

Covalent Capture of Collagen Triple Helices Using Lysine–Aspartate and Lysine–Glutamate Pairs

Sarah A. H. Hulgán, Abhishek A. Jalan, I-Che Li, Douglas R. Walker, Mitchell D. Miller, Abigail J. Kosgei, Weijun Xu, George N. Phillips, Jr., and Jeffrey D. Hartgerink*



Cite This: *Biomacromolecules* 2020, 21, 3772–3781



Read Online

ACCESS |



Metrics & More

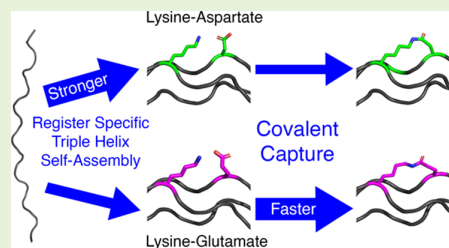


Article Recommendations



Supporting Information

ABSTRACT: Collagen mimetic peptides (CMPs) self-assemble into a triple helix reproducing the most fundamental aspect of the collagen structural hierarchy. They are therefore important for both further understanding this complex family of proteins and use in a wide range of biomaterials and biomedical applications. CMP self-assembly is complicated by a number of factors which limit the use of CMPs including their slow rate of folding, relatively poor monomer–trimer equilibrium, and the large number of competing species possible in heterotrimeric helices. All of these problems can be solved through the formation of isopeptide bonds between lysine and either aspartate or glutamate. These amino acids serve two purposes: they first direct self-assembly, allowing for composition and register control within the triple helix, and subsequently can be covalently linked, fixing the composition and register of the assembled structure without perturbing the triple helical conformation. This self-assembly and covalent capture are demonstrated here with four different triple helices. The formation of an isopeptide bond between lysine and glutamate (K–E) is shown to be a faster and higher yielding reaction than lysine with aspartate (K–D). Additionally, K–E amide bonds increase the thermal stability, improve the refolding capabilities, and enhance the triple helical structure as compared to K–D supramolecular interactions, observed by circular dichroism. In contrast, covalent capture of triple helices with K–D amide bonds occurs slower, and the captured triple helices do not have enhanced helical structure. The crystal structure of a triple helix captured through the formation of three K–E isopeptide bonds unequivocally demonstrates the connectivity of the amide bonds formed while also confirming the preservation of the canonical triple helix. The rate of reaction and yield for covalently captured K–E triple helices along with the excellent preservation of triple helical structure demonstrate that this approach can be used to effectively capture and stabilize this important biological motif for biological and biomedical applications.



INTRODUCTION

Self-assembly is a powerful method for creating complex ensembles of molecules with both fascinating structures and important applications.¹ However, the use of weak, non-covalent interactions in the assembly process indicates that the formed structures may have limited stability. Challenges associated with stability lead to at least two major problems. First, to borrow a term from traditional organic synthesis, methods of self-assembly are usually limited to “one-step syntheses” as subsequent steps of self-assembly require conditions that are not compatible with the stability of the initial structure. Thus, while self-assembly shows great promise as a synthetic method, more sophisticated multistep approaches are out of reach. Second, many applications, particularly those which require use at low concentrations, are not possible as the noncovalent structure disassembles under the necessary conditions. One method for overcoming both of these difficulties is a process of self-assembly, followed by covalent capture.^{2,3} This technique uses self-assembly to control the structure formation while subsequent covalent bond formation stabilizes that structure. Success requires both

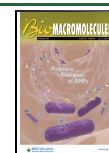
(1) reaction conditions which do not disrupt the noncovalent assembly and (2) covalent bonds that, when formed, do not significantly disrupt the desired structure. Thus, often systems designed exclusively for supramolecular assembly are not amenable for subsequent covalent capture. In this work, we show how collagen mimetic peptides (CMPs) can be self-assembled and subsequently covalently captured.

Collagens are the most abundant proteins by mass in the human body and play an important role in human health and disease.⁴ Although most collagens are well known as extracellular matrix proteins, there are many collagen-like proteins present in viruses, bacteria, and fungi as well as in the human innate immune system.^{5–7} It is therefore important to investigate the structures and functions of collagens, such as

Received: June 8, 2020

Revised: July 27, 2020

Published: August 21, 2020



ACS Publications

© 2020 American Chemical Society

3772

<https://dx.doi.org/10.1021/acs.biomac.0c00878>
Biomacromolecules 2020, 21, 3772–3781

protein binding and cellular interactions. However, because of the large size, frequent cross-links, and poor solubility of natural collagens, they are difficult to purify and study.⁸ Therefore, it is common practice to use CMPs to help understand or to mimic aspects of natural collagens' structures and functions. A collagen is characterized by a triple helix composed of three protein strands each in a left-handed polyproline type II (PPII) helix which together form a right-handed supercoil. Triple helices can be homotrimers, with three identical protein strands, or heterotrimers, with two or three different strands. Each of these protein strands contains the triplet motif (Xaa-Yaa-Gly), where Xaa and Yaa can be any amino acid but are frequently proline (Pro, P) and 4-(R)-hydroxyproline (Hyp, O), respectively.^{9,10} Glycine is present at every third amino acid of each peptide's sequence as this position's side chain points toward the center of the triple helix and anything larger than glycine's hydrogen would sterically disrupt the interstrand hydrogen bonds which stabilize the helix.^{11–14} When the three strands assemble into the trimer, glycine must be present at each cross section. To achieve this, each strand must be offset by a single amino acid differentiating a leading, middle, and trailing peptide strand. Because of this offset, the register determines the relative three-dimensional presentation of groups along the triple helical structure.^{15,16}

When designing heterotrimeric helices, the desired species must be stabilized above competing compositions and registers. The mixture of two peptides, A and B, for example, could fold into homotrimers, AAA or BBB, or heterotrimers, A₂B or AB₂. These heterotrimers additionally could fold into alternative registers: AAB, ABA, BAA, ABB, BAB, and BBA. Therefore, this mixture can result in eight structurally unique triple helices. However, charge–pair interactions have been used by our group, and others, to stabilize designed heterotrimers of a single composition and register.^{15–26} Xaa and Yaa amino acids may be substituted to form stabilizing interstrand charge–pair interactions between lysine and either glutamate or aspartate. As these amino acids are quite destabilizing to a triple helix when not charge-paired, this allows for both positive design of stabilizing the desired composition and register in addition to simultaneous negative design of destabilizing competing systems. In collagen, two charge–pair geometries have been identified: “axial” and “lateral”. An axial interaction is from the Yaa position of one strand to an Xaa position of the adjacent strand one triplet removed in the C-terminal direction (Figure 1a).^{27,28} In contrast, a lateral interaction can form between the Yaa position of one strand to an Xaa position of the adjacent strand in the same triplet. In lysine–aspartate (K–D) and lysine–glutamate (K–E) pairs, the axial geometry is substantially

more favorable, while lateral interactions provide only minimal benefits.^{27,29} However, this axial geometry is not preferred for all amino acid pairs. For example, experiments using unnatural amino acids have shown potential for lateral interaction stabilization of the triple helix.²⁶

Although charge–pair interactions have been helpful in stabilizing designed helices, the use of CMPs is still limited because of the concentration-dependent equilibrium between the triple helix and the monomeric peptides. This is particularly problematic in situations which require use at low concentrations, such as in vivo applications. This difficulty is worsened by the exceptionally slow folding kinetics of CMPs compared to globular proteins. For example, CMP homotrimers may take hours to fold while heterotrimers can take days or weeks to fold.^{30–32}

Previously, covalent techniques have been used to tether the three strands of the triple helix together, increasing the effective concentration of peptide and thus improving both the rate of folding and the equilibrium constant of triple helix formation. One such strategy uses a small molecule containing either three carboxylic acids or three amine groups to react with either the N- or C-termini of peptides.^{33–41} Another approach is to use the amino acid side chains themselves to ligate the three peptides together, such as cysteine knot formation where oxidation leads to homotrimer formation or a complex protecting group scheme is employed for a heterotrimer synthesis.^{42–48} Another technique is to use a di-lysine strategy during solid-phase peptide synthesis. In this method, a C-terminal sequence of Lys–Lys allows for branching during peptide synthesis and results in a stabilized homotrimer.^{49–54} Finally, other complex reaction schemes have been employed for heterotrimer production such as oxime ligation and click chemistry.^{55,56} Recently, homotrimers bearing hydrophobic moieties were used as a noncovalent stabilization method.⁵⁷ Though all these techniques improve the folding rate and equilibrium constant of species, they are often cumbersome synthetic routes, most are only useful for homotrimers, and none have been proved to account for the register of the helix formed via structural characterization.

