide bonds; a piecemeal accumulation of such local microfolding events leads to global folding. In the quasi-stochastic mechanism on the other hand, a set of disulfide bonds forms initially following the proximity rule for loop formation in early fast oxidation and reshuffling steps. This conformation is then locked in and protected from further rearrangements by local or global conformational folding. While this second mechanism applies well to most of the folding pathways of globular proteins, formation of the type III collagen disulfide knot, at least in model peptides, follows the folded precursor mechanism as it requires a proper spatial alignment of all the cysteine residues in the triple helix. Since the folded triple helix is in equilibrium with the monomeric state, competing disulfide formation according to the proximity rule (25) leads to the oxidized monomer, most likely a kinetically trapped intermediate.

We observed that upon folding of a collagen model peptide containing the $\alpha 1\beta 1$ and $\alpha 2\beta 1$ integrin adhesion motif of collagen type I into the triple helix conformation, oxidation almost exclusively produced the disulfide-crossbridged homotrimer (Fig. 5) (9). This result was surprising as the triple helix of the self-associated monomer exhibited a T_m value of only 17°C [*i.e.*, even slightly lower than those of the collagenous peptides of Fig. 3 with $n = 5$ ($T_m = 20.3$°C) or $n = 7$ ($T_m = 41.6$°C)]. The residues Asp/Glu and Arg/Lys are often found in collagens at positions X and Y, respectively (93). Although such residues reduce the triple helix stability as shown for the $(GER)_{15}$CPCCG trimer (80), as guest triplets of the type (GPR), (GEO), or (GER) in $(GPO)_n$ sequences the negative effects on thermal stabilities of the triple helical self-associated host–guest peptides are largely compensated by pH-

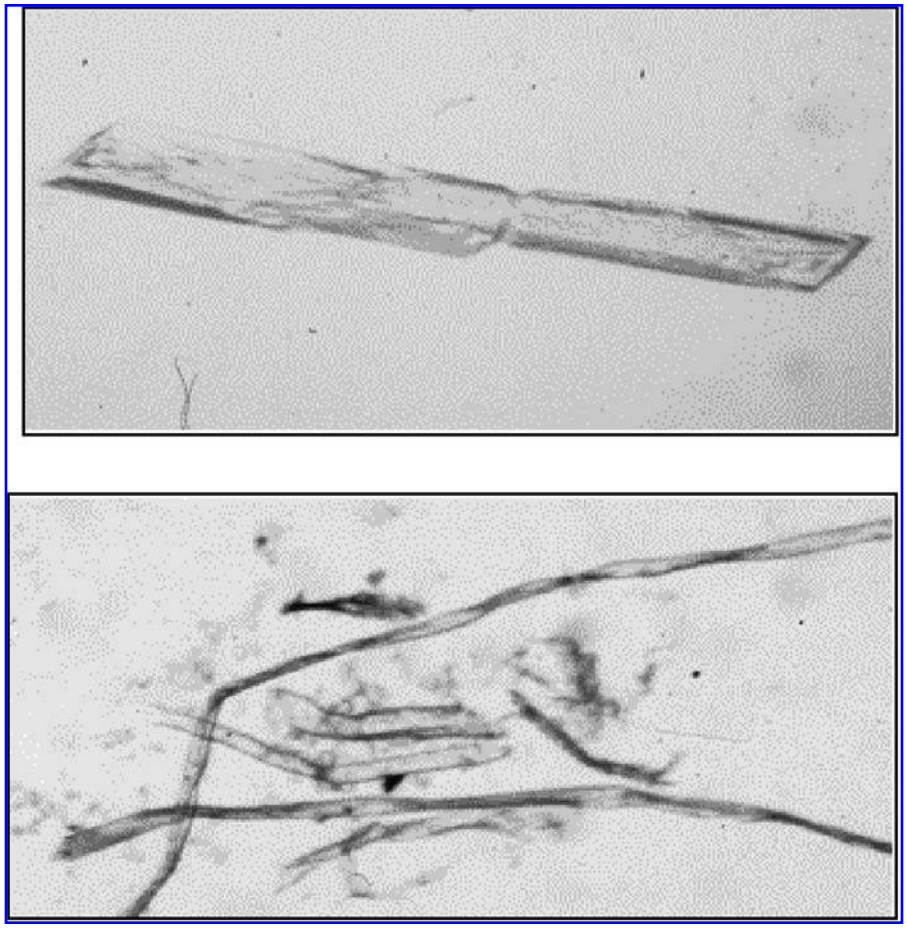

FIG. 4. Crystals of the disulfide crosslinked homotrimer Ac-$(POG)_5CCGGG\text{-}NH_2$ (*top*) and fibers of Ac-$(POG)_7CCGGG\text{-}NH_2$ (*bottom*) (7).

dependent electrostatic interactions between these charged residues, and possibly by hydrogen bonding with peptide backbones and interactions with the water shell (94, 116). In the model peptide of Fig. 5, the Arg and Glu residues may favor the staggered alignment by electrostatic interactions without enhancing *per se* the thermal stability of the triple helix. This would increase the probability of the cysteine residues being in juxtaposition for the correct disulfide formation. Such electrostatic interactions were exploited most recently in order to successfully self-assemble heterotrimeric collagen triple helices from chains containing Arg in position Y of one chain, Glu in X of the second chain, while the third peptide strand consisted of only (GPO) repeats (45).

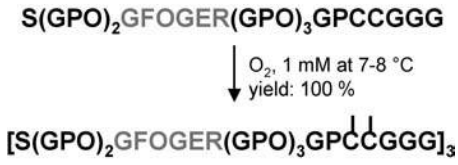

FIG. 5. Oxidative folding of a collagen peptide containing the $\alpha1\beta1$ and $\alpha2\beta1$ integrin adhesion motif (*gray*) and the C-terminal collagen type III cysteine motif produced the homotrimer in almost quantitative yield.

CYSTINE KNOTS OF FACIT COL1-NC1 JUNCTIONS

FACITs (fibril-associated collagen with interrupted triple helices) (Table 1) contain short triple-helical domains interdispersed with nontriple-helical domains. Collagen type IX is composed of three different α-chains while all others are homotrimers. Unlike the fibril-forming collagens, the FACITs have significantly shorter C-terminal NC domains, but share a remarkable sequence homology in the first collagenous domain (COL1) with two imperfections within the triple helix. Moreover, the FACITs contain in their COL1/NC1 junctions two strictly conserved cysteine residues separated by four residues (Table 2). In the folded molecules, these cysteine residues form an interchain disulfide knot, so far, of unknown disulfide connectivities. Extensive studies of *in vivo* folding of different constructs of the homotrimeric collagens XII and XIV strongly support the notion that the COL1 domain at the COL1/NC1 junction, which contains the first cysteine residue (Table 2) and the NC1 domain containing the second conserved cysteine are involved in the mechanism of chain selection and registration and thus in the assembly of FACITs (73, 75–78).

