J|A|C|S
ARTICLES
Published on Web 02/13/2007

Self-Assembled Heterotrimeric Collagen Triple Helices Directed through Electrostatic Interactions

Varun Gauba and Jeffrey D. Hartgerink*

Contribution from the Department of Chemistry and Department of Bioengineering, Rice University, 6100 Main Street, Mail Stop 60, Houston, Texas 77005

Received November 22, 2006; E-mail: jdh@rice.edu

Abstract: Collagen, a fibrous protein, is an essential structural component of all connective tissues such as cartilage, bones, ligaments, and skin. Type I collagen, the most abundant form, is a heterotrimer assembled from two identical α1 chains and one α2 chain. However, most synthetic systems have addressed homotrimeric triple helices. In this paper we examine the stability of several heterotrimeric collagen-like triple helices with an emphasis on electrostatic interactions between peptides. We synthesize seven 30 amino acid peptides with net charges ranging from −10 to +10. These peptides were mixed, and their ability to form heterotrimers was assessed. We successfully show the assembly of five different AAB heterotrimers and one ABC heterotrimer. The results from this study indicate that intermolecular electrostatic interactions can be utilized to direct heterotrimer formation. Furthermore, amino acids with poor stability in collagen triple helices can be "rescued" in heterotrimers containing amino acids with known high triple helical stability. This mechanism allows collagen triple helices to have greater chemical diversity than would otherwise be allowed.

Introduction

The number of protein folding motifs that are well understood in terms of a sequence−structure relationship is limited. One of the best understood motifs is the α-helical coiled coil which can be designed with a very high probability of success as both homomers and heteromers.[1−6] This understanding has had a broad impact in areas as diverse as protein folding,[1−6] catalyst design, nanotechnology,[7−9] and origins of life research.[10−13] More recently, β-sheet design has also become relatively straight forward allowing greater understanding of the mechanism behind neurodegenerative diseases, nanotechnology, and tissue engineering.[14−18] The design of these protein motifs is largely based on hydrophobic packing interactions and encompasses two of the most prevalent protein secondary structures. Collagen, which is the most abundant protein in the human body, has been less intensively studied, particularly with respect to de novo design. It has a unique triple helical structure in which three left-handed poly-proline type II helices wind around one another to form a right-handed super helix.[19−23] The sequence design criteria for this folding motif is not based on hydrophobicity, but instead the stability of the triple helix is a result of an extensive network of $CO_{(X)}-NH_{(Gly)}$ hydrogen bonds and the tightly packed nature of the triple helix permitted by the presence of glycine every third residue in an X−Y−Gly repeating motif. Glycine is completely buried in the interior of the triple helix, while the side chains of the X and Y residues are oriented away from the helix core where they are exposed to the solvent[24] and can take part in intra- and intermolecular side chain interactions.[25−27] Hydroxyproline (Hyp = O) also plays an

(1) Crick, F. H. C. *Acta Crystallogr.* **1953**, *6*, 689−697.
(2) DeGrado, W. F.; Lear, J. D. *J. Am. Chem. Soc.* **1985**, *107*, 7684−7689.
(3) Engel, M.; William, R. W.; Erickson, B. W. *Biochemistry* **1991**, *30*, 3161−3169.
(4) Graddis, T. J.; Myszka, D. G.; Chaiken, I. M. *Biochemistry* **1993**, *32*, 12664−12671.
(5) Monera, O. D.; Zhou, N. E.; Kay, C. M.; Hodges, R. S. *J. Biol. Chem.* **1993**, *268*, 19218−19227.
(6) Dong, H.; Hartgerink, J. D. *Biomacromolecules* **2006**, *7*, 691−695.
(7) Woolfson, D. N. *Adv. Protein Chem.* **2005**, *70*, 79.
(8) Smith, A. M.; Acquah, S. F. A.; Bone, N.; Kroto, H. W.; Ryadnov, M. G.; Stevens, M. S. P.; Walton, D. R.; Woolfson, D. N. *Angew. Chem., Int. Ed.* **2005**, *44*, 325−328.
(9) Pandya, M. J.; Spooner, G. M.; Sunde, M.; Thorpe, J. R.; Rodger, A.; Woolfson, D. N. *Biochemistry* **2000**, *39*, 8728−8734.
(10) Lee, D. H.; Granja, J. R.; Martinez, J. A.; Severin, K.; Ghadiri, M. R. *Nature* **1996**, *382*, 525−528.
(11) Lee, D. H.; Severin, K.; Yokobayashi, Y.; Ghadiri, M. R. *Nature* **1997**, *390*, 591−594.
(12) Severin, K.; Lee, D. H.; Kennan, A. J.; Ghadiri, M. R. *Nature* **1997**, *389*, 706−709.
(13) Saghatelian, A.; Yokobayashi, Y.; Soltani, K.; Ghadiri, M. R. *Nature* **2001**, *409*, 797−801.
(14) Altman, M.; Lee, P.; Rich, A.; Zhang, S. G. *Protein Sci.* **2000**, *9*, 1095−1105.
(15) Holmes, T. C.; Lacalle, S. D.; Su, X.; Rich, G.; Zhang, S. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 6728−6733.
(16) Hwang, W.; Zhang, S.; Kamm, R. D.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12916−12921.
(17) Lamm, M. S.; Rajagopal, K.; Schneider, J. P.; Pochan, D. J. *J. Am. Chem. Soc.* **2005**, *127*, 16692−16700.
(18) Yokoi, H.; Kinoshita, T.; Zhang, Shuguang. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 8414−8419.
(19) Cowan, P. M.; Mcgavin, S. *Nature* **1955**, *176*, 501−503.
(20) Sasisekharan, V. *Acta Crystallogr.* **1959**, *12*, 897−903.
(21) Rich, A.; Crick, F. H. C. *J. Mol. Biol.* **1961**, *3*, 483−506.
(22) Ramachandran, G. N. *Intl. Rev. Conn. Tis. Res.* **1963**, *1*, 127−182.
(23) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. *Science* **1994**, *266*, 75−81.
(24) Jones, E. Y.; Miller, A. *J. Mol. Biol.* **1991**, *218*, 209−219.
(25) Kramer, R. Z.; Bella, J.; Brodsky, B.; Berman, H. M. *J. Mol. Biol.* **2001**, *311*, 131−147.
(26) Kramer, R. Z.; Venugopal, M. G.; Bella, J.; Mayville, P.; Brodsky, B.; Berman, H. M. *J. Mol. Biol.* **2000**, *301*, 1191−1205.
(27) Kramer, R. Z.; Bella, J.; Mayville, P.; Brodsky, B.; Berman, H. M. *Nat. Struct. Biol.* **1999**, *6*, 454−457.

important role in triple helix stability as it is suggested to form water bridges with unused carbonyl groups in the helix[28−33] and has inductive effects[34] which lead to enhanced stability of the helix.