An ideal synthetic system would (1) allow for stabilization of trimer, (2) leave the triple helical conformation intact, and (3) control heterotrimer registration. Previously, we demonstrated the covalent capture of a heterotrimeric helix using lysine–aspartate (K–D) axial isopeptide bond formation.⁵⁸ Subsequently, we employed this covalent capture technique to stabilize a short type I collagen sequence in varying registers.⁵⁹ It is important to note that in our system, we identified that the register is indeed AAB as intended,⁵⁸ and also if our cross-links are made, then only one possible register can be adopted while still maintaining triple helical conformation, an important distinction compared to previous covalent techniques. However, isopeptide bond formation is known in natural proteins between Lys and Asx or Glx.^{60,61} Therefore, here we expand and improve this covalent capture technique with the use of lysine–glutamate (K–E) isopeptide bond formation. In this work, we show K–D bond formation effectivity, yet observe that the reaction occurs slowly and is unable to be driven to completion. We hypothesize that the required contraction in distance between charge-paired K–D side chains and covalently bonded K–D side chains may induce strain, leading to slow reaction rates and low yield (Figure 1). To overcome these challenges, we subsequently use K–E pairs for covalent capture. The extra methylene group in the

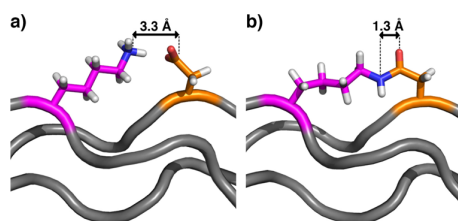


Figure 1. Lysine–aspartate (a) axial charge–pair interaction and (b) isopeptide bond highlighting the contraction of distance between carbonyl carbon and nitrogen before and after amide bond formation.

glutamic acid side chain provides additional flexibility and length, permitting more effective isopeptide bond formation. By comparing the use of K–D and K–E bond formation in four designed triple helices, we argue the superiority of the K–E usage via enhanced reaction rates and yields, increased triple helical character, and improved stability. Additionally, we demonstrate by X-ray crystal structure analysis the isopeptide bond connectivity and show that the triple helix backbone is not disrupted by the formation of this unusual bond. Through these comparisons, we observe supramolecular and covalent capture design criteria differences. Together, this work demonstrates that triple helix self-assembly directed by charge–pair interactions and subsequent covalent capture to form isopeptide bonds is a powerful method for CMP design and application.

■ EXPERIMENTAL DETAILS

Materials. Fmoc-protected amino acids and resin were purchased from EMD Chemicals, and 2-(1*H*-7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyl uranium hexafluorophosphate methanaminium (HATU) was purchased from P3Bio. All chemicals not otherwise specified were purchased from Sigma-Aldrich.

Peptide Synthesis. Peptides were synthesized using standard Fmoc-protected amino acids using a low loading rink amide MBHA resin to give C-terminal amidation. A mixture of 25% v/v piperidine in dimethylformamide (DMF) was used for deprotecting steps. Coupling was performed using HATU and diisopropylethylamine (DiEA) in DMF in the ratio of 1:4:4:6 (resin/amino acid/HATU/DiEA respectively). Acetylation of the N-terminus was performed twice with an excess of acetic anhydride and DiEA in dichloromethane. Cleavage was performed with 10% v/v scavengers (triisopropylsilane, H₂O, and ethanedithiol) in trifluoroacetic acid (TFA). TFA was removed from the reaction mixture by evaporation under nitrogen. Cold diethyl ether was used to triturate the crude peptide. After centrifugation, the crude pellet was washed with cold diethyl ether two more times.

Peptide Purification. Peptides were dissolved in H₂O to a concentration of 22 mg/mL. This was sonicated and then filtered before purification by reverse-phase high-pressure liquid chromatography (HPLC) with water and acetonitrile with 0.05% TFA at a gradient of 0.7% per minute on a Waters Atlantis T3 column. Samples were roto-evaporated to remove acetonitrile and then lyophilized. Matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-ToF MS) (Bruker Instruments) was used to confirm the peptide mass.

Sample Preparation. Peptide samples were dissolved in 100 mM aqueous 2-(*N*-morpholino)ethanesulfonic acid (MES) buffer pH 6.1 to a concentration of 3 mM peptide. The concentrations of KGE and KGD peptides were determined by mass; for C and D peptides, the concentration was determined by UV–vis spectroscopy using a Thermo Scientific Nanodrop 2000c. The heterotrimer sample was made by mixing C and D peptides in a 2:1 ratio. All samples were preheated at 85 °C for 15 min, cooled to room temperature, and then stored at 5 °C for 1 week to ensure complete folding before characterization or reaction.

Circular Dichroism Spectroscopy. Circular dichroism (CD) spectroscopy was performed on a Jasco J-810 spectropolarimeter equipped with a Peltier temperature-controlled stage. The samples were diluted with Milli-Q water to 0.03 mM for spectra and 0.3 mM for melts immediately before running on a CD spectrometer. The samples (200 μ L) were transferred to a quartz cuvette of 0.1 cm path length. Wavelength scans were performed between 180 and 250 nm. The maximum which falls near 225 nm is then followed as the temperature is increased from 5 to 85 °C at 10 °C/h. After holding for 15 min at 85 °C, refolding curves for heterotrimers were obtained by cooling from 85 to 5 °C at a rate of 10 °C/h. Using the Savitzky–Golay smoothing algorithm, the first derivative curve was calculated for melting and recovery curves. The minimum of the first derivative

is defined as the melting temperature (T_m). The molar residue ellipticity (MRE) was calculated as previously reported.⁵⁸

Homotrimer fold recovery was followed by first heating the 0.3 mM sample in an 85 °C hot water bath for 15 min. The samples were then transferred to the CD sample chamber already set to 5 °C. Immediately after transferring, ellipticity at 225 nm was measured every 20 s until the fold was recovered. Fraction folded was calculated using each samples' ellipticity at 5 °C as $\theta_{\max}$ (fully folded) and the ellipticity at 85 °C as $\theta_{\min}$ (fully unfolded) where

$$\text{Fraction folded } (\alpha) = \frac{\theta - \theta_{\min}}{\theta_{\max} - \theta_{\min}}$$

Size Exclusion Chromatography. The 3 mM peptide samples (in 100 mM MES buffer pH 6.1, as previously described) were diluted to 1 mM peptide in 100 mM MES pH 6.1 to be analyzed on a Superdex™ 30 Increase size exclusion column 10/300 GL (GE Healthcare, Chicago, IL). Elution was carried out in 100 mM MES buffer pH 6.1. Sodium azide was added to a final concentration of 0.02% (w/v) in MES to prevent microbial contamination. The flow rate was 0.8 mL/min with the elution pattern, followed by UV absorption at 220 nm for KGE and KGD peptides and at 274 nm for CCD (tyrosine absorption).