These studies provide evidence that by preventing 4-hydroxylation of the Pro residues, the interchain disulfide bridges are not formed and the monomers are secreted by the cultured cells. However, it was observed in the presence of ascorbate,

Table 2. Human and Chicken cDNA-Derived Sequences of the NC1 Domains

Type IX Collagen			
α1 chicken (P12106)	GPPGPPGYCEPSSCRMQAGQRAAGKNMKGP		920
α1 human (P20849)	GPPGLPGFCEPASCTMQAGQRAFNKGPDP	921	
α2 chicken (P12108)	GPAGLPGFCEPAACLGALPTPRHG		677
α2 human (Q14055)	GPVGLPGFCEPAACLGASAYASARLTEPGSIKGP	689	
α3 chicken (P32017)	GAQGTPGICDTSACMGAVGASTSKKS		675
α3 human (Q14050)	GAQGTPGICDTSACQGAVLGGVGEKSGSRSS		684
Type XII Collagen			
α1 chicken (P13944)	GPPGPPGYCDSSQCASIPYNGQGFPEPYVPESGPY…		
Unprocessed precursor 3124 AA	3048 3053		
α1 human (Q99715)	GPPGPPGYCDSSQCASIPYNGQSYPGSG	3063	
Unprocessed precursor 3063 AA	3044 3049		
Type XIV Collagen			
α3 chicken (P32018)	GPAGPPGYCDPSSCAGYGMGGGYGEPTDQDIPVVQL…		
Unprocessed precursor 1888 AA	1769 1774		
α1 human (Q05707)	GPSGQPGYCDPSSCSAYGVRAPHPDQPEFTPVQDELEA…		
Unprocessed precursor 1796AA	1759 1764		

Included are the C-terminus of the COL1 domains of the α-chains of the heterotrimeric collagen type IX and the homotrimeric collagen type XII and XIV (SwissProt data bank, entry codes are given in brackets). The Pro residues in positions Y of the collagenous domains are most probably (4*R*)-hydroxylated in the mature collagens.

which is essential for proline hydroxylation, that homotrimeric collagen type XII constructs were obtained. The results of the *in vivo* folding experiments with constructs consisting mainly of the COL1 domain or of the NC1 domain would suggest that both domains retain trimerization properties to some extents. However, for an efficient *in vivo* trimerization of the α-chains in the correct register and thus oxidative formation of the native cystine knot, the sequence-encoded information of both domains is concurrently required.

Homotrimeric model peptides with the cystine knot of collagen type XIV

To investigate the intrinsic propensity of the highly conserved cystine knot sequences of FACITs (Table 2) for oxidative trimerization *in vitro*, van der Rest and coworkers have used a synthetic model peptide derived from the COL1/NC1 junction of collagen type XIV (compound I of Fig. 6) (71). This peptide was N-terminally extended with three (GPO) repeats to replace the GSQ–GPA–GPO sequence of type XIV collagen (compound II of Fig. 6). From studies of synthetic collagen model peptides (34, 95, 105) and native procollagen (12, 30, 86), such (GPO) repeats are known to most efficiently stabilize triple helices; moreover, it was determined that the C-terminus of COL1 domain of the α1-chain of collagen type IX consists of three (GPO) triplets (87).

All oxidative experiments performed with the bis-cysteinyl sequence itself (peptide I of Fig. 6) failed to produce the homotrimer in satisfactory yields. The majority of the oxidized product was the intrachain disulfide-bridged monomer. The high tendency of the reduced monomeric peptide I to oxidize intramolecularly to Ia is in full agreement with the amino acid composition between the two Cys residues as these residues among others most frequently occur in β-turns (31). Conversely, oxidative refolding of the N-terminally extended peptide II was found to produce the monomers IIa, the dimer IIb, and the trimer IIc at ratios strongly depending upon the experimental conditions. At room temperature, the intrachain disulfide-bridged monomer was by far the predominant product, whereas the trimer formation was minimal. By lowering the temperature and particularly by exploiting the known triple-helix stabilizing effect of alcohols (17, 34), formation of the trimer IIc was found to increase significantly. At a temperature of −20°C and with 30% MeOH in the buffer (pH 8.5) the trimer composed up to 55% of the product mixture. When the incubation time was extended under these conditions, a fast scrambling of the trimer to the thermodynamically more stable monomer IIa was observed. Similarly negative was the effect of diluting below 1 mg/mL (*i.e.*, by increasing the entropic penalty of self-association of the three peptide chains in the correct register). Indeed, at higher dilution the trimer IIc was only formed in trace amounts. Therefore, trimerization into the triple helix register via the $(GPO)_3$ repeats is apparently required for oxidative formation of the disulfide knot. The competing formation of the intramolecular disulfide bonds and a time-dependent scrambling of the disulfides in the presence of GSH/GSSG to form the oxidized monomer represent the main side reactions. The triple-helical structure of the trimer IIc as isolated product was confirmed by the CD spectrum and NMR structural analysis (71).

Goodman and coworkers have shown that with $n = 3$, only an incipient triple helix is formed in aqueous environments when collagenous peptides were N-terminally crosslinked with Kemp triacid and contained an increasing number of $(GPO)_n$ repeats (48). In light of this information, it can be hypothesized that the concentration-dependent self-associated triple-helical trimers of II are in equilibrium with the monomeric form. The correctly aligned trimer is removed from the equilibrium mixture by formation of the cystine knot in an oxidation reaction that is competing with the conformationally highly favored intramolecular cyclization to IIa. Thus it could be expected that an increase of the triple-helix

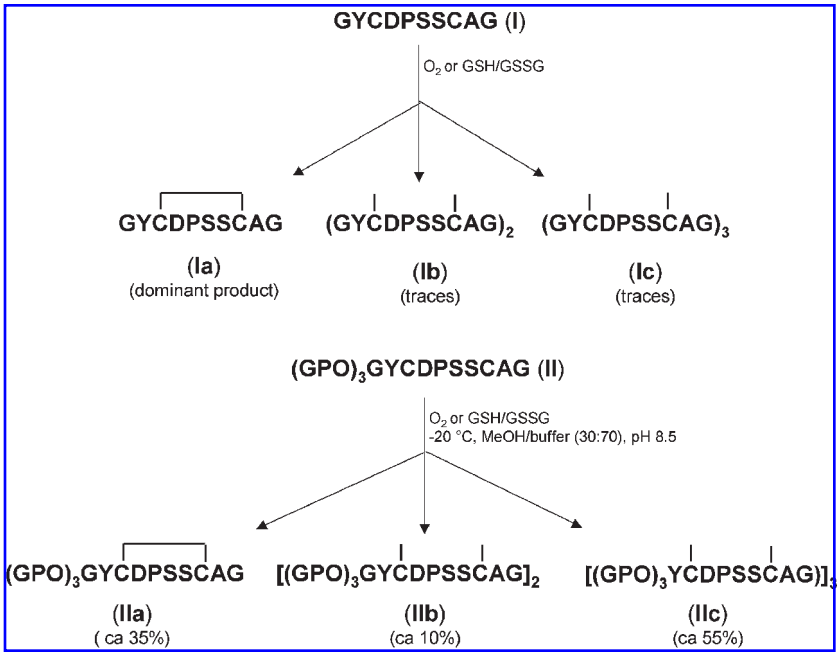

FIG. 6. Product distributions in the oxidative folding of peptides related to the collagen type XIV cystine knot (71).

stability may significantly reduce the content of oxidized monomer as a side product.

Disulfide crosslinking of three bis-cysteinyl-peptides leads to the two isomers A and B shown in Fig. 7. Their number, however, increases when the peptides are aligned in a collagen-type register by a one-residue shift as the chains are then no longer spatially equivalent. Six conformers are then possible for the disulfide knot A and two for the disulfide knot B. NMR conformational analysis of the trimer IIc confirmed the formation of a type A cystine knot although an unambiguous differentiation between the cystine knots A1 and A2 (Fig. 7) could not be made (71). Interestingly, the disulfides between the last Cys residue of the middle chain and the first Cys of the leading chain (isomer A2) or of the trailing chain (isomer A1) induce a locked-in bent structure that is very similar to the β-turn structure in the intrachain disulfide-bridged monomer Ia.