Naturally occurring collagens form triple helices composed of all identical (AAA, homotrimer), two different (AAB, heterotrimer), or three different (ABC, heterotrimer) polypeptides.[35] For example, collagen type I is an AAB heterotrimer formed from two identical α1(I) and one α2(I) chains.[35] However, the majority of studies on synthetic collagen-like peptides have been performed on self-assembling homotrimers.[36−45] Among these, the amino acid propensity for triple helix formation has been studied exhaustively.[46−50] These studies have shown that the repeating unit $(POG)_n$ yields the most stable triple helix. They have also shown that glutamic acid and arginine are the most commonly charged amino acids and are most stable in positions X and Y, respectively. Recently, some results have become available for heterotrimers obtained by cystine-knot strategy,[51−61] where cystine disulfide bonds covalently capture different polypeptides leading to the formation of heterotrimers. Only two studies have examined heterotrimer formation without covalently tethering the peptides together.[62,63] Furthermore, there are no studies which demonstrate the designed formation of noncovalent ABC triple helices. Nonetheless, these types of heterotrimeric structures are of critical importance for understanding the extracellular matrix as many forms of natural collagen are of the type AAB or ABC.[35]

In this paper we design, synthesize, and characterize a series of collagen-like peptides which utilize electrostatically charged amino acids to bias their self-assembly into collagen-like heterotrimeric helices. In addition to five AAB heterotrimers, for the first time we demonstrate the formation of an ABC heterotrimer utilizing only supramolecular interactions. We examine the melting temperature for each of these systems and find that the stability of a heterotrimer cannot be anticipated by individual amino acid propensities for triple helix formation. Instead we find that amino acids which are found to substantially destabilize the triple helix, such as glutamic acid, can form triple helices of high stability when paired against oppositely charged amino acids such as arginine or, more surprisingly, when paired against other neutral amino acids with high stability such as hydroxyproline. This results in the stability of the heterotrimeric triple helix being almost indistinguishable from the stability of those containing only amino acids with high triple helix stability. In the most dramatic example we show that $(EOG)_{10}$ and $(PRG)_{10}$, which are unable to form a triple helix in isolation, form high quality triple helices when mixed together. This work lays the groundwork for a better understanding of the thermal stability of heterotrimeric collagens, for the synthesis of a variety of synthetic extracellular matrix mimics[64−66] and molecular recognition devices based on the interactions reported herein.

Table 1. Sequence of Peptides Studied[a]

no.	sequence	abbreviation
1	$(POG)_{10}$	POG
2	$(POG)_2(\mathbf{E}OGPOG)_3(POG)_2$	E3
3	$(POG)_2(P\mathbf{R}GPOG)_3(POG)_2$	R3
4	$(\mathbf{E}OGPOG)_5$	E5
5	$(P\mathbf{R}GPOG)_5$	R5
6	$(\mathbf{E}OG)_{10}$	E10
7	$(P\mathbf{R}G)_{10}$	R10

[a] The N and C terminal of all peptides are acetylated and amidated, respectively. O = hydroxyproline.

Molecular and Experimental Design

We synthesized a set of $(POG)_{10}$ based peptides, with a general sequence $(POG)_{5-n}(EOGPOG)_n(POG)_{5-n}$ and $(POG)_{5-n}$-$(PRGPOG)_n(POG)_{5-n}$, where n represents the charge of each polypeptide chain, for n = 3, 5, and $(EOG)_{10}$ and $(PRG)_{10}$ for decacharged peptides. The polypeptides are abbreviated as POG, En, or Rn throughout the remainder of this paper, as shown in Table 1.

Glutamic acid, with a negative charge at neutral pH, replaces proline at X positions. In natural collagen, glutamic acid is found predominantly in the X position and is by far the most abundant

(28) Bella, J.; Brodsky, B.; Berman, H. M. *Structure* **1995**, *3*, 893−906.
(29) Privalov, P. L. *Adv. Protein Chem.* **1982**, *35*, 1−104.
(30) Vitagliano, L.; Berisio, R.; Mastrangelo, A.; Mazzarella, L.; Zagari, A. *Protein Sci.* **2001**, *10*, 2627−2632.
(31) Vitagliano, L.; Berisio, R.; Mazzarella, L.; Zagari, A. *Biopolymers* **2001**, *59*, 465−465.
(32) Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A. *Protein Sci.* **2002**, *11*, 262−270.
(33) Jenkins, C. L.; Raines, R. T. *Natl. Prod. Rep.* **2002**, *19*, 49−59.
(34) Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. *Nature* **1998**, *392*, 666−667.
(35) Gelse, K.; Poschl, E.; Aigner, T. *Adv. Drug Delivery Rev.* **2003**, *55*, 1531−1546.
(36) Kinberger, G. A.; Taulane, J. P.; Goodman, M. *Tetrahedron* **2006**, *62*, 5280−5286.
(37) Cai, W. B.; Kwok, S. W.; Taulane, J. P.; Goodman, M. *J. Am. Chem. Soc.* **2004**, *126*, 15030−15031.
(38) Kwak, J.; De Capua, A.; Locardi, E.; Goodman, M. *J. Am. Chem. Soc.* **2002**, *124*, 14085−14091.
(39) Wang, A. Y.; Mo, X.; Chen, C. S.; Yu, S. M. *J. Am. Chem. Soc.* **2005**, *127*, 4130−4131.
(40) Bann, J. G.; Bachinger, H. P.; Peyton, D. H. *Biochemistry* **2003**, *42*, 4042−4048.
(41) Babu, I. R.; Ganesh, K. N. *J. Am. Chem. Soc.* **2001**, *123*, 2079−2080.
(42) Yu, Y. C.; Berndt, P.; Tirrell, M.; Fields, G. B. *J. Am. Chem. Soc.* **1996**, *118*, 12515−12520.
(43) Berndt, P.; Fields, G. B.; Tirrell, M. *J. Am. Chem. Soc.* **1995**, *117*, 9515−9522.
(44) Lisman, T.; Raynal, N.; Groeneveld, D.; Maddox, B.; Peachey, A. R.; G., H. E.; Groot, P. G. d.; Farndale, R. W. *Blood* **2006**, *108*, 3753−3756.
(45) Raynal, N.; Hamaia, S. W.; Siljander, P. R.; Maddox, B.; Peachey, A. R.; Fernandez, R.; Foley, L. J.; Slatter, D. A.; Jarvis, G. E.; Farndale, R. W. *J. Biol. Chem.* **2006**, *281*, 3821−3831.
(46) Katz, E. P.; David, C. W. *Biopolymers* **1990**, *29*, 791−798.
(47) Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrick, A.; Brodsky, B. *Biochemistry* **2000**, *39*, 14960−14967.
(48) Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrick, A.; Brodsky, B. *J. Mol. Biol.* **2002**, *316*, 385−394.
(49) Persikov, A. V.; Ramshaw, J. A. M.; Brodsky, B. *J. Biol. Chem.* **2005**, *280*, 19343−19349.
(50) Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrick, A.; Brodsky, B. *Biochemistry* **2005**, *44*, 1414−1422.
(51) Slatter, D. A.; Foley, L. A.; Peachey, A. R.; Nietlispach, D.; Farndale, R. W. *J. Mol. Biol.* **2006**, *359*, 289−298.
(52) Sacca, B.; Renner, C.; Moroder, L. *J. Mol. Biol.* **2002**, *324*, 309−318.
(53) Sacca, B.; Moroder, L. *J. Pept. Sci.* **2002**, *8*, 192−204.
(54) Muller, J. C. D.; Ottl, J.; Moroder, L. *Biochemistry* **2000**, *39*, 5111−5116.
(55) Koide, T.; Homma, D. L.; Asada, S.; Kitagawa, K. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 5230−5233.
(56) Barth, D.; Kyrieleis, O.; Frank, S.; Renner, C.; Moroder, L. *Chem.—Eur. J.* **2003**, *9*, 3703−3714.
(57) Barth, D.; Musiol, H.-J.; Schutt, M.; Fiori, S.; Milbradt, A. G.; Renner, C.; Moroder, L. *Chem.—Eur. J.* **2003**, *9*, 3692−3702.
(58) Ottl, J.; Battistuta, R.; Pieper, M.; Tschesche, H.; Bode, W.; Kuhn, K.; Moroder, L. *FEBS Lett.* **1996**, *398*, 31−36.
(59) Ottl, J.; Gabriel, D.; Knauper, V.; Tominaga, Y.; Nagase, H.; Kroger, M.; Tschesche, H.; Bode, W.; Moroder, L. *Chem. Biol.* **2000**, *7*, 119−132.
(60) Ottl, J.; Moroder, L. *Tetrahedron Lett.* **1999**, *40*, 1487−1490.
(61) Ottl, J.; Musiol, H. J.; Moroder, L. *J. Pept. Sci.* **1999**, *5*, 103−110.
(62) Slatter, D. A.; Miles, C. A.; Bailey, A. J. *J. Mol. Biol.* **2003**, *329*, 175−183.
(63) Berisio, R.; Granata, V.; Vitagliano, L.; Zagari, A. *Biopolymers* **2004**, *73*, 682−688.
(64) Kotch, F. W.; Raines, R. T. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 3028−3033.
(65) Paramonov, S. E.; Gauba, V.; Hartgerink, J. D. *Macromolecules* **2005**, *38*, 7555−7561.
(66) Kishimoto, T.; Morihara, Y.; Osanai, M.; Ogata, S.; Kamitakahara, M.; Ohtsuki, C.; Tanihara, M. *Biopolymers* **2005**, *79*, 163−172.