Covalent Capture Reaction by EDC Coupling. Hydroxybenzotriazole (36 mM, HOBt) and 600 mM 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) solutions were prepared in 100 mM MES buffer pH 6.1. The 3 mM peptide solutions (in 100 mM MES buffer pH 6.1, as previously described) were mixed with the EDC and HOBt solutions in a 1:40:4 ratio (isopeptide bond/EDC/HOBt) in the MES buffer.⁵⁸ Thus, the final reaction mixture for KGE and KGD contained 1 mM peptide (1 mM isopeptide bonds), 40 mM EDC, and 4 mM HOBt in 100 mM MES, while the CCD reaction mixture contained 1 mM peptide (3 mM isopeptide bonds), 120 mM EDC, and 12 mM HOBt in 100 mM MES buffer pH 6.1. The reaction was mixed well by vortex and incubated at 5 °C. For the first 2 h, the solutions were mixed by vortex every half hour to ensure homogeneity. For preparing and purifying the KGDcc, additional activating agents (of the same concentration and volume) were added at day 4 to try increasing the yield of the trimer product. The reaction was quenched by addition of 1 M hydroxylamine in a 1:1 ratio v/v (reaction mixture/hydroxylamine), vortexed, and then incubated at room temperature for 24 h. Prior to characterization of the sample by MALDI or size exclusion chromatography (SEC), the same volume of 1 M HCl (as hydroxylamine) was added to neutralize pH prior to characterization or purification.

Other ratios of isopeptide bond/EDC/HOBt (all in MES buffer) were also tested on KGE and KGD samples. Increasing quantities of both EDC and HOBt were tested as 1:1:1, 1:2:2, 1:10:2, 1:20:2, 1:40:4, 1:80:8, and 1:120:12. 1:40:4 ratio was found to be optimal in this range. Additionally, it was tested to lower one or the other reagent to test the necessity of both using ratios of 1:40:1 and 1:2:8. The 1:40:4 ratio (isopeptide bond/EDC/HOBt) previously used⁵⁸ was still found to be optimal.

Purification and Preparation of Covalently Captured Triple Helices. The CCDcc and KGDcc samples were purified by SEC by the same method detailed above. The CCDcc was subsequently washed with MES buffer and concentrated using Pierce Protein Concentrator PES with 3K molecular weight cutoff. Thus, a final sample was obtained of 3 mM peptide (1 mM CCDcc trimer) in 100 mM MES pH 6.1 with 10% MeOH (added to increase solubility). Concentration was calculated by UV–vis absorption on Thermo Scientific Nanodrop 2000c. KGDcc was subsequently dialyzed to remove excess salts, followed by lyophilization. The peptide concentration was then calculated by mass and adjusted based on the SEC peak area to approximately 3 mM peptide in 100 mM MES pH 6.1. KGEcc samples were purified by HPLC, followed by roto-evaporation of the acetonitrile and lyophilization. The peptide concentration was calculated by mass to a final concentration of 3 mM peptide in 100 mM MES pH 6.1.

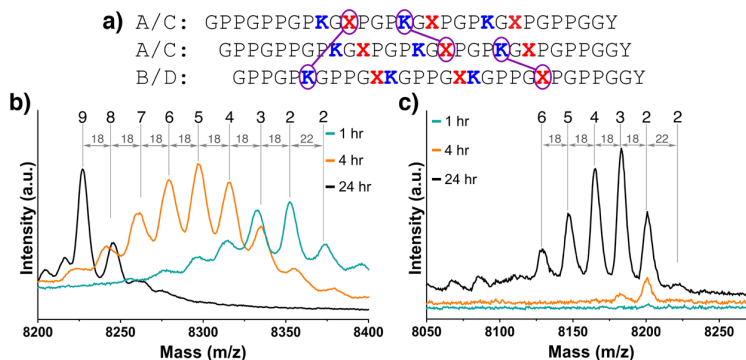


Figure 2. Comparison of K–E and K–D covalent capture. (a) Sequence of designed peptides, where X represents aspartate in the case of A/B peptides and glutamate in the case of C/D peptides. Lassos highlight three representative axial interactions: One each from leading to middle, middle to trailing, and trailing to leading peptide strand. Peptides are N-terminally acetylated and C-terminally amidated. (b) MALDI-MS spectra of the trimer region for C₂D from 1 to 24 h. The number of K–E amide bonds formed increases from 2 (the minimum necessary for trimer) to 9 (the maximum that can be formed). (c) Trimer region spectra for A₂B from 1 to 24 h. The average number of K–D amide bonds present increases from 0 to only 3. Data for (c) were obtained by us, replotted from ref 58.

Crystallization and Data Collection. Purified KGEcc collagen peptide powder was dissolved in H₂O at a concentration of 8 mg/mL and then diluted to 5 mg/mL for further crystallization experiments. A Mosquito liquid handling robot (SPT Labtech, Melbourn, UK) was used to set up sitting drop crystallization experiments using standard crystallization screens including PEG-RX HT, index HT, SaltRx HT (Hampton Research, Aliso Viejo, CA) Wizard I, and II (Rigaku Reagents, Bainbridge Island, WA). Similar hits producing sea urchin-like clusters of needle-shaped crystals were obtained against conditions containing either 3.5 M sodium formate or 50% tacsimate (1.83 M malonic acid, 0.25 M ammonium citrate tribasic, 0.12 M succinic acid, 0.3 M DL-malic acid, 0.4 M sodium acetate trihydrate, 0.5 M sodium formate, and 0.16 M ammonium tartrate dibasic). These hits were further optimized using sitting drop and hanging drop vapor diffusion methods producing clusters of needles with spines of roughly 100–400 μ m long by 1–8 μ m in the small dimensions grew in a well condition of 3.5 M sodium formate, 0.1 M Bis-Tris, and pH 7.0 (Figure S22a). The peptide crystals were harvested under paraffin oil (Hampton Research) using Mitegen MicroLoop E and MicroMesh mounts (Ithaca, NY) and flash-cooled in liquid nitrogen (Figure S22b). Data were collected with a Dectris Eiger 16M detector (Baden-Daettwil, Switzerland) on GM/CA-CAT beamline 23ID-B at the Advanced Photon Source (APS, Argonne National Lab, IL) using a 10 μ m microbeam collimator and 12 keV X-rays.

The crystals diffracted to a resolution of 1.4 Å. The best data set was obtained by merging data from two needle-shaped crystals with short edges of roughly 2–4 μ m and lengths of 60–70 μ m. The diffraction data were integrated with *DIALS*⁶² and scaled using *aimless*.^{63,64} The diffraction data were indexed in the monoclinic space group C2 with unit cell dimensions of $a = 134.70$, $b = 14.10$, and $c = 24.54$, $\beta = 90.3$. Data reduction statistics are shown in Table S2.

Structure Determination and Refinement. The collagen peptide structure was solved by molecular replacement with Phaser⁶⁵ using the previously solved KGE peptide structure without covalent cross-linkages (PDB id: 3t4f)²⁷ as the search model. The initial model was autobuilt using ARP/wARP,⁶⁶ followed by refinement and rebuilding using phenix.refine⁶⁷ and coot.⁶⁸ Stereochemical restraints were validated against the QM ones generated by the program DivCon.⁶⁹ The *R*-factor and the *R*-free were 0.156 and 0.180, respectively. The previously determined structure of KGE (PDB id: 3t4f) triple helix was re-refined using the same restraints to facilitate comparison. The structures were viewed and analyzed using a collaborative three-dimensional visualization system.⁷⁰ Refinement statistics for the structures are shown in Table S3. Coordinates and structure factors for the KGEcc peptide have been deposited in the worldwide protein data bank (wwPDB)⁷¹ under id 6vzx along with

updated coordinates for the re-refined KGE triple helix (PDB id: 3t4f).