Heterotrimeric model peptides with the cystine knot of collagen type IX

A mini-α-chain of FACIT collagen type XII, consisting of the COL1 domain and the N-terminal portion of the NC1 domain (Table 2), has been successfully self-assembled into a triple-helical homotrimer (76). Based on these findings, a similarly low molecular weight pepsin-resistant fragment of the heterotrimeric FACIT type IX was prepared (70). Upon reduction and isolation of the three truncated α-chains, these were then individually analyzed for their *in vitro* oxidative self-association into homotrimers, and as stoichiometric mixture of all three α-chains into a heterotrimer. All three chains are able to form triple-helical homotrimeric reassociation products. But most surprising was the observation that a mixture of the α1, α2, and α3 chain in the presence of reduced and oxidized glutathione (GSH/GSSG) leads to an oxidized product distribution consisting of the heterotrimer $\alpha 1\alpha 2\alpha 3$ in addition to the homodimers and homotrimers of the α1 and α2 chain, while homodimers and homotrimers of the α3 chain are formed to much lesser extents. Although the heterotrimer $\alpha 1\alpha 2\alpha 3$ constitutes maximally 25% of the product mixture, this value is significantly higher than that expected from a random oxidation of the three chains, which should lead to 16 species (6 dimers and 10 trimers). Whether the heterotrimer represents a mixture of all the possible conformers resulting from different chain alignments could not be determined in this study. Nevertheless, in view of the sequence composition around the cystine knot (Table 2), it is rather surprising that in absence of most of the noncollagenous C-terminal domain NC1, but retaining its cysteine-containing N-terminus, sufficient information is encoded in the three chains for a preferred self-association into the correctly registered triple-helical and interchain-crosslinked heterotrimer (70).

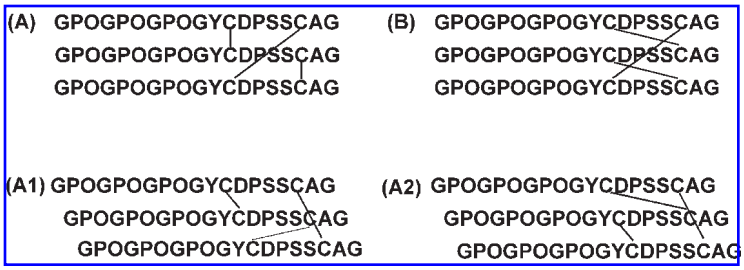

FIG. 7. Trimers A and B contain the possible cystine knot between three identical bis-cysteinyl peptides without a defined sequence alignment. Taking into account a collagen-type chain register, six conformers are possible for type A and two for type B cysteine connectivities. The conformers A1 and A2 belong to the type A and represent the two indistinguishable cystine knots determined by NMR for the trimer IIc (71).

These findings would suggest a more or less explicit involvement of the COL1 domains in the preferred heterotrimeric assembly of collagen type IX, which does not coincide with the self-association and oxidation experiments performed with synthetic peptides corresponding to the sequences of the three α-chains of the COL1/NC1 junction and the short NC1 domains of human collagen type IX (Table 2) (81). A stoichiometric mixture of these three chains was unable to self-assemble into a heterotrimer. Therefore, the peptides were N-terminally extended with the adjacent natural $(GXY)_5$ sequences of the COL1 domains with all the proline residues in the Y position in the 4-hydroxylated form. Operating at 25°C where CD spectra confirmed the absence of a triple helical conformation of the short COL1 extensions, the heterotrimer comprised 10% of the mixture among monomers, dimers, and higher oligomers. By lowering the temperature to 4°C, the N-terminal short triple helix is induced in the isolated trimer, but at least with the entire NC1 domain a prefolding does not seem to be essential for the selective heterotrimerization and cystine knot formation. With heterotrimer-specific antibodies, even the correct registration of the three α1, α2, and α3 NC1 domains was confirmed (81).

These most interesting results of Bächinger and associates were so far not applied to the production of synthetic or recombinant heterotrimeric collagen models. Yields of disulfide-crossbridged and thus stabilized triple-helical heterotrimers may possibly be enhanced by incorporation of specific interchain interactions, such as the electrostatic interactions used recently for self-association of three different synthetic α-chains into collagenous heterotrimers (45).

FOLDONS AS TRIMERIZATION DOMAINS

In addition to the synthetic scaffolds such as those depicted in Fig. 2 for covalent homotrimerization of collagenous peptides, foldon domains can equally be used for noncovalent assembly of trimers. Indeed by C- or N-terminal extension of the model peptide $(GPP)_{10}$ with the short foldon domain of bacteriophage T4 fibritin, a remarkable thermal stabilization of the triple helix could be achieved (13, 43, 44). However, compared to a C- or N-terminal cystine knot of collagen type III, the observed effect was weaker. This foldon domain forms a β-propeller-like structure with a hydrophobic interior that originates from the small β-sheet formed by three-fold-related β-hairpins (111). It folds autonomously and reversibly into an obligatory trimer (72), but it shows sufficient plasticity to adapt to the requirements of a triple helix at the interface. This was well assessed by the X-ray structure of the $(GPP)_{10}$-foldon construct (109). Because of its relatively small size (27 residues) it represents an interesting tool in the design of homotrimeric synthetic and recombinant collagen molecules. The usefulness of this foldon domain was particularly evident in the case of collagenous peptides with low triple helix propensity. All attempts to install the cystine knot in the recombinant construct of thioredoxin A containing the C-terminal collagenous sequence GS-$(GXY)_{13}$-GPCCGGG corresponding to the C-terminus of collagen type III and expressed in *E. coli* with the proline residues in the nonhydroxylated form have failed (14). However, using a construct with the bacteriophage T4 minifibritin, air oxidation produced the disulfide crossbridged homotrimer in satisfactory yields. A similarly efficient approach could possibly be derived from a three-stranded coiled-coil as demonstrated for folding of the recombinant N-propeptide of type IIA collagen (79).

COLLAGEN MOLECULES WITH ARTIFICIAL CYSTINE KNOTS

Except for the COL1/NC1 junction of collagen type IX (81), all the trimerization procedures discussed so far are limited to the preparation of homotrimeric collagen molecules. Therefore, for studies of structural and functional properties of heterotrimeric collagens such as collagen type I and IV, a new strategy was required that would allow crosslinking of three peptide chains in a regioselective manner to establish the correct registration of the α-chains. For this purpose, a simplified artificial C-terminal cystine knot was designed in our laboratory and produced by regioselective cysteine pairing procedures to align the peptide chains in the desired register and with minimal steric clashes within the triple helix (89). As shown in Fig. 8, the assembly of the heterotrimers from the side-chain deprotected single chains is based on regioselective cysteine activation as pyridyl- or nitropyridylsulfenyl derivatives followed by thiol-disulfide exchange reactions in slightly acidic aqueous media to prevent scrambling of already formed disulfides (89, 91).

A more facile synthetic route was proposed by Farndale and coworkers, which is based on the use of thiol-activated species of the monocysteinyl-α1-chains in excess over the bis-cysteinyl α2-chain (106). The incorporation of the artificial cystine knot is then controlled by reaction kinetics rather than regioselective sulfur chemistry. Using this procedure, the authors were able to achieve formation of the desired heterotrimer in up to 30% yield, although extensive chromatographic separations were required to isolate the trimer from the still complex mixture of reaction products.

Various heterotrimers containing the collagenase cleavage sequences 772–784 and 772–785 of the α1 and α2 chain of collagen type I, respectively, were obtained by the regioselective synthetic strategy. N-Terminal or N- and C-terminal extensions of the natural sequences with $(POG)_n$ repeats served to induce and stabilize the triple-helical structure (91). The trimers were folded into triple helices with sharp cooperative thermal transitions from the folded state to random coil. The heterotrimers proved to be well-suited substrates for analyzing the different mechanisms of digestion of collagen and gelatine by collagenase and gelatinase (90).