of the two possible negatively charged amino acids.[47] Arginine, with a positive charge at neutral pH, replaces hydroxyproline at Y positions. Arginine is found predominantly in the Y position in natural collagen and forms the most stable helix after hydroxyproline.[47] These two types of substitutions allow us to prepare oppositely charged peptides. Previously it was observed that ion pairs can be formed in homotrimeric helices between peptide strands with glutamic acid in the X position of one polypeptide and arginine in the Y position of the adjacent polypeptide.[46,67,68] In our system, we expect the presence of opposite charges on the two polypeptides to favor heterotrimeric assembly over homotrimeric assembly. All the peptides used in the study are N-terminal acetylated and C-terminal amidated in order to avoid charge interactions between the termini.[69]

Experimental Section

Peptide Synthesis and Purification. All the peptides were synthesized on an Advanced ChemTech 396 multipeptide automated synthesizer using Fmoc solid-phase chemistry based on a 0.15 mmol scale. An MBHA-rink amide resin was used. Amino acid couplings were 45 min long with the following amount of reagents: 4 equiv of amino acids, 4 equiv of 1-hydroxybenzotriazole hydrate (HoBt), 4 equiv of *O*-benzotriazole-*N*,*N*,*N*′,*N*′-tetramethyluroniumhexafluorophosphate (HBTU), 6 equiv of *N*,*N*-diisopropylethylamine (DiEA) with *N*,*N*-dimethlyformamide (DMF) as a solvent. All proline residues were double-coupled. Fmoc was deprotected with two 7-min treatments of 25% (by volume) piperidine in DMF. After completion of the peptide synthesis, the peptide was acetylated in the presence of acetic anhydride and DiEA at a molar ratio of 8.3:1 in dichloromethane (DCM). Cleavage of the peptide was accomplished by treatment of the resin with 20 mL of trifluoroacetic acid (TFA)/triisopropylsilane/H_2O (38:1:1 by volume) for 3 h at room temperature for POG and En peptides and with 20 mL of TFA/ethanedithiol/H_2O/triisopropylsilane (37.6:1:1:0.4 by volume) for 3 h at room temperature for Rn peptides. This TFA solution was collected, followed by rinsing the resin twice with neat TFA. All washings were combined and rotoevaporated to a thick solution of approximately 5 mL. The peptide was triturated by addition of 50 mL of cold diethyl ether. The precipitate was collected by centrifugation, and the pellet was washed two times with cold ether. The pellet was then dried under a vacuum overnight and redissolved in deionized water for purification by reversed phase HPLC using a C-18 column with a linear gradient of acetonitrile and water containing 0.05% TFA.

Circular Dichroism Spectroscopy. CD measurements were performed with a Jasco J-810 spectropolarimeter, equipped with a Peltier temperature control system, using quartz cells with a path length of 0.1 cm. Thermal unfolding curves were obtained by monitoring the decrease in ellipticity at 225 nm in a range of 5 °C to 95 °C at a heating rate of 10 °C/h for all the peptides except E10. Ellipticity at 218 nm was monitored for E10, as its CD spectra show a positive maximum at 218 nm. The fraction folded was calculated for the melting curves according to the following equation: $F_T = ([\theta]_T - [\theta]_{unfolded})/([\theta]_{folded} - [\theta]_{unfolded})$, where F_T is the fraction folded at temperature T, $[\theta]_T$ is the molar residue ellipticity (MRE) at temperature T, $[\theta]_{folded}$ and $[\theta]_{unfolded}$ are the MRE of the folded and unfolded forms, respectively, calculated as follows $[\theta] = (\theta * m)/(c * l * n_r)$, where θ is ellipticity in mdeg, m is molecular weight in g/mol, c is concentration in mg/mL, l is path length of the cuvette in mm, and n_r is the number of amino acid residues in the peptide.

For homotrimers, peptide solutions with a concentration of 0.2 mM in 10 mM phosphate buffer (pH 7) were used for all the experiments. For heterotrimers, various peptides were mixed in desired ratios in such a way that the final total peptide concentration was 0.2 mM, and a neutral pH was maintained by using 10 mM phosphate buffer solution.