RESULTS AND DISCUSSION

Heterotrimer Comparison. We designed two heterotrimers, identical with the exception that one pairs lysine with aspartate and the other with glutamate. The nine axial charge pairs between lysine and either aspartate or glutamate stabilize the triple helices in the desired register while fewer interactions are formed for alternative registers and compositions, thus promoting register specificity.⁵⁸ To make these two triple helices, four peptides A, B, C, and D were prepared which self-assemble to form the triple helices AAB and CCD (Table S1, Figure 2a). The triple helix AAB (containing aspartates) was previously published and characterized, however, is here included for thorough comparison to new helix CCD (containing glutamates).⁵⁸ Peptides were synthesized by solid-phase peptide synthesis with N-terminal acetylation and C-terminal amidation to prevent charge repulsion at the termini. Peptides were purified by HPLC following the procedure described previously (peptides A and B containing aspartates) or in experimental methods (peptides C and D containing glutamates) (Figure S1).⁵⁸

Peptides C and D were mixed in a 2:1 molar ratio in order to form the desired heterotrimer. This mixture was confirmed to form triple helix via SEC and CD spectroscopy. SEC can be used to observe relative monomer to trimer ratios. First, the individual peptides C and D are seen to contain only monomer (Figure S3). By comparison, the mixture of the two peptides shows the presence of a trimer, indicating that only a combination of the two peptides results in a triple helix (Figure S4a). Additionally, CD can be used to observe the presence of the PPII backbone characteristic of collagen. At 5 °C, the C and D peptides and C₂D mixtures all show the characteristic PPII spectra maxima at 225 nm and minima at 190 nm; however, the C₂D mixture shows a larger MRE than the individual peptides alone, indicating that a stronger (more abundant) triple helix is formed (Figures S5 and S6). Additionally, temperature-dependent CD was used to study the relative stabilities of triple helices by following the 225 nm peak as the sample is slowly heated to 85 °C. The minima of the first derivative of this melting curve is then defined as the melting temperature (T_m). The single peptides C and D show

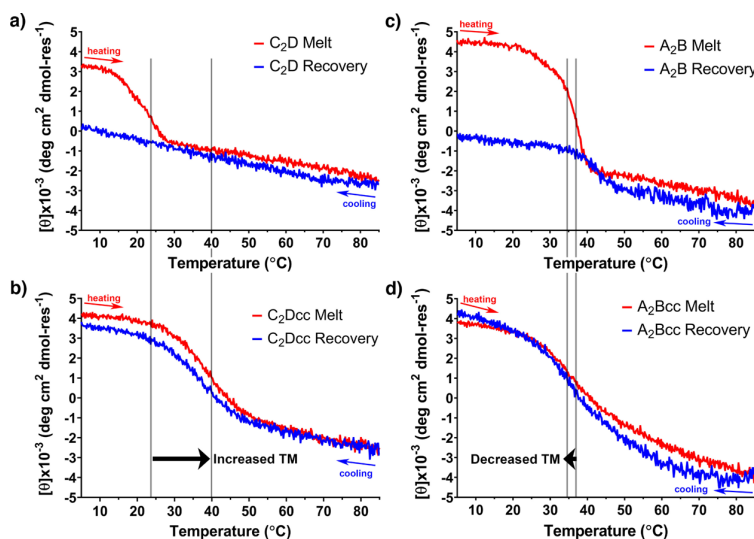


Figure 3. Thermal melting and refolding data for supramolecular and covalently captured triple helices. (a) C_2D melt and recovery. (b) C_2Dcc melt and recovery showing an improved melting temperature and recovery hysteresis over supramolecular C_2D . (c) A_2B melt and recovery curves. (d) A_2Bcc melt and recovery curves showing no melting temperature improvement over supramolecular A_2B . Red indicates melting curve with temperature increased over time; blue indicates recovery curve with temperature decreased over time. Data for (c,d) were obtained by us, replotted from ref 58.

low melting temperatures of 16.8 and 16.6 °C, respectively (Figure S5). The C_2D mixture shows that a more stable triple helix is formed than the single peptides alone with a T_m of 24.0 °C, again confirming the presence of the desired heterotrimer construct (Figure S6).

This C_2D assembly was then covalently captured using the carboxylate activating reagents EDC and HOBt. Briefly, the reaction mixture contained 1:40:4 isopeptide bond/EDC/HOBt. MALDI-MS was used to assess the rate and success of the reaction. The mass corresponding to the trimer shows that the triple helix was indeed covalently captured (Figure S2) because supramolecular self-assembled triple helices dissociate under the harsh conditions of MALDI-MS experiments and only appear as monomers. Furthermore, this trimer region can be studied to determine the extent of cross-links formed and compared to the previous A_2B system (Figure 2). Each amide bond formed corresponds to a loss of a water molecule, thus each peak spacing of 18 amu corresponds to one isopeptide bond formed. After only 1 h of reaction, the trimer peak is already present with two water molecules lost, indicating two isopeptide bonds formed, the minimum number needed for a covalent trimer. As the reaction time increases, so does the number of amide bonds formed. At 4 h, the major peak has five amide bonds. In theory, the CCD construct can form nine isopeptide bonds, and remarkably, after 24 h, the major species shows all of the nine possible K–E amide bonds. In contrast, no trimer is observed in the K–D heterotrimer spectra until after 4 h.⁵⁸ At 24 h, the major species still only contains three K–D bonds, much fewer than the K–E heterotrimer. Importantly, the species with nine amide bonds never become the major species for the A_2B system. Thus, the K–E covalent capture is shown to be faster and more efficient compared to K–D.

Upon purification by SEC, the covalently captured heterotrimer (C_2Dcc) can be characterized and compared to the supramolecular C_2D in order to ensure that K–E bond formation does not disrupt the intended structure. SEC can be

used to observe monomer to trimer ratios for both native and heated samples. Though the supramolecular C_2D shows trimer formation, as mentioned previously, there is still a large monomer population, indicating the poor equilibrium of folding that often plagues heterotrimeric CMP usage (Figure S4). Additionally, once C_2D is heated prior to SEC characterization, the sample falls apart to 100% monomer as expected for a supramolecularly stabilized sample. By comparison, covalent cross-links also allow for purification of only trimer species, so C_2Dcc is shown to be completely trimer natively and even upon heating (Figure S4), demonstrating the utility of covalently capturing the supramolecular structure.

CD was then used to characterize the secondary structure of the peptides. Both C_2D and C_2Dcc show the characteristic PPII spectra maxima at 225 nm and minima at 190 nm, indicating that the secondary structure has not been greatly disrupted by the formation of the covalent cross-links (Figure S6). C_2Dcc had a higher MRE than C_2D , indicating that the triple helical concentration is increased by the presence of the cross-links. As mentioned above, the supramolecular C_2D has a T_m of 24.0 °C. Though C_2Dcc is covalently tethered, the triple helical structure can nevertheless unfold; thus, a melting temperature is still observed for covalently captured species. The melting temperature of C_2Dcc was increased to a T_m of 39.6 °C, an increase in stability of over 15.6 °C above the supramolecular structure (Figure 3). Comparatively, the previous K–D system maintained the same T_m for the covalently captured helices as for the supramolecular A_2B , 35.6 and 37.0 °C, respectively (Figure 3).⁵⁸ Furthermore, the transition observed for the A_2Bcc system is very broad, indicating that multiple species are present because of the varying number of isopeptide bonds formed. These two observations reinforce the superiority of K–E bond stabilization over K–D.

Additionally, the recovery of the triple helical fold can be followed by CD as the sample is cooled to 5 °C. It is known that heterotrimer CMPs have slow refolding rates, and indeed,

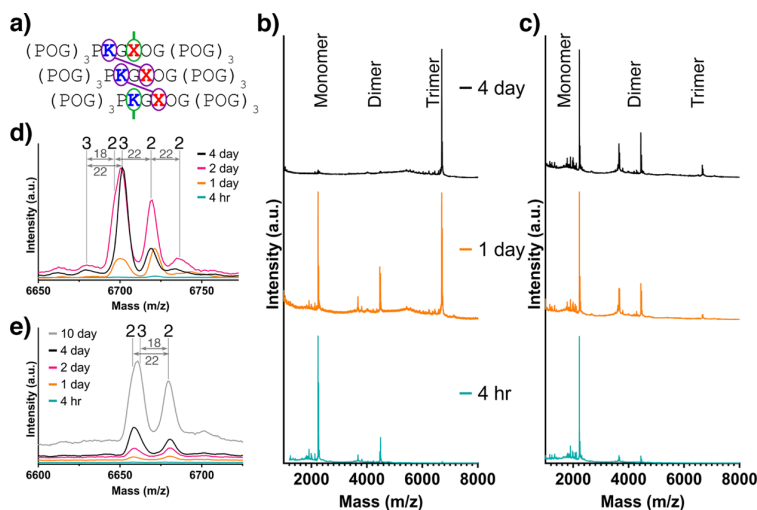


Figure 4. Comparison of K–D and K–E amide bond formation in homotrimers. (a) Sequence of peptides highlighting the guest region of peptides where X is either E or D. Two axial and one lateral bonds are highlighted in purple and green, respectively. Peptides are N-terminally acetylated and C-terminally amidated. (b–e) MALDI-TOF MS analysis. (b) KGE homotrimer after 4 h, 2 days, and 4 days of covalent capture reaction showing an increase in the trimer peak as the monomer peak decreases. (c) KGD homotrimer after 4 h, 2 days, and 4 days of covalent capture reaction, showing a very slow increase in the trimer peak over time. Zoom-in of trimer region for (d) KGE and (e) KGD.

