pubs.acs.org/Biomac

Article

Charge-Free, Stabilizing Amide−π Interactions Can Be Used to Control Collagen Triple-Helix Self-Assembly

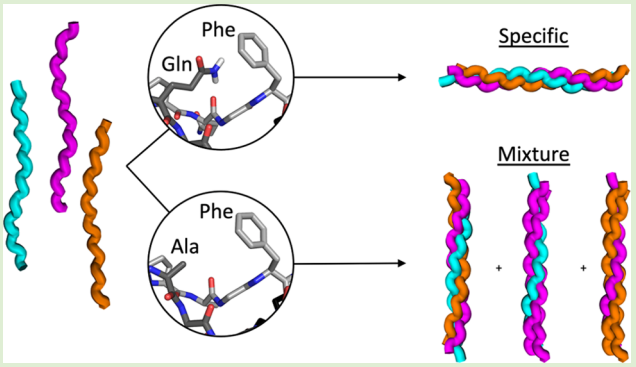
Douglas R. Walker, Ali A. Alizadehmojarad, Anatoly B. Kolomeisky, and Jeffrey D. Hartgerink*

Cite This: *Biomacromolecules* 2021, 22, 2137−2147

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: There is a noted lack of understood, controllable interactions for directing the organization of collagen triple helices. While the field has had success using charge−pair interactions and cation−π interactions in helix design, these alone are not adequate for achieving the degree of specificity desirable for these supramolecular structures. Furthermore, because of the reliance on electrostatic interactions, designed heterotrimeric systems have been heavily charged, a property undesirable in some applications. Amide−π interactions are a comparatively understudied class of charge-free interactions, which could potentially be harnessed for triple-helix design. Herein, we propose, validate, and utilize pairwise amino acid amide−π interactions in collagen triple-helix design. Glutamine−phenylalanine pairs, when arranged in an axial geometry, are found to exhibit a moderately stabilizing effect, while in the lateral geometry, this pair is destabilizing. Together this allows glutamine−phenylalanine pairs to effectively set the register of triple helices. In contrast, interactions between asparagine and phenylalanine appear to have little effect on triple-helical stability. After deconvoluting the contributions of these amino acids to triple-helix stability, we demonstrate these new glutamine−phenylalanine interactions in the successful design of a heterotrimeric triple helix. The results of all of these analyses are used to update our collagen triple-helix thermal stability prediction algorithm, Scoring function for Collagen Emulating Peptides' Temperature of Transition (SCEPTTr).



INTRODUCTION

Collagen mimetic peptides (CMPs) have long been used as proxy for studying natural collagen sequences. Natural collagen is unwieldy, long, insoluble, and cross-linked. In contrast, CMPs are adaptable, short, soluble, and easily prepared. CMPs have been used to study collagen disease states, collagen-binding domains, and collagen folding and assembly.[1−11] To a limited extent, CMPs have also been used in designing supramolecular nanomaterials.[12−16] The unique secondary structure of collagens, as compared to α-helices and β-sheets, serves to increase the variety of structural motifs available to materials design using only peptides as building blocks. However, both biological and materials applications have been limited due to the narrow selection of predictable interactions for controlling the folded arrangement of triple helices.

Collagens adhere to an (Xaa−Yaa−Gly) triplet repeat with a high propensity for proline in the Xaa position and hydroxyproline in the Yaa position that encourages a polyproline type II (PPII) secondary structure and a triple-helical tertiary structure (Figure 1). Each peptide strand adopts a staggered register offset by one amino acid from the next strand in the triple helix to allow the glycine residues to reside in the core of the triple helix and facilitate the packing of the strands. Beyond these constraints, there are very few known rules for predictable triple-helix design in contrast to the extensive design criteria known for α-helices and β-sheets.[17−25] This is due in part to the unique secondary and tertiary structures of collagen: unlike most peptides and proteins, nearly all of the interactions in collagen are intermolecular. The consequences of this fact are that, in contrast to intramolecular folding such as in α-helices where the sequential distance of two amino acids cannot change (for example, an i, i + 1 relation cannot become an i, i + 4), in collagens, the interaction geometries can be much more dynamic. As we described previously,[26] interactions depend on the triplet location of the two amino acids involved rather than their sequential relationship. For example, it is known that positively charged amino acids in Yaa positions will preferentially form intermediate to strong complementary interactions with negatively charged or aromatic amino acids in Xaa positions, a neighboring chain in the succeeding triplet (axial