Unfolding studies were performed with and without preheating. For preheating studies, peptides were mixed in desired ratios, heated to 85 °C, and incubated for 15 min. The peptide solution was then slowly cooled to 25 °C at a rate of 1 °C/min and then incubated overnight at room temperature before performing the unfolding studies. For non-preheating studies, peptides were mixed in the desired ratios and the unfolding studies were performed immediately. The minimum of the derivative of the fraction folded plot indicates the steepest slope of the unfolding process and is used in this paper to indicate the melting temperature (T_m) under the conditions described above. This was calculated using the Jasco Spectra Manager software. Repeated experiments show that the T_m value for each system varies by ±1 °C or less.

Results

Homotrimers. POG, En, and Rn homotrimers, at a concentration of 0.2 mM, were characterized by CD thermal unfolding studies at neutral pH using 10 mM phosphate buffer, as shown in Figure 1.

As expected, POG forms the most stable triple helix, with a $T_m = 67.5$ °C which is consistent with the values reported for similar peptides.[49,63] The triple helix stability decreases as the number of substitutions at the X and Y positions with Glu and Arg, respectively, increase. In the case of E10 and R10 the triple helix is completely eliminated while more moderately charged peptides are found to have intermediate melting temperatures (see Table 2). As expected from previous mutation studies,[47] substitutions with glutamic acid are found to be more destabilizing than arginine. In the extreme example of E10 and R10, no triple helix formation was observed by CD studies (see Supporting Information Figure S3). CD spectra for R10 does not show any positive maximum but show a negative peak around 190 nm which resembles polyproline II spectra. E10 spectra show a blue-shifted positive peak at 218 nm, instead of the regular 225 nm, and the unfolding studies indicate a linear decrease in ellipticity at 218 nm, instead of the co-operative unfolding as observed for a normal triple helix. In summary, the CD spectra and the unfolding curve data for E10 and R10 show that these peptides do not form triple helices under the time frame of the experiments but instead form weak polyproline II helices or disordered structures.

AAB Heterotrimers. 1:1 mixtures of E5 with POG and R5 with POG were prepared so that the final peptide concentration of the solution was 0.2 mM, and a neutral pH was maintained using 10 mM phosphate buffer. Thermal unfolding studies were performed with and without preheating. Without preheating, two separate transitions were observed in the unfolding curve, corresponding to the unfolding of E5 and POG homotrimers with a T_m of 46.5 °C and 67.5 °C, respectively (Figure 2a−c). The two homotrimeric species remain individual triple helices and are well resolved. After preheating, however, only one transition is observed in the unfolding curve, corresponding to a T_m of 64 °C, which is slightly lower than the T_m for the POG homotrimer while significantly higher than the T_m for the E5 homotrimer (Figure 2b). Importantly, the maximum MRE of these samples are similar to one another indicating that the majority of the sample folds into a triple helix as opposed to one component (for example POG alone) folding into a helix while the other remains unfolded. A similar behavior is observed in the case of the R5/POG mixture, with the corresponding T_m

(67) Salem, G.; Traub, W. *FEBS Lett.* **1975**, *51*, 94−99.
(68) Chan, V. C.; Ramshaw, J. A. M.; Kirkpatrick, A.; Beck, K.; Brodsky, B. *J. Biol. Chem.* **1997**, *272*, 31441−31446.
(69) Venugopal, M. G.; Ramshaw, J. A. M.; Braswell, E.; Zhu, D.; Brodsky, B. *Biochemistry* **1994**, *33*, 7948−7956.

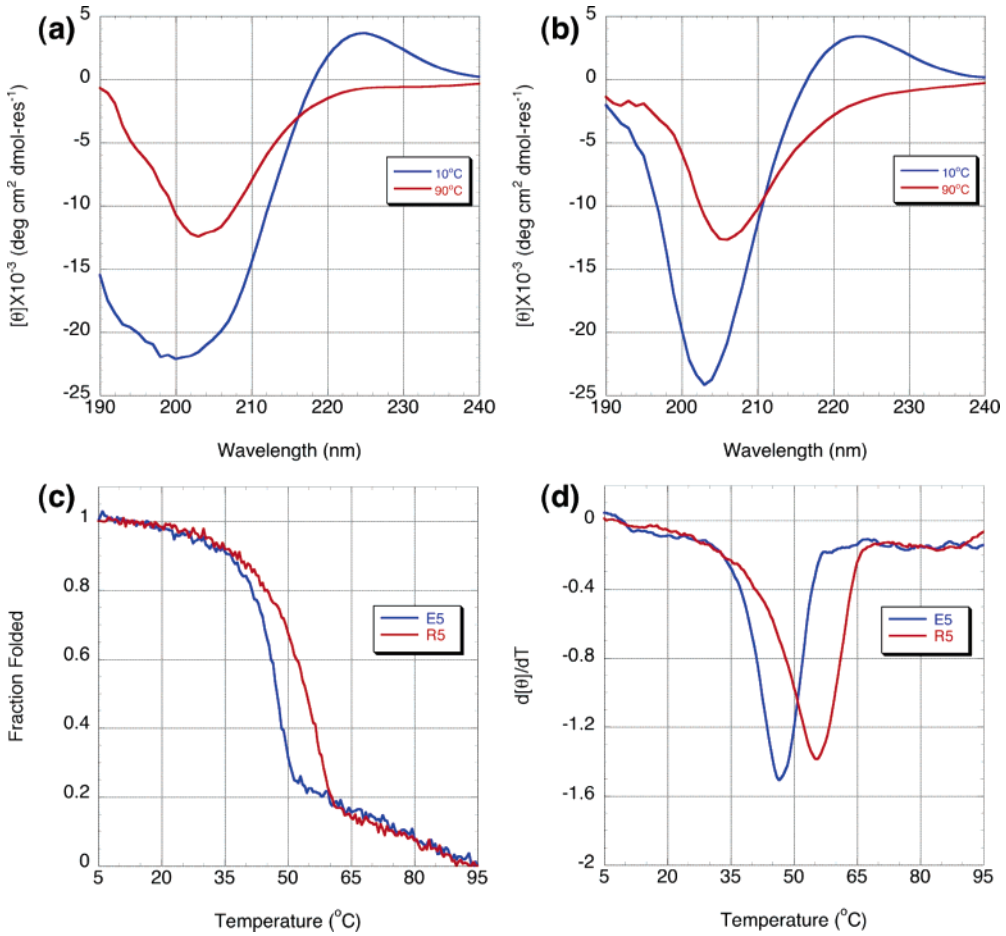

Figure 1. Circular dichroism for homotrimeric triple helices of E5 and R5. The spectra for both (a) E5 and (b) R5 go from a good poly-proline II helix at low temperature characterized by maxima near 224 nm and minima near 202 nm to a disordered structure at high temperature with only a single minima near 204 nm. (c) Thermal unfolding shows that the characteristic peak is eliminated in a cooperative transition which indicates the presence of a triple helix. (d) The first derivative of MRE is used to indicate a melting temperature, under the described conditions of 46.5 °C and 55.5 °C for E5 and R5, respectively. Equivalent data for POG, E3, R3, E10, and R10 are found in Supporting Information Figures S1−S3.