In order to analyze whether peptides that contain the native sequences of the heterotrimeric collagen type I retain sufficient information for self-association at least to some extents into a heterotrimer under the influence of a $(POG)_5$ extension, reduction and oxidative refolding experiments were performed at 4°C. The heterotrimer shown in Fig. 9 is characterized by a T_m value of 32°C. As well assessed by LC–MS, only trace amounts of the heterotrimer, even far below the statistical product distribution, were detected (89). In this case, a negatively charged side chain as counterpart for electrostatic interactions with the

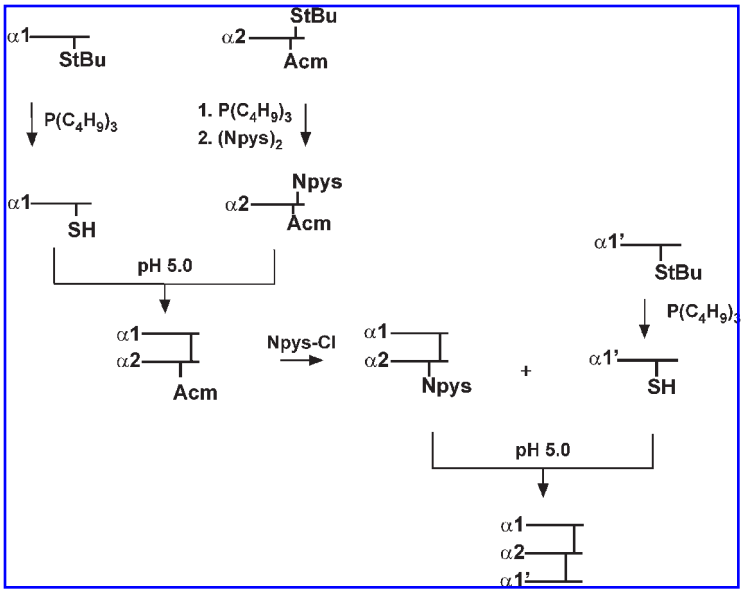

FIG. 8. Regioselective assembly of the α1, α2, and α1′ cysteine peptides with an artificial cystine knot into heterotrimers with the α1α2α1′ register (91).

Arg residue located on two chains and thus capable of favoring heterotrimerization is missing.

The triple-helical fold of the disulfide-crosslinked heterotrimer of Fig. 9 may therefore be attributed primarily to the $(POG)_5$ sequence and at negligible extent to a specific recognition and association of the native sequences. This was confirmed by a detailed NMR conformational analysis of a similar heterotrimer, which confirmed the presence of a triple helix only in the $(POG)_5$ extension, whereas a loose chain alignment was determined for the collagenase-cleavage epitope (42). Such findings clearly confirm that natural collagen sequences do not in all cases retain sufficient structural information for correct alignment into heterotrimers and thus for their oxidative cross-bridging with cystine knots. Possibly, the heterotrimeric cystine knot present in FACIT type IX (Table 2) could act more efficiently as a helping hand for self-association into heterotrimers stabilized by the subsequent cystine knot formation than the artificial cystine framework.

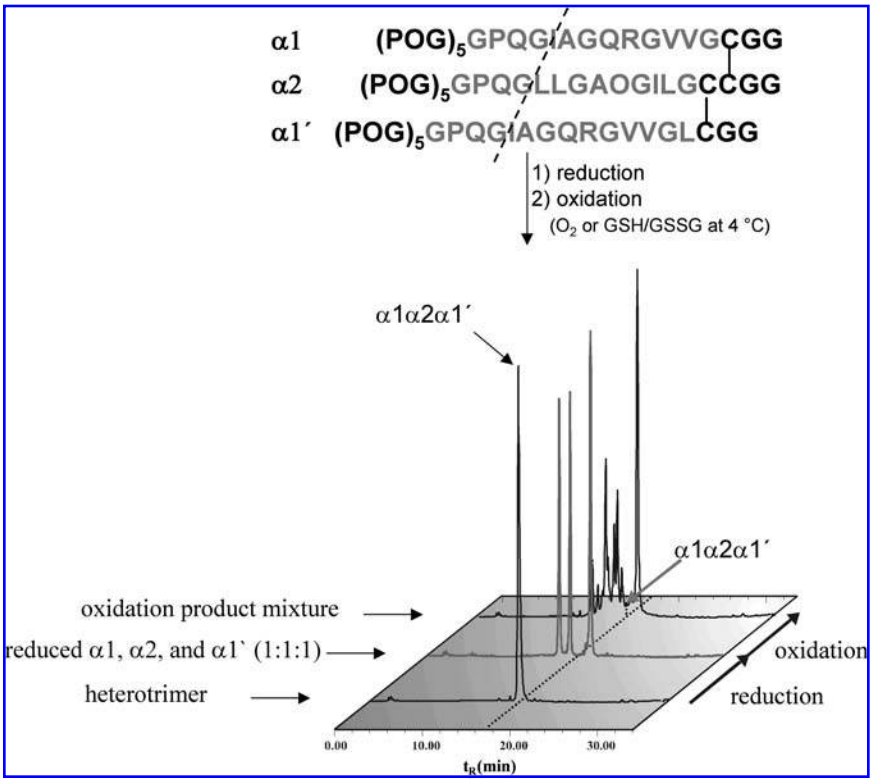

FIG. 9. Reduction and oxidative refolding of a collagenous heterotrimer with an artificial cystine knot (89). The natural collagenase cleavage site sequences of collagen typ I are highlighted in *gray*.

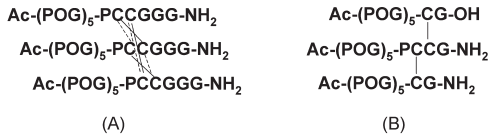

FIG. 10. Collagen models with (A) the native collagen type III and (B) an artificial cystine knot. The artificial cystine knot contains well-defined disulfide connectivities whereas the disulfide pairings of the native cystine knot are still not unambiguously assigned (7).

The sequence of reactions required for regioselective disulfide assembly of the heterotrimers in our hands was found to proceed smoothly as long as the single chains or dimers were lacking a higher tendency to self-associate into triple helices (84, 89, 91, 102). In fact, with increasing propensity for triple helix formation in aqueous solution, conformation-dependent side reactions were encountered (100), which became critical in the synthesis of the heterotrimer consisting of only five (POG) triplets adjacent to the artificial cystine knot (trimer B of Fig. 10) (8). Attempts to crosslink the unprotected chains with disulfides in successive steps failed completely when the route described in Fig. 8 was followed. However, operating with O-*tert*-butyl protected peptides, which are unable to self-associate into homotrimers, the heterotrimer was obtained in satisfactory yields (8). This heterotrimer represents the counterpart of the homotrimer with the collagen type III cystine knot of Fig. 3, as both differ only in the type of cystine framework. The simplified and thus more flexible artificial cystine knot of the trimer B (Fig. 10) leads to a thermal stability of the triple helix (T_m = 55.7°C) only slightly inferior to the more rigid native collagen type III knot of trimer A (T_m = 66.2°C) (7). NMR structural analysis of the artificial cystine knot B of Fig. 10 was supportive for an extension of the triple-helical hydrogen bonding network into the crosslinked cysteines, a fact that may affect the accessibility of larger size reagents required in the regioselective sulfur chemistry. On the other hand, the ability of our artificial cystine knot to adapt itself to the triple helix has been successfully exploited most recently in the synthesis of self-complimentary collagenous peptides (Fig. 11) that are able to aggregate into collagen-like triple-helical structures of sizes mimicking those of natural collagen (62, 68).