Received: February 21, 2021
Revised: March 29, 2021
Published: April 21, 2021

https://doi.org/10.1021/acs.biomac.1c00234
Biomacromolecules 2021, 22, 2137−2147

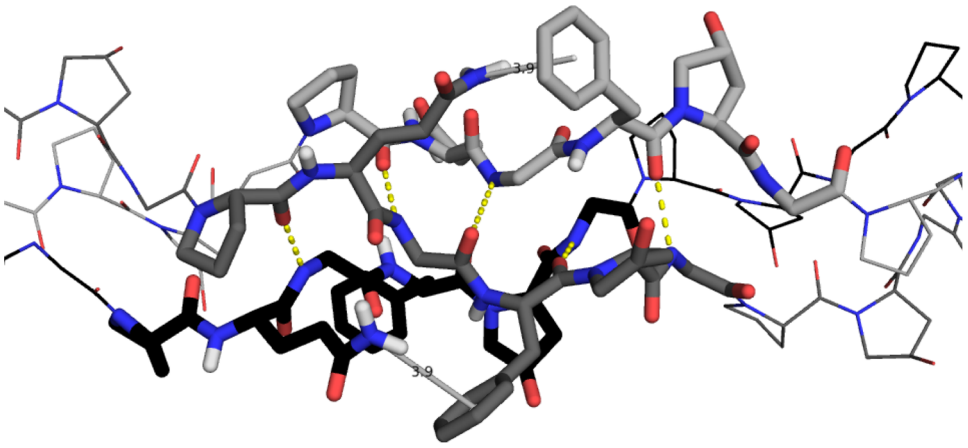


Figure 1. Collagen triple helix highlighting the three different strands (leading in black, middle in dark gray, trailing in light gray), the intermolecular hydrogen bonds (yellow dotted lines), and the staggered nature of collagen. Potential orientation of a Gln−Phe axial interaction in collagen. The distance from nitrogen to the ring center as shown is 3.9 Å.

interactions),[27,28] as illustrated by the gray bars in Figure 1. It is also known that negatively charged amino acids in Yaa positions can form weak to intermediate complementary interactions with positively charged amino acids in Xaa positions on neighboring chains in the same triplet (lateral interactions).[29]

In addition to charge−pair interactions and cation−π interactions, amide−π interactions play an important role in structural biology. In 1936, Wulf et al. suggested, for the first time, that aromatic systems could accept hydrogen bonds.[30] But it took 50 years before Burley and Petsko[31] published an analysis of 33 protein crystal structures highlighting the fact that side-chain amino groups interact with aromatic side chains in folded proteins, indicating an important role for these interactions in protein stability and/or function. This analysis included lysine and arginine, but also glutamine, asparagine, and histidine. Two years later, Levitt and Perutz[32] proposed, through energetic calculations of an N−H interaction with benzene, that this interaction is a hydrogen bond with an energy minimum distance of 2.9−3.6 Å between the nitrogen and the ring center, the N−H situated directly over the center of the benzene ring, and the N−H bond perpendicular to the benzene ring. These calculations were supported by Rodham et al.[33] in 1993 through gas-phase optical and microwave spectroscopy. Since these reports, there have been many subsequent energetic computational,[34−51] crystallographic computational,[35−40,52−60] and gas-phase spectroscopic[41−45] analyses of these interactions, but fewer analyses in experimental aqueous conditions.[40,61−63] These reports indicate that the N−H−π interaction ranges from ∼3 to 5 Å and is sufficiently strong to serve an important role in stabilizing protein and peptide structures. However, of the four accounts that report aqueous analysis, none utilizes these interactions to design new systems that benefit from the included amide−π interactions.