Table 2. Homotrimeric Melting Temperatures

homotrimer	melting temperature (T_m)
POG	67.5 ˚C
E3	51.5 ˚C
R3	58.5 ˚C
E5	46.5 ˚C
R5	55.5 ˚C
E10	no triple helix
R10	no triple helix

values of 56 °C and 67 °C for non-preheating and 65 °C after preheating (Figure 2d−f). This shows that heterotrimeric species are observed when samples are preheated and equilibrated as compared to the non-preheating case where well resolved individual homotrimers are observed.

In another set of experiments, En and Rn peptides were mixed to analyze the effect of electrostatics on heterotrimer formation. E3 was mixed with R3 in 2:1, 1:1, and 1:2 ratios, followed by preheating and unfolding studies (Figure 3a−c). T_m values of 56, 56.5, and 57 °C were observed for 2:1, 1:1, and 1:2 (E3/R3), respectively, as compared to 51.5 °C and 58.5 °C for E3 and R3 homotrimers, respectively. The E5/R5 series behaved in a similar fashion showing that preheating leads to the formation of a heterotrimeric species by allowing the system to equilibrate (Figure 3d−f). T_m values observed were 54, 54, and 55.5 °C for 2:1, 1:1, and 1:2 (E5/R5), respectively, with values of 46.5 °C and 55.5 °C for E5 and R5 homotrimers, respectively. The T_m for heterotrimers is observed to cluster near the T_m of the more stable Rn homotrimers, while maximum MRE is observed to be approximately the same in all cases (Figure 3b and e). Finally, heterotrimers of E10/R10 were prepared. Although neither E10 nor R10 form homomeric triple helices on their own (Supporting Information Figure S3), when they are mixed together and preheated they form a triple helix with a T_m of 41 °C (Figure 4). This triple helix is a heterotrimer composed of both positively charged R10 and negatively charged E10 polypeptides which interact favorably by electrostatic interactions. When the pH of the E10/R10 mixture is reduced to 3, the triple helical conformation is observed to be destroyed as indicated by the reduction of the maximum at 225 nm (Figure 4c).

Next, unfolding studies were performed on the E10/POG mix and R10/POG mix in a 1:1 ratio (Figure 5). The unfolding curve in both cases showed a single transition with a corresponding T_m of 68.5 °C and 68 °C for the E10/POG mix and R10/POG mix, respectively. POG homotrimers show a nearly identical T_m value of 67.5 °C (Figure 5b). Importantly, the MRE for these mixtures drops to approximately one-half of that observed for POG alone (Figure 5a). Furthermore it was observed that there was essentially no difference between the CD spectra of E10/

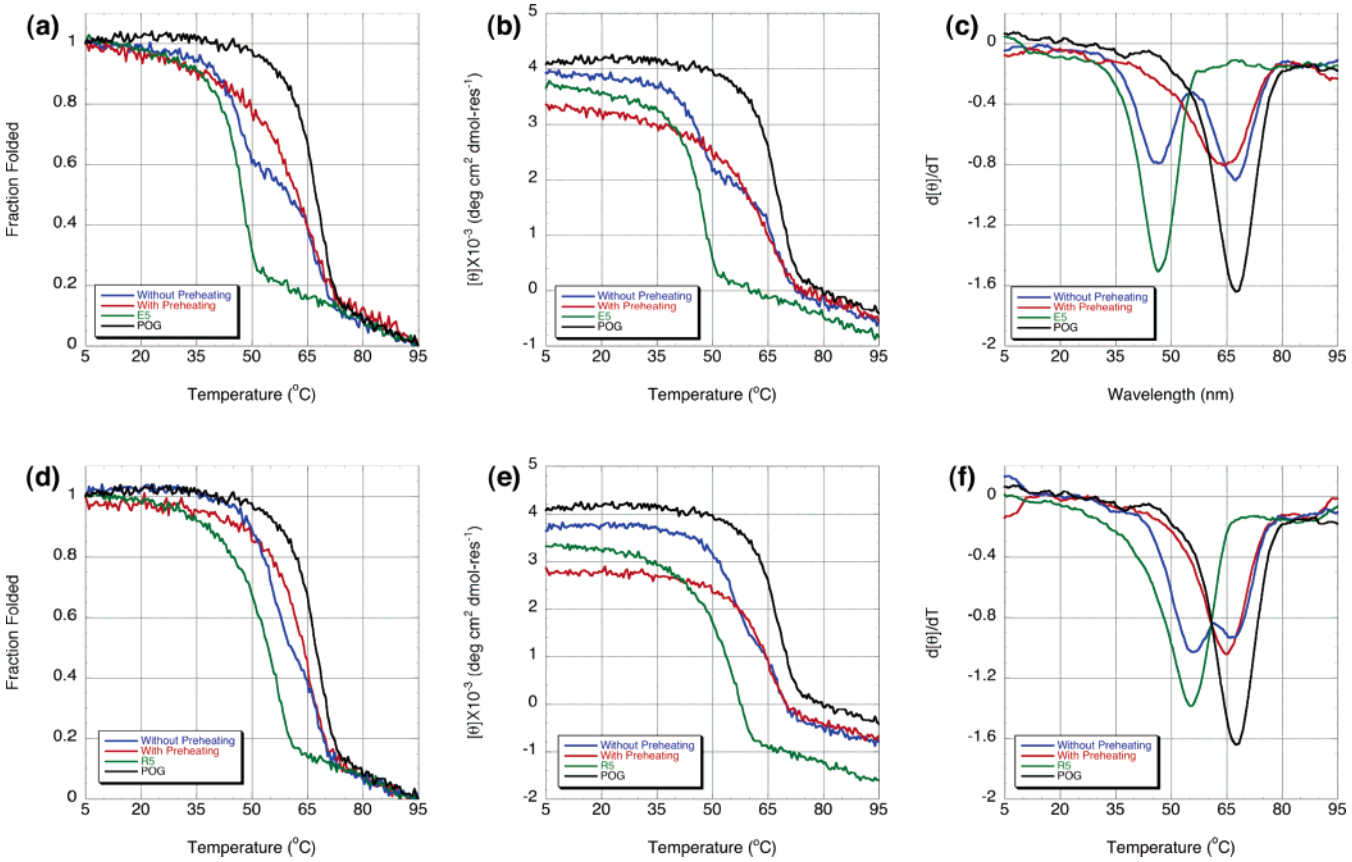

Figure 2. Circular dichroism of charged peptides mixed with neutral peptide. (a) Thermal unfolding shows cooperative transition for POG, E5, and their mixture. (b) Mean residue ellipticity (MRE) versus temperature for POG, E5, and their mixture. (c) The first derivative of unfolding clearly shows the individual homotrimeric species remain when preheating is not applied while these merge to a single heterotrimeric species after heating. (d) Thermal unfolding shows cooperative transition for POG, R5, and their mixture. (e) MRE versus temperature for POG, R5, and their mixture. (f) The first derivative of unfolding clearly shows the individual homotrimeric species remain when preheating is not applied while these merge to a single heterotrimeric species after heating.