CONCLUSIONS

Because of the difficulties in handling native collagens and related fragments, both homo- and heterotrimeric collagen models stabilized by cysteine knots proved to be the most suited molecules for studying the folding kinetics of triple helices and for mimicking functional epitopes of native collagens. In particular, the understanding of folding kinetics gained in early studies on collagen and collagen fragments was substantially increased by studying model collagenous peptides (2, 6, 13). Similarly, interesting new insights were obtained into the spatial display of single side chains on the rod-shaped molecules, which are required for selective recognition by interacting proteins (61, 64, 65, 83, 101, 103, 104), and into the structural effects of single mutations known to generate serious pathological effects in humans (16, 35, 96). Whereas the homotrimeric collagenous peptides that are crosslinked N- or more effectively C-terminally with the collagen type III cystine knot, are readily accessible, the production of collagenous heterotrimers still remains a challenging task. Possible new perspectives may derive from the proper use of the heterotrimeric collagen type IX cystine knot, although a well-defined registration of the chains by this approach requires more detailed insights into the role of the NC1 and/or COL1 domain in chain selection. Since the registration of the α-chains has been found to significantly affect the plasticity of triple-helical functional epitopes in model heterotrimers (103), a defined chain alignment so far can only be achieved with regioselective interchain disulfide formation. Moreover, as the use of bovine collagen in clinical applications involves the risk of prion infection or even gelatine-related allergies (55, 60, 114), the development of synthetic or recombinant collagen molecules has become a task of great priority. In this context, thermally more stable artificial collagen-like fibrils may possibly be produced as new biomaterials by the use of native or artificial cystine knots.

ACKNOWLEDGMENTS

This work was supported by the EU grant Therapeuskin LSHB-CT-2005-511974.

ABBREVIATIONS

CD, circular dichroism; COL, triple-helical domains in collagens; FACIT, fibril-associated collagens with interrupted triple helices; GSH, glutathione; GSSG, oxidized glutathione; LC–MS, mass spectrometry coupled to liquid chromatography; MACIT, membrane-associated collagens with interrupted triple helices; MD, molecular dynamics; NC, noncollagenous domains; NMR, nuclear magnetic resonance; O, (4*R*)-hydroxyproline.

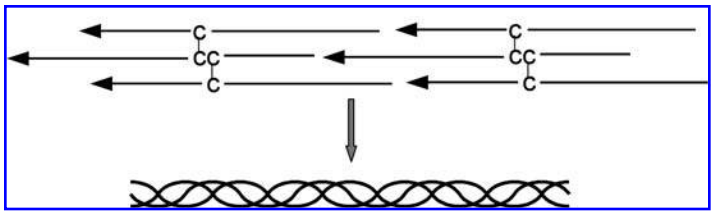

FIG. 11. Self-assembly of heterotrimeric collagenous peptides into larger-size triple helices (62, 68).

REFERENCES

1. Areida SK, Reinhardt DP, Müller PK, Fietzek PP, Köwitz J, Marinkovich MP, and Notbohm H. Properties of the collagen type XVII ectodomain. Evidence for N- to C-terminal triple helix folding. *J Biol Chem* 276:1594–1601, 2001.
2. Bächinger HP and Engel J. The thermodynamics and kinetics of collagen folding. In: *Protein Folding Handbook*, edited by Buchner, J, and Kiefhaber, T. Weinheim, Germany: Wiley-VCH, 2005, Vol. 2, pp. 1059–1110.
3. Bächinger HP, Bruckner P, Timpl R, and Engel J. Role of *cis-trans* isomerization of peptide bonds in coil ↔ triple helix conversion of collagen. *Eur J Biochem* 90: 605–613, 1978.
4. Bächinger HP, Bruckner P, Timpl R, Prockop DJ, and Engel J. Folding mechanism of the triple helix in type-III collagen and type-III pN-collagen. Role of disulfide bridges and peptide-bond isomerization. *Eur J Biochem* 106: 619–632, 1980.
5. Bächinger HP, Fessler LI, Timpl R, and Fessler JH. Chain assembly intermediate in the biosynthesis of type-III procollagen in chick-embryo blood-vessels. *J Biol Chem* 256: 3193–3199, 1981.
6. Bachmann A, Kiefhaber T, Boudko S, Engel J, and Bächinger HP. Collagen triple-helix formation in *all-trans* chains proceeds by a nucleation/growth mechanism with a purely entropic barrier. *Proc Natl Acad Sci USA* 102: 13897–13902, 2005.
7. Barth D, Kyrieleis O, Frank S, Renner C, and Moroder L. The role of cystine knots in collagen folding and stability. Part II. Conformational properties of $(Pro\text{-}Hyp\text{-}Gly)_n$ model trimers with N- and C-terminal collagen type III cystine knots. *Chem Eur J* 9: 3703–3714, 2003.
8. Barth D, Musiol HJ, Schütt M, Fiori S, Milbradt AG, Renner C, and Moroder L. The role of cystine knots in collagen folding and stability. Part I. Conformational properties of $(Pro\text{-}Hyp\text{-}Gly)_5$ and $(Pro\text{-}(4S)\text{-}FPro\text{-}Gly)_5$ model trimers with an artificial cystine knot. *Chem Eur J* 9: 3692–3702, 2003.
9. Barth L, Sinner EK, Cadamur SA. Renner C, Oesterhelt D, and Moroder L. Homotrimeric collagen peptides as model systems for cell adhesion studies. In: *Peptides for Youth*, edited by Escher E, Lubell WD, and Del Valle S. American Peptide Society, in press.
10. Beck K, Boswell BA, Ridgway CC, Bächinger HP. Triple helix formation of procollagen type I can occur at the rough endoplasmic reticulum membrane. *J Biol Chem* 271: 21566–21573, 1996.
11. Beier G and Engel J. Renaturation of soluble collagen—Products formed at different temperatures. *Biochemistry* 5: 2744–2755, 1966.
12. Berg RA and Prockop DJ. Purification of carbon-14-labeled protocollagen and its hydroxylation by prolyl-hydroxylase. *Biochemistry* 12: 3395–3401, 1973.
13. Boudko S, Frank S, Kammerer RA, Stetefeld J, Schulthess T, Landwehr R, Lustig A, Bächinger HP, and Engel J. Nucleation and propagation of the collagen triple helix in single-chain and trimerized peptides— Transition from third to first order kinetics. *J Mol Biol* 317: 459–470, 2002.
14. Boudko SP and Engel J. Structure formation in the C-terminus of type III collagen guides disulfide cross-linking. *J Mol Biol* 335: 1289–1297, 2004.
15. Boulègue C, Milbradt AG, Renner C, and Moroder L. Single proline residues can dictate the oxidative folding pathways of cysteine-rich peptides. *J Mol Biol* 358: 846–856, 2006
16. Brodsky B and Persikov AV. Molecular structure of the collagen triple helix. *Adv Prot Chem* 70: 301–339, 2005.
17. Brown FR, Blout ER, Corato AD, and Lorenzi GP. Synthesis and structural studies of two collagen analogs: poly-(L-Prolyl-L-Seryl-Glycyl) and poly-(L-Prolyl-L-Alanyl-Glycyl). *J Mol Biol* 63: 85–99, 1972.
18. Bruckner P, Bächinger HP, Timpl R, and Engel J. Three Conformationally distinct domains in amino-terminal segment of type-III procollagen and its rapid triple helix (coil transition. *Eur J Biochem* 90: 595–603, 1978.
19. Bruckner P and Eikenberry EF. Formation of the triple helix of type-I procollagen in *cellulo*—Temperature-dependent kinetics support a model based on *cis-trans* isomerization of peptide bonds. *Eur J Biochem* 140: 391–395, 1984.
20. Bruckner P, Eikenberry EF, and Prockop DJ. Formation of the triple helix of type-I procollagen in *cellulo*—A kinetic model based on *cis-trans* isomerization of peptide-bonds. *Eur J Biochem* 118: 607–613, 1981.
21. Bruckner P, Rutschmann B, Engel J, and Rothe M. Classical synthesis of collagen-like peptides with the sequence $Z(GlyPro\text{-}Pro)_nOBu^t$ and their characterization with circular-dichroism and ultracentrifugation. *Helv Chim Acta* 58: 1276–1287, 1975.
22. Buevich AV and Baum J. Residue-specific real-time NMR diffusion experiments define the association states of proteins during folding. *J Am Chem Soc* 124: 7156–7162, 2002.
23. Bulleid NJ, Dalley JA, and Lees JF. The C-propeptide domain of procollagen can be replaced with a transmembrane domain without affecting trimer formation or collagen triple helix folding during biosynthesis. *EMBO J* 16: 6694–6701, 1997.
24. Cai WB, Kwok SW, Taulane JP, and Goodman M. Metal-assisted assembly and stabilization of collagen-like triple helices. *J Am Chem Soc* 126: 15030–15031, 2004.
25. Camacho CJ and Thirumalai D. Theoretical predictions of folding pathways by using the proximity rule, with applications to bovine pancreatic trypsin-inhibitor. *Proc Natl Acad Sci USA* 92: 1277–1281, 1995.
26. Capasso S, Mattia C, Mazzarella L, and Puliti R. Structure of a *cis*-peptide unit—Molecular-conformation of cyclic disulfide L-cysteinyl-L-cysteine. *Acta Crystallogr B* 33: 2080–2083, 1977.
27. Carugo O, Čemažar M, Zahariev S, Hudáky I, Gáspári Z, Perczel A, and Pongor S. Vicinal disulfide turns. *Protein Engineer* 16: 637–639, 2003.
28. Čemažar M, Zahariev S, Lopez JJ, Carugo O, Jones JA, Hore PJ, and Pongor S. Oxidative folding intermediates with nonnative disulfide bridges between adjacent cysteine residues. *Proc Natl Acad Sci USA* 100: 5754–5759, 2003.
29. Chandrasekaran R and Balasubramamian R. Stereochemical studies of cyclic peptides. 4. Energy calculations of cyclic disulphide cysteinylcysteine. *Biochim Biophys Acta* 188:1–9, 1969.
30. Chopra RK and Ananthanarayanan VS. Conformational implications of enzymatic proline hydroxylation in collagen. *Proc Natl Acad Sci USA* 79: 7180–7184, 1982.
31. Chou PY and Fasman GD. Empirical predictions of protein conformation. *Annu Rev Biochem* 47: 251–276, 1978.
32. Creighton CJ, Reynolds CH, Lee DHS, Leo GC, and Reitz AB. Conformational analysis of the eight-membered ring of the oxidized cysteinyl-cysteine unit implicated in nicotinic acetylcholine receptor ligand recognition. *J Am Chem Soc* 123: 12664–12669, 2001.
33. Engel J and Bächinger HP. Structure, stability and folding of the collagen triple helix. *Top Curr Chem* 247: 7–33, 2005.
34. Engel J, Chen HT, Prockop DJ, and Klump H. Triple helix (coil conversion of collagen-like polytripeptides in aqueous and nonaqueous solvents—Comparison of thermodynamic parameters and binding of water to $(L\text{-}Pro\text{-}L\text{-}Pro\text{-}Gly)_n$ and $(L\text{-}Pro\text{-}L\text{-}Hyp\text{-}Gly)_n$. *Biopolymers* 16: 601–622, 1977.
35. Engel J and Prockop DJ. The zipper-like folding of collagen triple helices and the effects of mutations that disrupt the zipper. *Annu Rev Biophys Biophys Chem* 20: 137–152, 1991.
36. Feng YB, Melacini G, Taulane JP, and Goodman M. Acetyl-terminated and template-assembled collagen-based polypeptides composed of Gly-Pro-Hyp sequences. 2. Synthesis and conformational analysis by circular dichroism, ultraviolet absorbance, and optical rotation. *J Am Chem Soc* 118: 10351–10358, 1996.
37. Fessler LI, Timpl R, and Fessler JH. Assembly and processing of procollagen type-III in chick-embryo blood-vessels. *J Biol Chem* 256: 2531–2537, 1981.
38. Fields CG, Grab B, Lauer JL, and Fields GB. Purification and analysis of synthetic, triple-helical minicollagens by reversed-phase high-performance liquid-chromatography. *Anal Biochem* 231: 57–64, 1995.
39. Fields CG, Grab B, Lauer JL, Miles AJ, Yu YC, and Fields GB. Solid-phase synthesis of triple-helical collagen-model peptides. *Lett Pept Sci* 3: 3–16, 1996.
40. Fields CG, Lovdahl CM, Miles AJ, Hagen VLM, and Fields GB. Solid-phase synthesis and stability of triple-helical peptides incorporating native collagen sequences. *Biopolymers* 33: 1695–1707, 1993.