In our previous work, we reported on the development of our algorithm "Scoring function for Collagen Emulating Peptides' Temperature of Transition" (SCEPTTr) for predicting the thermal stability and registration of combinations of CMPs.[64] SCEPTTr utilizes the effects of single-amino-acid substitutions and pairwise interactions in addition to several other terms to predict the melting temperature of a given triple helix. SCEPTTr can analyze both homotrimers and heterotrimers, considers both canonical and noncanonical registrations, and analyzes all compositions and registers of a combination of CMPs before predicting the most stable conformation. This algorithm reproduces the experimental results of over 400 published triple helices. The parameters used to fit this data set suggested that glutamine/asparagine−phenylalanine axial interactions might be useful to control the registration of collagen triple helices, interactions not previously investigated in collagen. In our current work, we follow up on that indication, investigating the effects of both axial and lateral interactions of glutamine−phenylalanine (Q−F) (shown in Figures 1 and 6) and asparagine−phenylalanine (N−F) pairs. We find that the axial Q−F interaction is moderately stabilizing and the lateral Q−F interaction is destabilizing. In contrast, neither axial nor lateral N−F interaction significantly perturbs the triple-helix stability. This registrationally specific Q−F interaction is then used to design a new heterotrimer to illustrate the utility of the interaction. Finally, the twelve new triple helices reported here are used to further train SCEPTTr to produce SCEPTTr1.1. We expect that the new Q−F interactions and the improved SCEPTTr reported here will be useful for the field for designing new CMP systems, new materials, and for understanding sequences in natural collagens.

EXPERIMENTAL SECTION

Peptide Synthesis. Solid-phase peptide synthesis with Fmoc-protecting strategy on a Rink Amide MBHA resin was used for all peptides using the methodology previously described.[65] In peptide D (see Table 2), the glycine at position 18 was labeled with ^{15}N. Peptides were purified by cold ether wash, trituration, reversed-phase high-performance liquid chromatography (HPLC), and finally lyophilization. All peptides were N-terminally acetylated and C-terminally amidated. Correct synthesis was confirmed by matrix-assisted laser desorption ionization (MALDI) mass spectrometry and ultra-performance liquid chromatography (UPLC) (available in the Supporting Information).

Sample Preparation. Peptide solutions were prepared at 3 mM peptide, 10 mM phosphate buffer, and pH 7. For the mixed samples used for heterotrimeric formation and analysis, the peptides were mixed at an equimolar concentration to a total peptide concentration of 3 mM. All 3 mM solutions were heated to 85 °C for 15 min, allowed to cool to room temperature, and subsequently refrigerated at 4 °C for a minimum of 12 h before data collection. Directly prior to circular dichroism (CD) analysis, 200 μL of 0.3 mM samples was

https://doi.org/10.1021/acs.biomac.1c00234
Biomacromolecules 2021, 22, 2137−2147

prepped from the 3 mM sample by dilution into Milli-Q (mQ) water. Directly prior to nuclear magnetic resonance (NMR) analysis, 450 μL of the 3 mM sample was transferred to an NMR tube with 50 μL of D_2O and a small amount of 3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid sodium salt.

Circular Dichroism Spectroscopy. A Jasco J-810 spectropolarimeter equipped with a Peltier temperature controlled was employed to collect all CD spectra and thermal unfolding traces. The limit of the wavelength measurements used were 190 nm and 250 nm, and melting analyses were recorded using the PPII maximum at 225 nm. An intensity of 220 nm was monitored from 5 to 85 °C at a heating rate of 10 °C/h. The derivative of the unfolding curves was calculated using the Savitzky−Golay algorithm and the melting temperature is defined here as the temperature at which the derivative reaches a local minimum.