POG and R10/POG before or after preheating (Figure 5c). Together with the T_m value being approximately equal to that for POG alone this indicates that no heterotrimeric species are formed in either the E10/POG or R10/POG cases.

ABC Heterotrimers. When E10, R10, and POG are mixed in a 1:1:1 ratio, a T_m of 54 °C is observed after preheating (Figure 6). This melting temperature corresponds to an ABC triple helix composed of all three polypeptides. Again maximal MRE in the case of E10/R10/POG indicates that the majority of the sample has folded rather than one homotrimer (for example, POG) leaving the other peptides unfolded (Figure 6b). The E10/R10/POG mixture was further analyzed by comparing the non-preheated and preheated mixtures. Without preheating, three transitions were observed in the unfolding curve, corresponding to a POG homotrimer (68 °C), an E10/R10 heterotrimer (38.5 °C), and an improperly folded E10/R10 heterotrimer (19 °C, see Discussion below). On preheating, all of these transitions merged to a single T_m of 54 °C corresponding to an ABC type E10/R10/POG heterotrimer.

Discussion

The thermal stability of homotrimers, shown in Table 2, compares favorably with the previous literature data: POG forms the most stable helix with a melting temperature of 67.5 °C; substitution with three arginines reduces the melting point to 58.5 °C, while five arginines reduce the melting point to 55.5 °C. The use of 10 arginines destroys the ability of the peptide to form a triple helix within the time frames examined (as long as 1 week; data not shown). Glutamic acid substitution shows the same pattern but with greater destabilization per substitution. Using these results concerning the homotrimers, we were able to assess heterotrimeric helix assembly and stability. Our hypothesis was that pairing glutamic acid in the X position with arginine in the Y position would lead to the formation of a stabilized heterotrimer through charge paired hydrogen bonding. In homotrimeric systems, such an interaction has been shown to be possible if the charged residues are present in consecutive X and Y positions.[46,67,68] Instead of incorporating both the residues in one polypeptide, we incorporated them in two different polypeptides in such a way that these two residues can interact favorably only upon the formation of a heterotrimer. Mixtures of E3 with R3 appear to confirm this with the observation of a species that melts at a temperature intermediate to that of either E3 or R3 alone. However, the melting temperatures of E3, R3, and E3/R3 are close to one another (51.5, 58.5, and 56.5 °C, respectively) making analysis difficult. To help clarify the situation, E5 and R5 were prepared which have a wider separation of melting points. The mixture shows a melting point of 54.0 °C, surprisingly close to that for R5 alone (55.5 °C). Mixtures of E5/POG and R5/POG also show melting points surprisingly close to that for POG alone and are

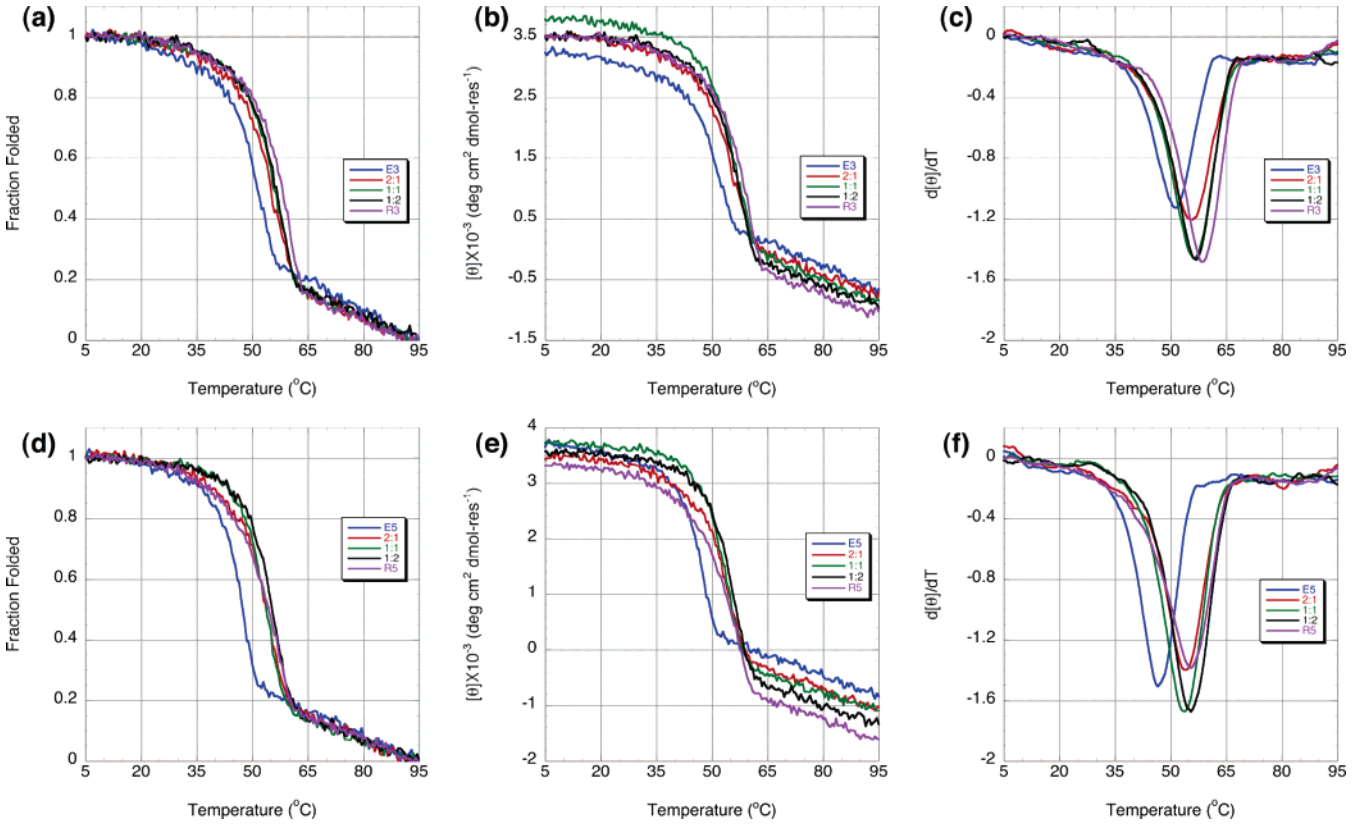

Figure 3. Thermal denaturation studies on multiply charged heterotrimeric helices showing the fraction folded, MRE, and first derivative plots versus temperature. (a) Fraction folded of E3 with R3. (b) MRE of E3 with R3. (c) First derivative of unfolding for E3 and R3 mixture. (d) Fraction folded of E5 with R5. (e) MRE of E5 with R5. (f) First derivative of unfolding for E5 and R5 mixture.