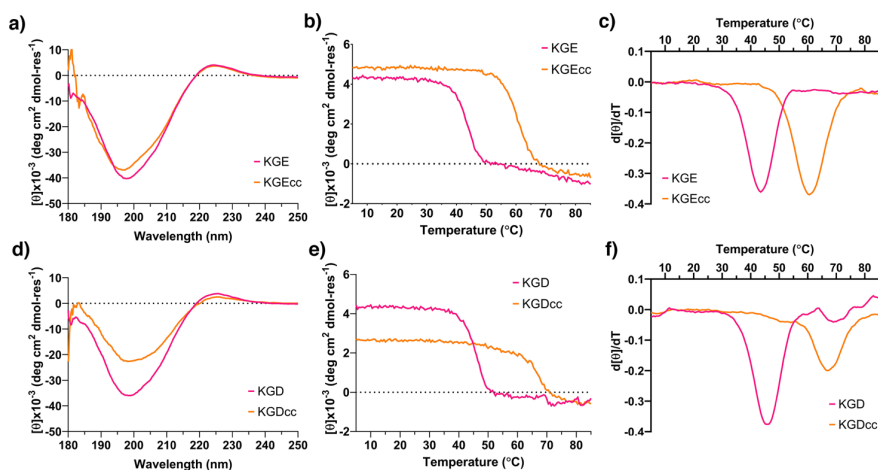


Figure 5. Comparison of K–E and K–D homotrimers by CD. KGE and KGEcc (a) spectra, (b) melting curve, and (c) first derivative of the melting curve showing an increase in both MRE and melting temperature following covalent capture. KGD and KGDcc (d) spectra, (e) melting curve, and (f) first derivative of the melting curve showing an increase in melting temperature but a decrease in MRE following covalent capture.

the supramolecular C_2D shows little fold recovery across the refolding curve and does not regain its full MRE until after nearly a week of refolding. Comparatively, C_2Dcc refolds with the decreasing temperature showing nearly zero hysteresis, a phenomenon previously observed in the A_2Bcc system (Figure 3). This quick fold recovery is yet another feature of covalent capture that improves CMP practicality.

Homotrimer Comparison. To simplify the comparison between K–D and K–E bond formation, we expanded our analysis into homotrimers. Host–guest peptides of the form Ac-(POG) $_3$ PKGXOG(POG) $_3$ -NH $_2$, where X is D or E, were synthesized and purified (Figures S8 and S9). When assembled, these homotrimers contain two axial and one lateral interactions (Figure 4a). Then, the homotrimers, termed KGD and KGE based on the guest sequence, were covalently captured.

Reaction Progression. Using the previously described 1:40:4 ratio (isopeptide bond/EDC/HOBt), the reaction progress was again followed by MALDI-MS. The KGD peptide shows the presence of trimer after 1 day of reaction (Figures 4, S10). After 5 days, the trimer peak becomes much broader possibly indicating the presence of three amide bonds overlapping with two amide bonds. Even after running the reaction for 10 days, the trimer never becomes the major species present. Additionally, there is still a substantial fraction of monomer present in the MALDI spectra. In order to prepare enough covalently captured KGD (KGDcc) for further characterization, large-scale reactions had additional activating reagents added on day 4 in hopes of pushing the reaction toward completion. However, as can be seen in Figure S10a, this did not result in trimer becoming a major species.

KGE, by comparison, shows the presence of trimer after only 4 h (Figure 4b). It is unclear if there are two amide bonds

formed or all three because of the lack of resolution at this mass range and the close mass difference between [two amide bonds formed + H]⁺ and [three amide bonds formed + Na]⁺ (Figure 4d). Notably, after 2 days, the monomer and trimer peaks are nearly equivalent in intensity. At day 4, there is minimal monomer left, showing that the reaction has gone nearly to completion, a drastic improvement over the KGD reaction rate.

Other ratios of activating reagents were tested for both KGD (Figures S12–S14) and KGE (Figures S15–S17). Lower amounts of reagents led to slower amide bond formation. Higher reagent concentrations led to quicker amide bond formation but with unwanted oligomerization. Therefore, the optimal condition was still determined to be the previously described 1:40:4 ratio (isopeptide bond/EDC/HOBt) used throughout this study.

Purification of Cross-Linked Triple Helices. KGDcc was subsequently purified by SEC with difficulty: heating the sample to separate unreacted monomer from covalently bonded trimer was not beneficial. The faster folding rates of the KGD homotrimer (as compared to heterotrimer) resulted in both supramolecular and covalently captured KGD being collected in the same fractions. Thus, the “pure” KGDcc MALDI spectra still showed some monomer to be present (Figure S10). KGEcc, by comparison, was easily purified by HPLC, as little monomer remained in the crude sample, and the final MALDI spectra did not show the presence of monomer (Figure S11).

Monomer to Trimer Equilibrium. The SEC chromatograms of the supramolecular homotrimers are notably different from those of the heterotrimers (Figures S18 and S19). Under native conditions, KGD and KGE are mostly trimers. When heated, both become mostly monomer with a smaller trimer peak. This is due to the increased folding kinetics of homotrimers, resulting in the KGD and KGE samples partially refolding during the separation process. KGDcc natively appears to be all trimer. However, upon heating, a monomer peak emerges, illustrating the difficulty in purifying KGDcc. KGEcc by comparison shows all trimer under native and heated conditions, showing negligible monomer present.

Structure Analysis. CD was then used to compare the secondary structure and stability (Figure 5). Both covalently captured homotrimers had characteristic PPII spectra (Figure 5a,d) with maxima near 225 nm. Additionally, both KGDcc and KGEcc showed substantially increased melting temperatures compared to their noncovalent counterparts of +19.5 °C (*T_m* of 45.5 °C for KGD and 67.0 °C for KGDcc) and +17.0 °C (*T_m* of 43.5 °C for KGE and 60.5 °C for KGEcc). This is in contrast to what was observed for heterotrimers where the A₂B heterotrimer containing Lys–Asp isopeptide bonds resulted in a modest decrease in melting temperature. Triple helices containing Lys–Glu isopeptide bonds resulted in increased melting temperatures in all cases. In addition to the differences in *T_m*, the homotrimers also showed differences in their MRE, an indicator of the percentage of the peptide adopting the PPII secondary structure. KGEcc displayed a modestly increased MRE compared to its supramolecular form, while KGDcc displayed a significantly decreased MRE. Based on this analysis, we speculate that the Lys–Asp bond may result in a small amount of strain on the PPII and triple helical structure, which is tolerated when there are only a few such isopeptide bonds, but results in mild destabilization and deviations from PPII conformation when in larger number.

Next, the rate of refolding was examined. Samples were heated to 85 °C, where structures are completely unfolded, for 15 min and then quickly transferred to the CD with the temperature set to 5 °C to observe the refolding as it occurred (Figure S21). As expected, the covalently captured samples exhibit faster folding recovery because of the strands being locked in close proximity. KGD returns to 50% of its original fold after 8 min, but it takes over 5.5 h to return to 100%. KGE by comparison takes 1.2 min to reach 50% and 3.4 h to reach 100%. KGDcc takes <10 s (the limit of our instrumentation) to reach 50% folded and 25 min to reach 100% folded. KGEcc takes <10 s to reach 50% folded and only 19 min to reach 100% folded. This again demonstrates the dramatic improvement covalent capture has on the ability of the triple helix to refold.