NMR Spectroscopy. All NMR experiments were collected on a Bruker Avance III HD 800 MHz spectrometer with a QCI CryoProbe and were run with samples incubating at 25 °C. Data collection and processing were performed using methods as previously described.[64] Briefly, all experiments were run with the first ^{1}H dimension transmitter frequency offset set to 4.81 ppm with a spectral width of 12 ppm. In the heteronuclear single quantum coherence (HSQC) experiments, the ^{15}N dimension transmitter frequency offset was set to 112 ppm with a spectral width of 25 ppm. Each HSQC was collected using 8 scans in 1922 transients in the ^{1}H dimension and 256 transients in the ^{15}N dimension after 32 dummy scans. In the three-dimensional (3-D) ^{1}H−^{1}H−^{15}N nuclear Overhauser effect spectroscopy (NOESY) HSQC experiment, the ^{15}N dimension transmitter frequency offset was set to 108 ppm with a spectral width of 10 ppm and for the second ^{1}H dimension the spectral width was set to 8 ppm. Sixty-four dummy scans were run followed by four scans in 2048 transients in the first ^{1}H dimension, 32 in the ^{15}N dimension, and 256 in the second ^{1}H dimension.

Molecular Dynamics. All-atom molecular dynamics (MD) simulations were utilized for a detailed understanding of side-chain interaction effects on the stability of the single- and double-substituted peptides. The initial structure of the homotrimer collagen was obtained from the crystal structure of $(GPO)_9$ with Protein Data Bank (PDB) id: 3B0S.[66] We then focused on the central portion of the triple helix and applied the mutations of interest for each structure. In the case of single-substituted peptides, we obtained initial structures of **OGFO** and **QGPO** by replacing 14th and 12th residues of the $(GPO)_9$ structure, respectively. To build **QGFO**, a double-substituted peptide, we replaced 12th and 14th residues of the $(GPO)_9$ structure with glutamine and phenylalanine, respectively. The **OGFQ** homotrimer structure was constructed by changing 14th and 15th residues of $(GPO)_9$ collagen to phenylalanine and glutamine amino acids, respectively. We placed each homotrimer in the center of a cubic box and solvated using the Solvate VMD plugin. The systems were totally composed of at least 167 000 atoms in a cubic box with a side length of 12 nm, where we used the TIP3P water model. The initial condition of a representative prepared system is shown in Figure S17.

NAMD 2.13 package[67] was used to carry out MD simulations using both CHARMM36[68] and AMBER-ffSB14[69] force fields. According to the experimental conditions, pressure and temperature were, respectively, kept at 1 bar and 300 K using Langevin dynamics in the *NPT* ensemble. Periodic boundary conditions were applied in all directions, and the particle mesh Ewald (PME) method was utilized to evaluate the electrostatic potentials.[70] Energy minimization was run for 5000 steps, and then the systems were warmed up to the room temperature before 100 ns production run. A time step of 2 fs was chosen in all simulations and snapshots of the systems were saved every 20 ps. Preparation of the systems and data analysis tasks were done using VMD.[71]

RESULTS AND DISCUSSION

Homotrimer Analysis. We synthesized the peptides shown in Table 1 to test the amide−π interactions between glutamine/asparagine and phenylalanine. Circular dichroism was used to measure the melting temperatures (Figure 2) of each of the homotrimers, which were then used to calculate the degree of stabilization of both the axial and lateral geometries of the interactions. The reference triple helix, which consists of a homotrimeric arrangement of three $(POG)_8$ peptides (here abbreviated as **OGPO**, see Table 1 below), unfolds at 50.0 °C. In comparison, the homotrimers of single-amino-acid substitutes in the **OGFO** peptides melt at 37.0 °C and the homotrimer of the **QGPO** peptides melts at 45.0 °C.