Table 3. Melting Temperatures of Heterotrimeric Helices[a]

peptide/peptide mix	melting temperature (T_m)
E3/R3 = 2:1	56 ˚C
E3/R3 = 1:1	56.5 ˚C
E3/R3 = 1:2	57 ˚C
E5/POG = 1:1*	46.5 and 67.5 ˚C
E5/POG = = 1:1	64 ˚C
R5/POG = 1:1*	55 and 67 ˚C
R5/POG = 1:1	65 ˚C
E5/R5 = 2:1	54 ˚C
E5/R5 = 1:1	54 ˚C
E5/R5 = 1:2	55.5 ˚C
E10/R10 = 1:1	41 ˚C
E10/R10/POG = 1:1:1*	19, 38.5, and 68 ˚C
E10/R10/POG = 1:1:1	54 ˚C

[a] Peptides were preheated to ensure equilibrium unless marked with an asterisk.

almost indistinguishable from one another. One possible interpretation of this is that the more stable peptide forms a homomeric triple helix, while the less stable peptide remains unfolded. This can be ruled out by examining the MRE values of the various mixtures which do not dramatically change depending on components being mixed (Figures 2b, 2e, 3b, and 3e).

In recent studies[62,63] performed on $(PPG)_{10}/(POG)_{10}$ mixtures (abbreviated here as PPG and POG), it was observed that the T_m value increased almost linearly in the order 3•PPG (44.9 °C), 2•PPG/POG(54.8 °C), PPG/2•POG (63.6 °C), and 3•POG (68.4 °C).[62] Our studies indicate that the T_m value increases *nonlinear*ly as the fraction of the more stable homotrimer is increased in the mixture (Table 3) with T_m values for all the mixtures clustered very close to the T_m for the more stable homotrimeric helix. This behavior can be attributed to the differences in electrostatic interactions between the oppositely charged monomers or between charged and neutral monomers, which are possible only in heterotrimers. In the extreme example of E10 and R10, no homotrimeric helices are formed in the time frame of the experiment. However, mixing the two peptides results in the formation of a triple helix with a melting temperature of 41.0 °C. Adjusting the pH of the solution to 3 to eliminate the charge on E10 also eliminates possible charge pairing and helix formation (Figure 4c).

Two factors must be considered while analyzing the formation of these heterotrimers: inherent helix propensity of the given amino acids and electrostatic interactions between them. It has been shown that P and O are preferred over E and R, respectively, and this predisposition dominates helix formation in most cases. Pairing E with R leads to the elimination of the repulsive like−like charge and incorporation of attractive opposite charge interactions. However, there is a competition between the destabilization caused by the substitutions and the stability incurred by the attractive interactions, and the final stability depends on the relative magnitude of both. Therefore the peptide E5 or R5 can be stabilized by mixing with POG which eliminates repulsive interstrand interactions or by mixing with one another where the interstrand interactions are favorable.

Mixing POG with E10 and R10 can theoretically result in the formation of seven types of triple helixes[70] including POG, E10, and R10 AAA homotrimers; E10/R10, E10/POG, and R10/POG AAB heterotrimers; and an E10/R10/POG ABC hetero-

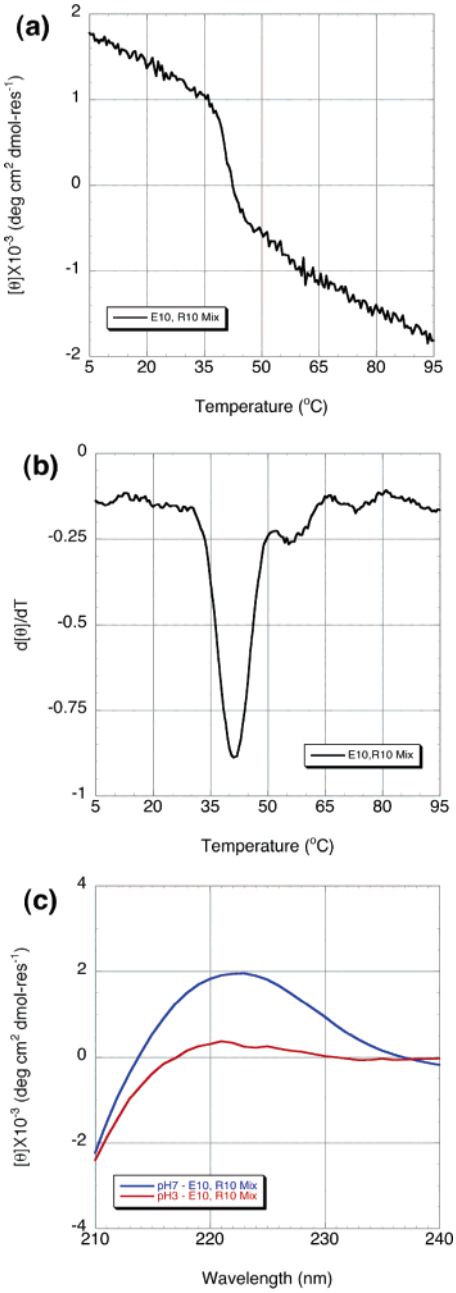

Figure 4. CD analysis of E10/R10 mixtures. (a) MRE versus temperature, (b) first derivative of MRE versus temperature, (c) MRE versus wavelength at pH 7 and pH 3.

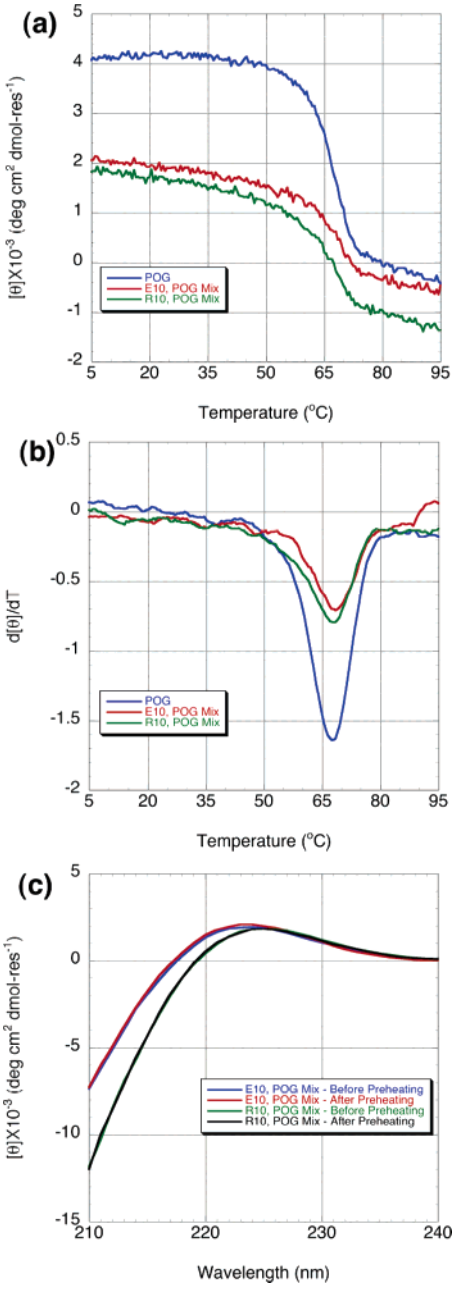

Figure 5. CD analysis of mixtures E10/POG and R10/POG. (a) MRE vs temperature, (b) first derivative of MRE vs temperature, (c) MRE vs wavelength.