Because of the improvements we observed for K–E covalent capture, particularly its rate of reaction, ability to be pushed to near 100% yield, and ease of purification, we sought to crystallize the KGEcc helix in order to subsequently solve its molecular structure (Figure 6). We had previously published

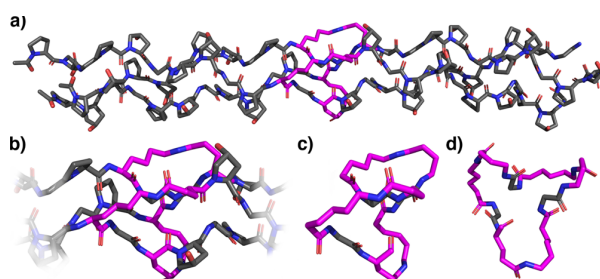


Figure 6. Crystal structure of KGEcc. (a) Full-length triple helix. (b) View of cross-linked region. (c) The 45-membered ring formed by cross-linking. (d) The 45-membered ring formed by cross-linking as viewed looking down the helix axis.

the crystal structure of KGE;²⁷ however, we re-refined that structure using more modern crystallographic refinement methods in order to critically compare the supramolecular starting material (KGE) and the covalently captured product (KGEcc). Crystallization was successful and the triple helical structure was solved to a resolution of 1.4 Å (see the Supporting Information for details). The major conclusion is that the backbone of the triple helix is almost entirely unaffected by the isopeptide bond formation. After superposition, the root-mean-square deviations (rmsds) of the positions of the backbone atoms in the captured side chain region (PKGEOG) were 0.25 and 0.33 Å for the two helices present in the KGE unit cell (Figure S24). The rmsds in positions of the two triplets immediately N-terminal to the cross-linked region were 0.18 and 0.34 Å and of the two triplets C-terminal were 0.31 and 0.24 Å. Comparatively, the rmsds in positions between the two KGE helices themselves were 0.24, 0.31, and 0.33 Å for the center, the two triplets N-terminal, and two triplets C-terminal, respectively, demonstrating no significant difference between the backbone of the covalently captured versus supramolecular structures. All interstrand backbone hydrogen bonding remains intact and with good hydrogen bonding distances (all between 2.0 and 2.3 Å from nitrogen to carbonyl oxygen). Additionally, the Φ , Ψ , and χ_1 dihedral angles are all comparable between the KGE and KGEcc structures (Figures S25–S27). This reinforces the

results as observed by CD that the triple helix is not disrupted by amide bond formation. Additionally, we are able to directly confirm that the covalent connectivity between the side chains is as expected: two axial charge pairs form isopeptide amide bonds between the leading and middle strand and from the middle to trailing strand. In addition to these two expected bonds, we clearly see the formation of one lateral K–E amide bond between the trailing and leading strand (Figure 7).

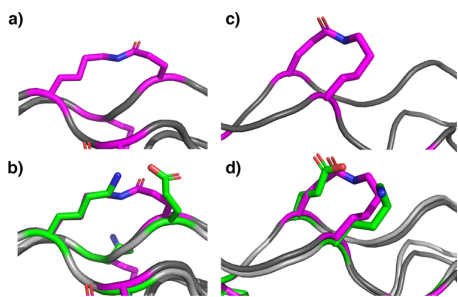


Figure 7. Comparison of isopeptide bonds and charge-pair interactions. (a) Axial isopeptide bond and (b) overlay of axial isopeptide bond and charge pair (c) lateral isopeptide bond and (d) overlay of lateral isopeptide bond and charge pair.

Interestingly, these three bonds together create a circular peptide loop spanning all three peptide chains, covering nine amino acids in the sequence KGE-KGE-KGE, and creating a 45-atom ring (Figure 6c,d).

From the above results and discussion, a few additional observations are notable. First, we and others have long noted that axial K–D charge pairs confer greater stability and much greater specificity to supramolecular heterotrimer design (over that of K–E charge pairs), which is why our early designs used this motif.^{28,58,59} Despite this superiority for supramolecular design, axial K–D is not as useful for covalent capture as K–E isopeptide bonds are formed more rapidly, in higher yield, result in improved structural fidelity, and result in comparably stable covalent structures. As noted above, we believe that this is due to the extra methylene in the glutamate side chain, which provides the conformational flexibility necessary to adopt a strain-free PPII secondary structure and the triple helix. Second, while the axial geometry for K–D and K–E charge pairs has been demonstrated repeatedly to be more effective at stabilizing a supramolecular triple helical structure, the lateral geometry was also captured in the K–E design (Figure 7). These insights underscore the idea that design considerations for supramolecular structures and their covalently captured counterparts are frequently not the same, which is an ongoing challenge in the development of any covalently captured system. Extending side chain length slightly decreased the supramolecular stability but greatly improved many aspects of covalent capture. Additionally, side chains not involved in supramolecular design might be unintentionally cross-linked, as is the case with the lateral K–E charge pair. Taking into consideration these differences can allow one to modify a supramolecular design for a covalently captured one.

CONCLUSIONS

Here, we have used isopeptide amide bond formation to covalently capture both homo- and hetero-trimeric collagen helices in order to compare the use of K–D and K–E bonds.

The reaction rate and yield observed for K–E isopeptide bond formation is improved for both homo- and hetero-trimers over that of K–D bond formation. K–E heterotrimer shows all possible amide bonds formed, while K–D heterotrimer does not attain complete cross-linking. Similarly, the K–E homotrimer reaction is able to go to completion, while K–D homotrimer still retains a significant monomer population. Both K–D and K–E cross-links improve trimer to monomer ratios in SEC. Both allow for purification of the trimer, however with much greater difficulty and lower yield for K–D homotrimer. While both K–D and K–E cross-links improve thermal stability and refolding of the triple helical backbone, only K–E cross-links increase the relative content of PPII helix. Thus, overall, covalent capture of collagen triple helices through K–E isopeptide bond formation is superior to that of K–D. Comparison of these systems allows elucidation of necessary design adjustments from supramolecular to covalent capture systems. Crystallization of the K–E triple helix revealed a triple helix nearly identical to that of the supramolecular helix, demonstrating that K–E covalent capture is a valid method for retaining and stabilizing the triple helical structure for future applications.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.biomac.0c00878>.

Additional characterization data of individual peptides, purification of covalently captured triple helices, MALDI-MS spectra of additional covalent capture reaction conditions, crystallographic data collection, and structural comparisons for KGE and KGEcc (PDF)

AUTHOR INFORMATION

Corresponding Author

Jeffrey D. Hartgerink – Department of Chemistry and Department of Bioengineering, Rice University, Houston, Texas 77005, United States; orcid.org/0000-0002-3186-5395; Email: jdh@rice.edu

Authors

Sarah A. H. Hulgán – Department of Chemistry, Rice University, Houston, Texas 77005, United States
Abhishek A. Jalan – Department of Biochemistry, University of Bayreuth, Bayreuth 95447, Germany
I-Che Li – Department of Chemistry, Rice University, Houston, Texas 77005, United States
Douglas R. Walker – Department of Chemistry, Rice University, Houston, Texas 77005, United States
Mitchell D. Miller – Department of Biosciences, Rice University, Houston, Texas 77005, United States
Abigail J. Kosgei – Department of Biosciences, Rice University, Houston, Texas 77005, United States
Weijun Xu – Department of Biosciences, Rice University, Houston, Texas 77005, United States
George N. Phillips, Jr. – Department of Biosciences and Department of Chemistry, Rice University, Houston, Texas 77005, United States

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.biomac.0c00878>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by grants from the Welch Foundation (C1557) and the NSF (CHE 1709631). Additional funding from the NSF Science and Technology Center BioXFEL contract number STC-1231306 and CPRIT grant RP160652 is gratefully acknowledged. GM/CA@APS has been funded in whole or in part with Federal funds from the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006). The Eiger 16M detector at GM/CA-XSD was funded by NIH grant S10 OD012289. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357. We thank Dr. Yizhi J. Tao with for assistance with the updating KGD (3u29) and KGE (3t4f) crystal structures.