Table 1. Homotrimeric Peptides Used for the Amide−π Interaction Deconvolution[b]

Abbrev.	Sequence	T_m (°C)	Δ Expected T_m
OGPO	POGPOGPOGPOGPOGPOGPOGPOG	**50.0[a]**	-
OGFO	POGPOGPOGPOGFOGPOGPOGPOG	**37.0[a]**	-
AGPO	POGPOGPOGPAGPOGPOGPOGPOG	**44.0[a]**	-
QGPO	POGPOGPOGPQGPOGPOGPOGPOG	**45.0**	-
NGPO	POGPOGPOGPNGPOGPOGPOGPOG	**34.0**	-
OGFA	POGPOGPOGPOGFAGPOGPOGPOG	**22.5**	**-8.5**
AGFO	POGPOGPOGPAGFOGPOGPOGPOG	**29.0[a]**	**-2.0**
OGFQ	POGPOGPOGPOGFQGPOGPOGPOG	**24.5**	**-7.5**
QGFO	POGPOGPOGPQGFOGPOGPOGPOG	**32.5**	**+0.5**
OGFN	POGPOGPOGPOGFNGPOGPOGPOG	**19.5**	**-1.5**
NGFO	POGPOGPOGPNGFOGPOGPOGPOG	**21.5**	**+0.5**

[a]Previously published.[64] [b]T_m is the experimental melting temperature. Expected T_m values for double-substituted peptides are calculated by subtracting the destabilization of each single amino acid from the basis triple helix and assumes no pairwise interactions exist. ΔExpected T_m indicates the deviation from the expected melting temperature. The gray highlighted amino acids signify the region that is altered for each different peptide, and consequently, the letters that are used to name each peptide.

The reference triple helix, which consists of a homotrimeric arrangement of three $(POG)_8$ peptides (here abbreviated as **OGPO**, see Table 1 below), unfolds at 50.0 °C. In comparison, the homotrimers of single-amino-acid substitutes in the **OGFO** peptides melt at 37.0 °C and the homotrimer of the **QGPO** peptides melts at 45.0 °C.

Because each amino acid present in a collagen mimetic peptide is present three times in a homotrimer formed by that peptide, these three melting temperatures show that a phenylalanine in the Xaa position of a collagen triplet destabilizes a triple helix by 4.3 °C (calculated from (50.0 − 37.0)/3) and a glutamine in the Yaa position of a collagen triplet destabilizes a triple helix by 1.7 °C (calculated from (50.0 − 45.0)/3). Assuming no pairwise interactions, this suggests that a homotrimer of peptides with a phenylalanine in an Xaa position and a glutamine in a Yaa position should form a triple helix destabilized by 18.0 °C and is therefore expected to melt at 32.0 °C.

Peptides with two adjacent Xaa, Yaa position amino acids mutated from the reference peptide $(POG)_n$ have the potential to form intermolecular pairwise amino acid interactions when they fold into a triple helix. Specifically, two lateral interactions can form, one from leading (L) Yaa to middle (M) Xaa and one from middle Yaa to trailing (T) Xaa (see Figure S27). We show (Table 1) that the homotrimer of the **OGFQ** peptides melts at 24.5 °C, much lower than predicted without considering pairwise interactions. This difference can be attributed to two lateral interactions, where each destabilizes the triple helix by 3.75 °C.