41. Fields GB and Prockop DJ. Perspectives on the synthesis and application of triple-helical, collagen-model peptides. *Biopolymers* 40: 345–357, 1996.
42. Fiori S, Saccà B, and Moroder L. Structural properties of a collagenous heterotrimer that mimics the collagenase cleavage site of collagen type I. *J Mol Biol* 319: 1235–1242, 2002.
43. Frank S, Boudko S, Mizuno K, Schulthess T, Engel J, and Bächinger HP. Collagen triple helix formation can be nucleated at either end. *J Biol Chem* 278: 7747–7750, 2003.
44. Frank S, Kammerer RA, Mechling D, Schulthess T, Landwehr R, Bann J, Guo Y, Lustig A, Bächinger HP, and Engel J. Stabilization of short collagen-like triple helices by protein engineering. *J Mol Biol* 308: 1081–1089, 2001.
45. Gauba V and Hartgerink JD. Self-assembled heterotrimeric collagen triple helices directed through electrostatic interactions. *J Am Chem Soc* 129: 2683–2690, 2007.
46. Glanville RW, Allmann H, and Fietzek PP. Cysteine residues mark C-terminal triple-helical-non-triple-helical junction in type III collagen. *Hoppe-Seylers Z Physiol Chem* 357: 1663–1665, 1976.
47. Goodman M, Bhumralkar M, Jefferson EA, Kwak J, and Locardi E. Collagen mimetics. *Biopolymers* 47: 127–142, 1998.
48. Goodman M, Feng YB, Melacini G, and Taulane JP. A template-induced incipient collagen-like triple-helical structure. *J Am Chem Soc* 118: 5156–5157, 1996.
49. Goodman M, Melacini G, and Feng YB. Collagen-like triple helices incorporating peptoid residues. *J Am Chem Soc* 118: 10928–10929, 1996.
50. Grab B, Miles AJ, Furcht LT, and Fields GB. Promotion of fibroblast adhesion by triple-helical peptide models of type I collagen-derived sequences. *J Biol Chem* 271: 12234–12240, 1996.
51. Greiche Y and Heidemann E. Collagen model peptides with antiparallel structure. *Biopolymers* 18: 2359–2361, 1979.
52. Harrington WF and Rao NV. Collagen structure in solution. I. Kinetics of helix regeneration in single-chain gelatins. *Biochemistry* 9: 3714–3724, 1970.
53. Heidemann E and Roth W. Synthesis and investigation of collagen model peptides. *Adv Polym Sci* 43: 143–203, 1982.
54. Hojo H, Akamatsu Y, Yamauchi K, Kinoshita M, Miki S, and Nakamura Y. Synthesis and structural characterization of triple-helical peptides which mimic the ligand binding site of the human macrophage scavenger receptor. *Tetrahedron* 53: 14263–14274, 1997.
55. Holmes TC. Novel peptide-based biomaterial scaffolds for tissue engineering. *Trends Biotech* 20: 16–21, 2002.
56. Horng JC, Hawk AJ, Zhao Q, Benedict ES, Burke SD, and Raines RT. Macrocyclic scaffold for the collagen triple helix. *Org Lett* 8: 4735–4738, 2006.
57. Khoshnoodi J, Cartailler JP, Alvares K, Veis A, and Hudson BG. Molecular recognition in the assembly of collagens—Terminal noncollagenous domains are key recognition modules in the formation of triple helical protomers. *J Biol Chem* 281: 38117–38121, 2006.
58. Kielthy CM and Grant ME. The collagen family. Structure assembly and organisation in the extracellular matrix. In: *Connective Tissue and its Heritable Disorders, Molecular, Genetic and Medical Aspects*, 2nd edition, edited by Royce PM and Steinmann B. New York: Wiley-Liss; 2002, pp. 159–222.
59. Kinberger GA, Cai WB, and Goodman M. Collagen mimetic dendrimers. *J Am Chem Soc* 124: 15162–15163, 2002.
60. Koide T. Triple helical collagen-like peptides. Engineering and applications in matrix biology. *Connect Tiss Res* 46: 131–141, 2005.
61. Koide T, Asada S, Takahara Y, Nishikawa Y, Nagata K, and Kitagawa K. Specific recognition of the collagen triple helix by chaperone HSP47. Minimal structural requirement and spatial molecular orientation. *J Biol Chem* 281: 3432–3438, 2006.
62. Koide T, Homma DL, Asada S, and Kitagawa K. Self-complementary peptides for the formation of collagen-like triple helical supramolecules. *Bioorg Med Chem Lett* 15: 5230–5233, 2005.
63. Koide T and Nagata K. Collagen biosynthesis. *Top Curr Chem* 247: 85–114, 2005.
64. Koide T, Nishikawa Y, Asada S, Yamazaki CM, Takahara Y, Homma DL, Otaka A, Ohtani K, Wakamiya N, Nagata K, and Kitagawa K. Specific recognition of the collagen triple helix by chaperone HSP47. II. The HSP47-binding structural motif in collagens and related proteins. *J Biol Chem* 281: 11177–11185, 2006.
65. Koide T, Nishikawa Y, and Takahara Y. Synthesis of heterotrimeric collagen models containing Arg residues in Y-positions and analysis of their conformational stability. *Bioorg Med Chem Lett* 14: 125–128, 2004.
66. Koide T, Yuguchi M, Kawakita M, and Konno H. Metal-assisted stabilization and probing of collagenous triple helices. *J Am Chem Soc* 124: 9388–9389, 2002.
67. Koivu J and Myllylä R. Interchain disulfide bond formation in type-I and type-II procollagen—Evidence for a protein disulfide isomerase catalyzing bond formation. *J Biol Chem* 262: 6159–6164, 1987.
68. Kotch FW and Raines RT. Self-assembly of synthetic collagen triple helices. *Proc Natl Acad Sci USA* 103: 3028–3033, 2006.
69. Kwak J, De Capua A, Locardi E, and Goodman M. TREN (tris(2-aminoethyl)amine)—An effective scaffold for the assembly of triple helical collagen mimetic structures. *J Am Chem Soc* 124: 14085–14091, 2002.
70. Labourdette L and van der Rest M. Analysis of the role of the Col1 domain and its adjacent cysteine-containing sequence in the chain assembly of type-IX collagen. *FEBS Lett* 320: 211–214, 1993.
71. Lesage A, Penin F, Geourjon C, Marion D, van der Rest M. Trimeric assembly and three-dimensional structure model of the FACIT collagen COL1-NC1 junction from CD and NMR analysis. *Biochemistry* 35: 9647–9660, 1996.
72. Letarov AV, Londer YY, Boudko SP, and Mesyanzhinov VV. The carboxy-terminal domain initiates trimerization of bacteriophage T4 fibritin. *Biochem (Mosc)* 64: 817–823, 1999.
73. Lunstrum GP, McDonough AM, Marinkovich MP, Keene DR, Morris NP, and Burgeson RE. Identification and partial-purification of a large, variant form of type-XII collagen. *J Biol Chem* 267: 20087–20092, 1992.
74. Mann K, Mechling DE, Bächinger HP, Eckerskorn C, Gaill F, and Timpl R. Glycosylated threonine but not 4-hydroxyproline dominates the triple helix stabilizing positions in the sequence of a hydrothermal vent worm cuticle collagen. *J Mol Biol* 261: 255–266, 1996.
75. Mazzorana M, Cogne S, Goldschmidt D, and Aubert–Foucher E. Collagenous sequence governs the trimeric assembly of collagen XII. *J Biol Chem* 276: 27989–27998, 2001.
76. Mazzorana M, Giry–Lozinguez C, and van der Rest M. Trimeric assembly of collagen-XII. Effect of deletion of the C-terminal part of the molecule. *Matrix Biol* 14: 583–588, 1995.
77. Mazzorana M, Gruffat H, Sergeant A, and van der Rest M. Mechanisms of collagen trimer formation—Construction and expression of a recombinant minigene in HeLa cells reveals a direct effect of prolyl hydroxylation on chain assembly of type-XII collagen. *J Biol Chem* 268: 3029–3032, 1993.
78. Mazzorana M, Snellman A, Kivirikko KI, van der Rest M, and Pihlajaniemi T: Involvement of prolyl-4-hydroxylase in the assembly of trimeric minicollagen XII—Study in a baculovirus expression system. *J Biol Chem* 271: 29003–29008, 1996.
79. McAlinden A, Crouch EC, Bann JG, Zhang PN, and Sandell LJ. Trimerization of the amino propeptide of type IIA procollagen using a 14-amino acid sequence derived from the coiled-coil neck domain of surfactant protein D. *J Biol Chem* 277: 41274–41281, 2002.
80. Mechling DE and Bächinger HP. The collagen-like peptide $(GER)_{15}GPCCG$ forms pH-dependent covalently linked triple helical trimers. *J Biol Chem* 275: 14532–14536, 2000.
81. Mechling DE, Gambee JE, Morris NP, Sakai LY, Keene DR, Mayne R, and Bächinger HP. Type IX collagen NC1 domain peptides can trimerize *in vitro* without forming a triple helix. *J Biol Chem* 271: 13781–13785, 1996.
82. Moroder L, Fiori S, Friedrich R, and Ottl J. Collagen mimics. Synthesis and properties of disulfide-bridged trimeric collagen peptides. In: *Synthetic Macromolecules with Higher Structural*