■ REFERENCES

- (1) Aida, T.; Meijer, E. W.; Stupp, S. I. Functional Supramolecular Polymers. *Science* **2012**, *335*, 813–817.
- (2) Hartgerink, J. D. Covalent capture: a natural complement to self-assembly. *Curr. Opin. Chem. Biol.* **2004**, *8*, 604–609.
- (3) Prins, L. J.; Scrimin, P. Covalent Capture: Merging Covalent and Noncovalent Synthesis. *Angew. Chem., Int. Ed.* **2009**, *48*, 2288–2306.
- (4) Brožek, J.; Grande, F.; Anderson, J. T.; Keys, A. Densitometric analysis of body composition: Revision of some quantitative assumptions. *Ann. N.Y. Acad. Sci.* **1963**, *110*, 113–140.
- (5) Yu, Z.; An, B.; Ramshaw, J. A. M.; Brodsky, B. Bacterial collagen-like proteins that form triple-helical structures. *J. Struct. Biol.* **2014**, *186*, 451–461.
- (6) Celerin, M.; Ray, J. M.; Schisler, N. J.; Day, A. W.; Stetler-Stevenson, W. G.; Laudénbach, D. E. Fungal fimbriae are composed of collagen. *EMBO J.* **1996**, *15*, 4445–4453.
- (7) Thiel, S. Complement activating soluble pattern recognition molecules with collagen-like regions, mannan-binding lectin, ficolins and associated proteins. *Mol. Immunol.* **2007**, *44*, 3875–3888.
- (8) Orgel, J. P. R. O.; Irving, T. C.; Miller, A.; Wess, T. J. Microfibrillar structure of type I collagen in situ. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 9001–9005.
- (9) Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrick, A.; Brodsky, B. Amino acid propensities for the collagen triple-helix. *Biochemistry* **2000**, *39*, 14960–14967.
- (10) Hodges, J. A.; Raines, R. T. Stereoelectronic effects on collagen stability: the dichotomy of 4-fluoroproline diastereomers. *J. Am. Chem. Soc.* **2003**, *125*, 9262–9263.
- (11) Acevedo-Jake, A. M.; Clements, K. A.; Hartgerink, J. D. Synthetic, Register-Specific, AAB Heterotrimers to Investigate Single Point Glycine Mutations in Osteogenesis Imperfecta. *Biomacromolecules* **2016**, *17*, 914–921.
- (12) Long, C. G.; Braswell, E.; Zhu, D.; Apigo, J.; Baum, J.; Brodsky, B. Characterization of Collagen-Like Peptides Containing Interruptions in the Repeating Gly-X-Y Sequence. *Biochemistry* **1993**, *32*, 11688–11695.
- (13) Clements, K. A.; Acevedo-Jake, A. M.; Walker, D. R.; Hartgerink, J. D. Glycine Substitutions in Collagen Heterotrimers Alter Triple Helical Assembly. *Biomacromolecules* **2017**, *18*, 617–624.
- (14) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. Crystal and Molecular Structure of a Collagen-Like Peptide at 1.9 Å Resolution. *Science* **1994**, *266*, 75–81.
- (15) Fallas, J. A.; Hartgerink, J. D. Computational design of self-assembling register-specific collagen heterotrimers. *Nat. Commun.* **2012**, *3*, 1087.
- (16) Xu, F.; Zahid, S.; Silva, T.; Nanda, V. Computational Design of a Collagen A:B:C-Type Heterotrimer. *J. Am. Chem. Soc.* **2011**, *133*, 15260–15263.
- (17) Jalan, A. A.; Demeler, B.; Hartgerink, J. D. Hydroxyproline-free single composition ABC collagen heterotrimer. *J. Am. Chem. Soc.* **2013**, *135*, 6014–6017.
- (18) Jalan, A. A.; Hartgerink, J. D. Simultaneous control of composition and register of an AAB-type collagen heterotrimer. *Biomacromolecules* **2013**, *14*, 179–185.
- (19) Fallas, J. A.; Lee, M. A.; Jalan, A. A.; Hartgerink, J. D. Rational design of single-composition ABC collagen heterotrimers. *J. Am. Chem. Soc.* **2012**, *134*, 1430–1433.
- (20) Russell, L. E.; Fallas, J. A.; Hartgerink, J. D. Selective assembly of a high stability AAB collagen heterotrimer. *J. Am. Chem. Soc.* **2010**, *132*, 3242–3243.
- (21) Hodges, J. A.; Raines, R. T. Stereoelectronic and steric effects in the collagen triple helix: toward a code for strand association. *J. Am. Chem. Soc.* **2005**, *127*, 15923–15932.
- (22) Koide, T.; Nishikawa, Y.; Takahara, Y. Synthesis of heterotrimeric collagen models containing Arg residues in Y-positions and analysis of their conformational stability. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 125–128.
- (23) Xiao, J.; Sun, X.; Madhan, B.; Brodsky, B.; Baum, J. NMR Studies Demonstrate a Unique AAB Composition and Chain Register for a Heterotrimeric Type IV Collagen Model Peptide Containing a Natural Interruption Site. *J. Biol. Chem.* **2015**, *290*, 24201–24209.
- (24) Xu, F.; Silva, T.; Joshi, M.; Zahid, S.; Nanda, V. Circular Permutation Directs Orthogonal Assembly in Complex Collagen Peptide Mixtures. *J. Biol. Chem.* **2013**, *288*, 31616–31623.
- (25) O’Leary, L. E. R.; Fallas, J. A.; Hartgerink, J. D. Positive and negative design leads to compositional control in AAB collagen heterotrimers. *J. Am. Chem. Soc.* **2011**, *133*, 5432–5443.
- (26) Hentzen, N. B.; Islami, V.; Köhler, M.; Zenobi, R.; Wennemers, H. A Lateral Salt Bridge for the Specific Assembly of an ABC-Type Collagen Heterotrimer. *J. Am. Chem. Soc.* **2020**, *142*, 2208–2212.
- (27) Fallas, J. A.; Dong, J.; Tao, Y. J.; Hartgerink, J. D. Structural insights into charge pair interactions in triple helical collagen-like proteins. *J. Biol. Chem.* **2012**, *287*, 8039–8047.
- (28) Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrick, A.; Brodsky, B. Electrostatic Interactions Involving Lysine Make Major Contributions to Collagen Triple-Helix Stability. *Biochemistry* **2005**, *44*, 1414–1422.
- (29) Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrick, A.; Brodsky, B. Peptide investigations of pairwise interactions in the collagen triple-helix. *J. Mol. Biol.* **2002**, *316*, 385–394.
- (30) Mizuno, K.; Boudko, S. P.; Engel, J.; Bächinger, H. P. Kinetic Hysteresis in Collagen Folding. *Biophys. J.* **2010**, *98*, 3004–3014.
- (31) Persikov, A. V.; Xu, Y. J.; Brodsky, B. Equilibrium thermal transitions of collagen model peptides. *Protein Sci.* **2004**, *13*, 893–902.
- (32) Xu, F.; Zhang, L.; Koder, R. L.; Nanda, V. De Novo Self-Assembling Collagen Heterotrimers Using Explicit Positive and Negative Design. *Biochemistry* **2010**, *49*, 2307–2316.
- (33) Rump, E. T.; Rijkers, D. T. S.; Hilbers, H. W.; de Groot, P. G.; Liskamp, R. M. J. Cyclotrimeratrylene (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.* **2002**, *8*, 4613–4621.
- (34) Goodman, M.; Feng, Y.; Melacini, G.; Taulane, J. P. A Template-Induced Incipient Collagen-Like Triple-Helical Structure. *J. Am. Chem. Soc.* **1996**, *118*, 5156–5157.
- (35) Kwak, J.; Capua, A. D.; Locardi, E.; Goodman, M. TREN (Tris(2-aminoethyl)amine): An Effective Scaffold for the Assembly of Triple Helical Collagen Mimetic Structures. *J. Am. Chem. Soc.* **2002**, *124*, 14085–14091.
- (36) Kemp, D. S.; Petrakis, K. S. Synthesis and conformational analysis of cis,cis-1,3,5-trimethylcyclohexane-1,3,5-tricarboxylic acid. *J. Org. Chem.* **1981**, *46*, 5140–5143.
- (37) Feng, Y.; Melacini, G.; Taulane, J. P.; 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.* **1996**, *118*, 10351–10358.