Peptides with Yaa, Xaa position amino acids separated by one glycine and mutated from the reference peptide $(POG)_n$ also have the potential to form intermolecular pairwise interactions when they fold into a triple helix. Specifically, two axial and one lateral interactions can form: one axial from leading Yaa to middle Xaa, one axial from middle Yaa to trailing Xaa, and the lateral from trailing Yaa to leading Xaa (see Figure S27). The homotrimer of the **QGFO** peptides, which melts at 32.5 °C, is used to elucidate the contribution of

https://doi.org/10.1021/acs.biomac.1c00234
Biomacromolecules 2021, 22, 2137−2147

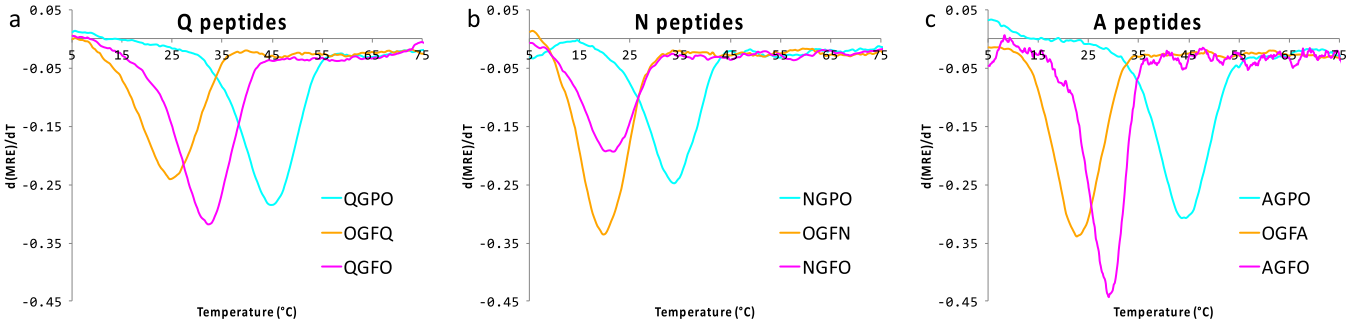


Figure 2. Triple-helix melting as indicated by the first derivative of CD for each of the homotrimers synthesized for this study. (a) Triple helices containing glutamine. (b) Triple helices containing asparagine. (c) Triple helices containing alanine. AGPO and AGFO as previously published.[64]

the axial Q−F interaction. This homotrimer contains one lateral and two axial interactions. Using the above value for the lateral contribution, we can determine that each axial interaction stabilizes the triple helix by 2.13 °C. Repeating this analysis for asparagine-containing homotrimers reveals lateral interactions worth −0.75 °C and axial worth +0.63 °C, neither of which is considered significant when compared to the precision of our method.

This method of analysis of single- and double-substitution derivatives of the basis triple-helix **OGPO** allows us to isolate the pairwise interactions found in both the lateral and axial spatial relationship to the exclusion of what each amino acid contributes toward the stability of the helix alone. This demonstrates that while Q−F axial pairwise interactions stabilize the triple helix by just over 2 °C, Q−F lateral pairwise interactions are strongly destabilizing and both axial and lateral N−F pairwise interactions have a minimal impact on helix stability.

The alanine containing homotrimers, **AGFO** and **OGFA**, melting temperatures indicate a significant lateral destabilization worth −4.25 °C per interaction and a mildly stabilizing axial interaction of +1.13 °C per interaction. These values are surprisingly close to the lateral and axial interactions found in glutamine−phenylalanine pairs at −3.75 and +2.13 °C per interaction, respectively. The similarly strong destabilization between the Q−F and A−F lateral interactions suggests a similar mode of action between the two. Because of the lack of possible strong intermolecular interactions for the alanine side chain, it is unlikely that the destabilizing mode of action is due to side-chain interactions, but likely instead is due to compounding backbone conformational perturbations. In the case of **OGXY**-patterned homotrimers, each peptide contains a contiguous run of four noncyclic amino acids. Because the most stable Xaa and Yaa amino acids in collagen are proline variants, it is reasonable to consider that multiple successive deviations from that pattern may have a compounding effect on the stability of a triple helix. These may be better thought of as the lack of stabilizing proline and hydroxyproline residues rather than actual destabilizing pairwise interactions. Thus, we hypothesize that the lateral destabilization of both Q−F combination and the A−F combination is due in large part to this kind of compounding backbone structural changes more so than to side-chain interactions.