trimer. E10 and R10 homotrimers can be excluded from consideration by studies on these species in isolation (Supporting Information Figure S3) which show that they do not form helices. Likewise, unfolding studies on E10/POG and R10/POG mixtures show that neither E10 nor R10 interacts with POG (Figure 5). Instead unfolding curves show a transition corresponding to POG homotrimers in both mixtures. Additionally, the MRE value is approximately half of the observed value for pure POG indicating the POG homotrimeric helix is forming and being observed at half the normal concentration, while the charged peptide (either E10 or R10) remains unfolded. The other species (POG AAA homotrimer, E10/R10 AAB heterotrimer, and E10/R10/POG ABC heterotrimer) may be possible. A 1:1:1 mixture of these peptides without preheating results in the observation of three distinct transitions at 68, 38.5, and 19 °C (Figure 6c). These can be attributed to a POG homotrimer, an E10/R10 heterotrimer, and a much less stable component which cannot yet be identified but may be an out of register or misaligned triple helix. Preheating merges these three transitions into one new transition with T_m = 54.0 °C. This can only be attributed to the formation of an ABC heterotrimer. This ABC heterotrimer is significantly more stable than the E10/R10 heterotrimer because addition of a POG component allows the complete triple helix to have an overall neutral charge. It also has an MRE value almost double that observed for AAB

(70) The definition of seven triple helical types groups together with heterotrimers which are composed by the same units but in different ratios. For example, AAB and ABB heterotrimers are grouped together in this analysis.

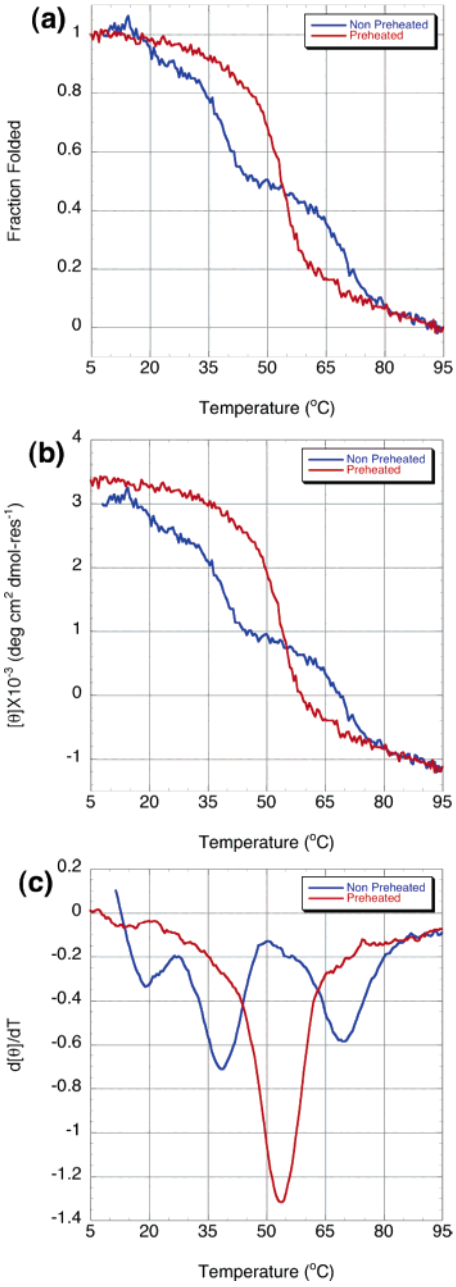

Figure 6. Thermal denaturation studies on preheated and non-preheated E10/R10/POG mixtures at pH 7. (a) Fraction folded versus temperature, (b) MRE versus temperature, (c) first derivative of MRE vs temperature.

heterotrimers in the E10/R10 mixture which indicates that nearly all of the peptide in the solution forms the ABC helix.

Perhaps surprisingly, a POG homotrimer is not observed in these mixtures despite it being the most stable individual triple helix. However, if the POG homotrimer is to form, it must also lead to formation of the less stable E10/R10 AAB type heterotrimers as our studies show that E10/POG and R10/POG heterotrimers and E10 and R10 homotrimers do not form. The preferred formation of the ABC type E10/R10/POG heterotrimer over a mixture of the POG homotrimer and the AAB type E10/R10 heterotrimer suggests that ABC type heterotrimer formation leads to greater systemwide energy stabilization even though POG homotrimer is more stable in isolation. These results suggest that triple helix stability is not simply due to an amino acid's propensity to adopt a particular conformation but a convolution of amino acid conformational propensity, electrostatic attraction or repulsion, and cross strand pairing of amino acids. This allows highly stable motifs such as POG to significantly stabilize poor triple helix forming sequences such as EOG far beyond what one would expect from simply diluting out the bad amino acid. Similarly, oppositely charged amino acids E and R can form heterotrimers with a stability comparable to that for R alone. We observe an effect in which a weak triple helix forming peptide can be "rescued" by mixing it with a stronger triple helix forming peptide. The extent of rescue depends on the individual polypeptides and their ability to interact with each other by favorable electrostatic interactions. Heterotrimer formation therefore allows the introduction of greater chemical diversity in the collagen triple helix, as amino acids with weak individual propensities for triple helix formation can be paired with strong ones.

Conclusions

In this paper we have designed, synthesized, and characterized a series of seven peptides which utilize charged amino acids and the ubiquitous X−Y−Gly sequence to bias their self-assembly into collagen-like heterotrimeric helices. In addition to five novel AAB heterotrimers, we demonstrate for the first time the formation of an ABC heterotrimer utilizing only supramolecular interactions. Melting studies show that amino acids which substantially destabilize the triple helix, such as glutamic acid, can form triple helices of high stability when paired against oppositely charged amino acids such as arginine or, more surprisingly, when paired against the neutral amino acid hydroxyproline. This results in the stability of the heterotrimeric helix being almost indistinguishable from the stability of those containing only amino acids with high triple helix stability. Heterotrimer formation therefore allows the introduction of greater chemical diversity in the collagen triple helix as amino acids with weak individual propensities for triple helix formation can be paired with, and rescued by, strong ones. These results lay the groundwork for a better understanding of the thermal stability of heterotrimeric collagens, for the synthesis of synthetic extracellular matrix mimics[64−66] and molecular recognition devices based on these interactions.

Acknowledgment. J.D.H. gratefully acknowledges the Kinship Foundation for his Searle Scholar award. This work was funded in part by the Robert A. Welch Foundation Research Grant Number C1557 and The Alliance for Nanohealth.

Supporting Information Available: MALDI-MS data, HPLC and additional CD analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0683640