Order, edited by Khan IM. ACS Symposium Series, vol. 812. Washington, DC: ACS, 2002, pp. 103–116.
83. Moroder L, Musiol HJ, Götz M, and Renner C. Synthesis of single- and multiple-stranded cystine-rich peptides. *Biopolymers* 80: 85–97, 2005.
84. Müller JCD, Ottl J, and Moroder L. Heterotrimeric collagen peptides as fluorogenic collagenase substrates—Synthesis, conformational properties, and enzymatic digestion. *Biochemistry* 39: 5111–5116, 2000.
85. Myllyharju J and Kivirikko KI. Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet* 20: 33–43, 2004.
86. Myllylä R, Kuuttisavolainen ER, and Kivirikko KI. Role of ascorbate in prolyl hydroxylase reaction. *Biochem Biophys Res Comm* 83: 441–448, 1978.
87. Ninomiya Y and Olsen BR. Synthesis and characterization of cDNA encoding a cartilage-specific short collagen. *Proc Natl Acad Sci USA* 81: 3014–3018, 1984.
88. Olsen BR. Collagen biosynthesis. In: *Cell Biology of the Extracellular Matrix*, edited by Hay ED. New York: Plenum Press, 1991, pp. 177–220.
89. Ottl J, Battistuta R, Pieper M, Tschesche H, Bode W, Kühn K, and Moroder L. Design and synthesis of heterotrimeric collagen peptides with a built-in cystine-knot. Models for collagen catabolism by matrix-metalloproteases. *FEBS Lett* 398: 31–36, 1996.
90. Ottl J, Gabriel D, Murphy G, Knauper V, Tominaga Y, Nagase H, Kröger M, Tschesche H, Bode W, and Moroder L. Recognition and catabolism of synthetic heterotrimeric collagen peptides by matrix metalloproteinases. *Chem Biol* 7: 119–132, 2000.
91. Ottl J and Moroder L. Disulfide-bridged heterotrimeric collagen peptides containing the collagenase cleavage site of collagen type I. Synthesis and conformational properties. *J Am Chem Soc* 121: 653–661, 1999.
92. Park C and Raines RT. Adjacent cysteine residues as a redox switch. *Protein Engineer* 14: 939–942, 2001.
93. Persikov AV, Ramshaw JAM, Kirkpatrick A, and Brodsky B. Amino acid propensities for the collagen triple-helix. *Biochemistry* 39: 14960–14967, 2000.
94. Persikov AV, Ramshaw JAM, Kirkpatrick A, and Brodsky B. Peptide investigations of pairwise interactions in the collagen triple-helix. *J Mol Biol* 316: 385–394, 2002.
95. Privalov PL. Stability of proteins. Proteins which do not present a single cooperative system. *Adv Prot Chem* 35: 1–104, 1982.
96. Prockop DJ and Kivirikko KI. Collagens. Molecular biology, diseases, and potentials for therapy. *Annu Rev Biochem* 64: 403–434, 1995.
97. Roth W and Heidemann E. Triple helix-coil transition of covalently bridged collagen-like peptides. *Biopolymers* 19: 1909–1917, 1980.
98. Roth W, Heppenheimer K, and Heidemann ER. Structure of collagen-like homopolytripeptides and heteropolytripeptides. IV. Polytripeptides by repetitive peptide-synthesis and bridging of oligopeptides. *Macromol Chem Phys* 180: 905–917, 1979.
99. Rump ET, Rijkers DTS, Hilbers HW, de Groot PG, and Liskamp RMJ. Cyclotriveratrylene (CTV) as a new chiral triacid scaffold capable of inducing triple helix formation of collagen peptides containing either a native sequence or Pro-Hyp-Gly repeats. *Chem Eur J* 8: 4613–4621, 2002.
100. Saccà B, Barth D, Musiol HJ, and Moroder L. Conformation-dependent side reactions in interstrand-disulfide bridging of trimeric collagenous peptides by regioselective cysteine chemistry. *J Pept Sci* 8: 205–210, 2002.
101. Saccà B, Fiori S, and Moroder L. Studies of the local conformational properties of the cell-adhesion domain of collagen type IV in synthetic heterotrimeric peptides. *Biochemistry* 42: 3429–3436, 2003.
102. Saccà B and Moroder L. Synthesis of heterotrimeric collagen peptides containing the a1b1 integrin recognition site of collagen type IV. *J Pept Sci* 8: 192–204, 2002.
103. Saccà B, Renner C, and Moroder L. The chain register in heterotrimeric collagen peptides affects triple helix stability and folding kinetics. *J Mol Biol* 324: 309–318, 2002.
104. Saccà B, Sinner EK, Kaiser J, Lübken C, Eble JA, and Moroder L. Binding and docking of synthetic heterotrimeric collagen type IV peptides with a1b1 integrin. *ChemBioChem* 3: 904–907, 2002.
105. Sakakibara S, Inouye K, Shudo K, Kishida Y, Kobayashi Y, and Prockop DJ. Synthesis of $(Pro\text{-}Hyp\text{-}Gly)_n$ of defined molecular weights. Evidence for stabilization of collagen triple helix by hydroxyproline. *Biochim Biophys Acta* 303: 198–202, 1973.
106. Slatter DA, Foley LA, Peachey AR, Nietlispach D, and Farndale RW. Rapid synthesis of a register-specific heterotrimeric type I collagen helix encompassing the integrin a2b1 binding site. *J Mol Biol* 359: 289–298, 2006.
107. Snellman A, Keränen MR, Hägg PO, Lamberg A, Hiltunen JK, Kivirikko KI, and Pihlajaniemi T. Type XIII collagen forms homotrimers with three triple helical collagenous domains and its association into disulfide-bonded trimers is enhanced by prolyl 4-hydroxylase. *J Biol Chem* 275: 8936–8944, 2000.
108. Snellman A, Tu HM, Väisänen T, Kvist AP, Huhtala P, and Pihlajaniemi T. A short sequence in the N-terminal region is required for the trimerization of type XIII collagen and is conserved in other collagenous transmembrane proteins. *EMBO J* 19: 5051–5059, 2000.
109. Stetefeld J, Frank S, Jenny M, Schulthess T, Kammerer RA, Boudko S, Landwehr R, Okuyama K, and Engel J. Collagen stabilization at atomic level. Crystal structure of designed $(GlyPro\text{-}Pro)_{10}$ foldon. *Structure* 11: 339–346, 2003.
110. Tanaka T, Wada Y, Nakamura H, Doi T, Imanishi T, and Kodama T. A synthetic model of collagen structure taken from bovine macrophage scavenger receptor. *FEBS Lett* 334: 272–276, 1993.
111. Tao YZ, Strelkov SV, Mesyanzhinov VV, and Rossmann MG. Structure of bacteriophage T4 fibritin—A segmented coiled coil and the role of the C-terminal domain. *Structure* 5: 789–798, 1997.
112. Wang XH, Connor M, Smith R, Maciejewski MW, Howden MEH, Nicholson GM, Christie MJ, and King GF. Discovery and characterization of a family of insecticidal neurotoxins with a rare vicinal disulfide bridge. *Nature Struct Biol* 7: 505–513, 2000.
113. Welker E, Wedemeyer WJ, Narayan M, and Scheraga HA. Coupling of conformational folding and disulfide-bond reactions in oxidative folding of proteins. *Biochemistry* 40: 9059–9064, 2001.
114. Woolfson DN and Ryadnov MG. Peptide-based fibrous biomaterials—Some things old, new and borrowed. *Curr Opin Chem Biol* 10: 559–567, 2006.
115. Xia ZX, He YN, Dai WW, White SA, Boyd GD, and Mathews FS. Detailed active site configuration of a new crystal form of methanol dehydrogenase from *Methylophilus W3A1* at 1.9 Ä resolution. *Biochemistry* 38: 1214–1220, 1999.
116. Yang W, Chan VC, Kirkpatrick A, Ramshaw JAM, and Brodsky B. Gly-Pro-Arg confers stability similar to Gly-Pro-Hyp in the collagen triple-helix of host-guest peptides. *J Biol Chem* 272: 28837–28840, 1997.