(38) Goodman, M.; Melacini, G.; Feng, Y. Collagen-Like Triple Helices Incorporating Peptoid Residues. *J. Am. Chem. Soc.* **1996**, *118*, 10928–10929.

(39) Li, Y.; Mo, X.; Kim, D.; Yu, S. M. Template-tethered collagen mimetic peptides for studying heterotrimeric triple-helical interactions. *Biopolymers* **2011**, *95*, 94–104.

(40) Horng, J.-C.; Hawk, A. J.; Zhao, Q.; Benedict, E. S.; Burke, S. D.; Raines, R. T. Macrocyclic scaffold for the collagen triple helix. *Org. Lett.* **2006**, *8*, 4735–4738.

(41) Priem, C.; Geyer, A. Reversible Covalent End-Capping of Collagen Model Peptides. *Chem.—Eur. J.* **2019**, *25*, 14278–14283.

(42) Ottl, J.; Moroder, L. Disulfide-Bridged Heterotrimeric Collagen Peptides Containing the Collagenase Cleavage Site of Collagen Type I. Synthesis and Conformational Properties. *J. Am. Chem. Soc.* **1999**, *121*, 653–661.

(43) Mechling, D. E.; Bächinger, H. P. The Collagen-like Peptide (GER)15GPCCG Forms pH-dependent Covalently Linked Triple Helical Trimers. *J. Biol. Chem.* **2000**, *275*, 14532–14536.

(44) Ottl, J.; Gabriel, D.; Murphy, G.; Knäuper, V.; Tominaga, Y.; Nagase, H.; Kröger, M.; Tschesche, H.; Bode, W.; Moroder, L. Recognition and catabolism of synthetic heterotrimeric collagen peptides by matrix metalloproteinases. *Chem. Biol.* **2000**, *7*, 119–132.

(45) Krishna, O. D.; Kiick, K. L. Supramolecular Assembly of Electrostatically Stabilized, Hydroxyproline-Lacking Collagen-Mimetic Peptides. *Biomacromolecules* **2009**, *10*, 2626–2631.

(46) Golser, A. V.; Röber, M.; Börner, H. G.; Scheibel, T. Engineered Collagen: A Redox Switchable Framework for Tunable Assembly and Fabrication of Biocompatible Surfaces. *ACS Biomater. Sci. Eng.* **2018**, *4*, 2106–2114.

(47) Tanrikulu, I. C.; Raines, R. T. Optimal interstrand bridges for collagen-like biomaterials. *J. Am. Chem. Soc.* **2014**, *136*, 13490–13493.

(48) Barth, D.; Musiol, H.-J.; Schütt, M.; Fiori, S.; Milbradt, A. G.; Renner, C.; Moroder, L. The Role of Cystine Knots in Collagen Folding and Stability, Part I. Conformational Properties of (Pro-Hyp-Gly)5 and (Pro-(4S)-FPro-Gly)5 Model Trimers with an Artificial Cystine Knot. *Chem.—Eur. J.* **2003**, *9*, 3692–3702.

(49) Fields, C. G.; Lovdahl, C. M.; Miles, A. J.; Hageini, V. L. M.; Fields, G. B. Solid-Phase synthesis and stability of triple-helical peptides incorporating native collagen sequences. *Biopolymers* **1993**, *33*, 1695–1707.

(50) Tanaka, Y.; Suzuki, K.; Tanaka, T. Synthesis and stabilization of amino and carboxy terminal constrained collagenous peptides. *J. Pept. Res.* **1998**, *51*, 413–419.

(51) Henkel, W.; Vogl, T.; Echner, H.; Voelter, W.; Urbanke, C.; Schleuder, D.; Rauterberg, J. Synthesis and Folding of Native Collagen III Model Peptides. *Biochemistry* **1999**, *38*, 13610–13622.

(52) Sweeney, S. M.; Guy, C. A.; Fields, G. B.; Antonio, J. D. S. Defining the domains of type I collagen involved in heparin-binding and endothelial tube formation. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 7275–7280.

(53) Fields, C. G.; Grab, B.; Lauer, J. L.; Fields, G. B. Purification and Analysis of Synthetic, Triple-Helical “Minicollagens” by Reversed-Phase High-Performance Liquid Chromatography. *Anal. Biochem.* **1995**, *231*, 57–64.

(54) Malcor, J.-D.; Juskaite, V.; Gavrilidou, D.; Hunter, E. J.; Davidenko, N.; Hamaia, S.; Sinha, S.; Cameron, R. E.; Best, S. M.; Leitinger, B.; Farndale, R. W. Coupling of a specific photoreactive triple-helical peptide to crosslinked collagen films restores binding and activation of DDR2 and VWF. *Biomaterials* **2018**, *182*, 21–34.

(55) Hentzen, N. B.; Smeenk, L. E. J.; Witek, J.; Riniker, S.; Wennemers, H. Cross-Linked Collagen Triple Helices by Oxime Ligation. *J. Am. Chem. Soc.* **2017**, *139*, 12815–12820.

(56) Byrne, C.; McEwan, P. A.; Emsley, J.; Fischer, P. M.; Chan, W. C. End-stapled homo and hetero collagen triple helices: a click chemistry approach. *Chem. Commun.* **2011**, *47*, 2589–2591.

(57) Egli, J.; Siebler, C.; Köhler, M.; Zenobi, R.; Wennemers, H. Hydrophobic Moieties Bestow Fast-Folding and Hyperstability on Collagen Triple Helices. *J. Am. Chem. Soc.* **2019**, *141*, 5607–5611.

(58) Li, I.-C.; Hulgán, S. A. H.; Walker, D. R.; Farndale, R. W.; Hartgerink, J. D.; Jalan, A. A. Covalent Capture of a Heterotrimeric Collagen Helix. *Org. Lett.* **2019**, *21*, 5480–5484.

(59) Jalan, A. A.; Sammon, D.; Hartgerink, J. D.; Brear, P.; Stott, K.; Hamaia, S. W.; Hunter, E. J.; Walker, D. R.; Leitinger, B.; Farndale, R. W. Chain alignment of collagen I deciphered using computationally designed heterotrimers. *Nat. Chem. Biol.* **2020**, *16*, 423–429.

(60) Kang, H. J.; Coulibaly, F.; Clow, F.; Proft, T.; Baker, E. N. Stabilizing Isopeptide Bonds Revealed in Gram-Positive Bacterial Pilus Structure. *Science* **2007**, *318*, 1625–1628.

(61) Greenberg, C. S.; Birckbichler, P. J.; Rice, R. H. Transglutaminases: Multifunctional cross-linking enzymes that stabilize tissues. *FASEB J.* **1991**, *5*, 3071–3077.

(62) Winter, G.; Waterman, D. G.; Parkhurst, J. M.; Brewster, A. S.; Gildea, R. J.; Gerstel, M.; Fuentes-Montero, L.; Vollmar, M.; Michels-Clark, T.; Young, I. D.; Sauter, N. K.; Evans, G. DIALLS : implementation and evaluation of a new integration package. *Acta Crystallogr., Sect. D: Struct. Biol.* **2018**, *74*, 85–97.

(63) Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 282–292.

(64) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP 4 suite and current developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 235–242.

(65) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.

(66) Langer, G.; Cohen, S. X.; Lamzin, V. S.; Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **2008**, *3*, 1171–1179.

(67) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 213–221.

(68) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486–501.

(69) Borbulevych, O. Y.; Plumley, J. A.; Martin, R. I.; Merz, K. M.; Westerhoff, L. M. Accurate macromolecular crystallographic refinement: incorporation of the linear scaling, semiempirical quantum-mechanics program DivCon into the PHENIX refinement package. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2014**, *70*, 1233–1247.

(70) Yennamalli, R.; Arangarasan, R.; Bryden, A.; Gleicher, M.; Phillips, G. N., Jr. Using a commodity high-definition television for collaborative structural biology. *J. Appl. Crystallogr.* **2014**, *47*, 1153–1157.

(71) wwPDB consortium. Protein Data Bank: the single global archive for 3D macromolecular structure data. *Nucleic Acids Res.* **2019**, *47*, D520–D528.