Heterotrimer Analysis. The combination of destabilizing Q−F lateral interactions and stabilizing Q−F axial interactions as found in our homotrimer analysis suggests that the glutamine−phenylalanine interactions will be useful for controlling the register of heterotrimeric triple helices. To test this hypothesis, we synthesized peptide D shown in Table 2, which has been adapted from peptide C of the ABC heterotrimer previously studied by us.[64] Peptide D is designed to form a triple helix with peptides A and B, which employs four Q−F axial interactions. Tables 2 and 3 show the melting temperatures of the various combinations of the three designed peptides A, B, and D. As predicted, the ABD mixture forms the most stable triple helix and melts at 33.5 °C (Figure 3b). This system has specificity (difference in melting temperature between the most stable and second most stable triple-helical assembly) of 12.5 °C with the most stable competitor being the heterotrimer of the AB peptide mixture, which melts at 21.0 °C.

Table 2. Peptides Used to Assemble the Heterotrimers and the Expected Registration of the ABD Triple Helix[b]

Abbrev.	Sequence	Tm (°C)
A	PKGPKGFOGPOGFKGFKGPKGPOGFKGPOG	NT[a]
B	PKGDOGDKGPOGPPGDKGDOGDKGPKGDOG	18.5[a]
C	PRGEPGPRGERGPPGPPGERGPPGEPGEPG	16.0[a]
D	PQGEPGPQGEQGPPGPPGEQGPPGEPGEPG	NT
E	PAGEPGPAGEAGPPGPPGEAGPPGEPGEPG	NT
Y	PKGPKGYOGPOGYKGYKGPKGPOGYKGPOG	NT
ABD	PKGPKGFOGPOGFKGFKGPKGPOGFKGPOG	33.5
Expected registration	PKGDOGDKGPOGPPGDKGDOGDKGPKGDOG	
	PQGEPGPQGEQGPPGPPGEQGPPGEPGEPG	

[a]Previously published.[64] [b]Sequences and abbreviations of the peptides used in the heterotrimers discussed. The expected registration of the ABD triple helix is shown at the bottom. Amino acids involved in Q−F axial interactions highlighted in gray. NT indicates "no transition".

This implies that, while producing lower stability, the Q−F interaction has a similar or even stronger utility for the specificity of a set of CMPs than does R−F. Furthermore, the use of the Q−F interaction results in a triple helix with net charge 0 rather than net charge +4 as when using the R−F interaction. Finally, positive and negative designs are both important for specificity. Arginine is very weakly destabilizing and requires further negative design in the form of additional destabilizing mutations to promote specificity where glutamine is more destabilizing than arginine and therefore better for negative design.

To characterize the composition and register of the ABD heterotrimer, we utilized the ^{15}N labels in each peptide and performed a series of NMR experiments. ^{1}H, ^{15}N HSQC experiments of each combination of peptides, as shown in Figure 4, indicate that only one composition and only one registration of triple helix exist in these experiments and that it

https://doi.org/10.1021/acs.biomac.1c00234
Biomacromolecules 2021, 22, 2137−2147

Table 3. Melting Temperatures and Specificities of ABC, ABD, and ABE Heterotrimer Mixtures

combination	ABC T_m (°C)	ΔABC	combination	ABD T_m (°C)	ΔABD	combination	ABE T_m (°C)	ΔABE
ABC	39.5[a]		ABD	33.5		ABE	24.0	
AB	21.0[a]	−18.5	AB	21.0	−12.5	AB	21.0	−3.0
AC	23.5[a]	−16.0	AD	12.0	−21.5	AE	NT[b]	
BC	20.5[a]	−19.0	BD	16.5	−17.0	BE	20.5	−3.5
specificity		16.0	specificity		12.5	specificity		3.0

[a]Previously published.[64] [b]NT indicates "no transition".