Address reprint requests to:
Luis Moroder
Max-Planck-Institut für Biochemie
Am Klopferspitz 18
D-82152 Martinsried, Germany

E-mail: moroder@biochem.mpg.de

Date of first submission to ARS Central, August 2, 2007; date of acceptance, August 9, 2007.

This article has been cited by:

1. Gregg B. Fields. 2010. Synthesis and biological applications of collagen-model triple-helical peptides. *Organic & Biomolecular Chemistry* **8**:6, 1237. [CrossRef]
2. Sergei P. Boudko, Takako Sasaki, Jürgen Engel, Thomas F. Lerch, Jay Nix, Michael S. Chapman, Hans Peter Bächinger. 2009. Crystal Structure of Human Collagen XVIII Trimerization Domain: A Novel Collagen Trimerization Fold. *Journal of Molecular Biology* **392**:3, 787-802. [CrossRef]
3. Rita Berisio, Alfonso De Simone, Alessia Ruggiero, Roberto Improta, Luigi Vitagliano. 2009. Role of side chains in collagen triple helix stabilization and partner recognition. *Journal of Peptide Science* **15**:3, 131-140. [CrossRef]
4. Salvador Ventura . 2008. Oxidative Protein Folding: From the Test Tube to In Vivo Insights. *Antioxidants & Redox Signaling* **10**:1, 51-54. [Citation] [PDF] [PDF Plus]