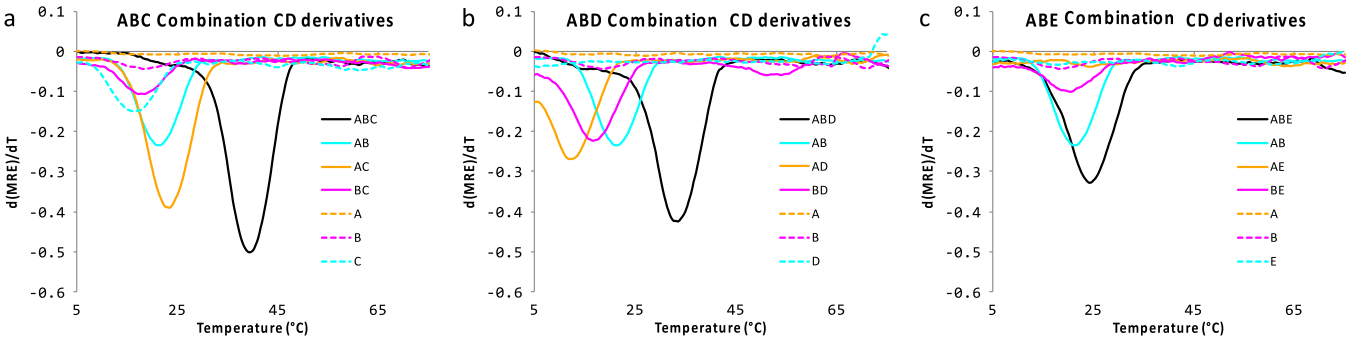


Figure 3. Triple-helix melting as indicated by the first derivative of CD for each possible mixture of the peptides forming heterotrimers. (a) ABC peptides,[64] (b) ABD peptides, and (c) ABE peptides. Note that AB, A, and B are all reproduced in all panels.

is one of the ABD composition heterotrimers (some residual monomeric peptides are observed). The ^{1}H, ^{15}N HSQC experiment, however, cannot distinguish between the six possible registrations.

A 3-D ^{1}H, ^{1}H, and ^{15}N NOESY HSQC experiment was performed to determine the registration of the self-assembling triple helix observed in the HSQC. The full 3-D cube of data is shown in Figure S16. In a NOESY HSQC of a triple-helical assembly, each labeled glycine's amide proton will possess an NOE to the preceding α proton. For example, in the D peptide, the labeled 18G amide N−H should possess an NOE to the 17P α proton, as seen in Figure 5c at (8.14, 4.60) ppm in the 104.0 ppm ^{15}N plane (circled in orange). These strong NOEs allow for the designation of each glycine N−H to its respective peptide strand. Furthermore, each glycine amide N−H also will possess a weaker NOE to the α proton of the Yaa position amino acid in the same helical cross-section as the glycine in question. For example, in the designed triple helix, the D peptide's labeled 18G amide N−H should possess an NOE to the A peptide 20K α proton, as seen in Figure 5c at (8.14, 4.35) ppm in the 104.0 ppm ^{15}N plane (circled in purple). Following these NOEs will inform the arrangement of the triple helix from leading to middle to trailing strands. This analysis confirms that the triple helix that self-assembles is the triple helix that was intended. As shown in the two-dimensional (2-D) slice at 104.0 ppm ^{15}N shift, the D peptide 17P α proton shift is at 4.60 ppm (circled in orange). The D peptide 18G amide with this NOE also shows a weak NOE at 4.35 ppm (circled in purple), which corresponds to the A peptide 20K α proton also seen in the 2-D slice at 109.7 ppm ^{15}N shift (circled in orange). In this slice, also visible is a weak NOE at 4.84 ppm (circle in purple), which corresponds to the B peptide 20O α proton also seen in the 2-D slice at 105.4 ppm ^{15}N shift (circled in orange). Finally, the weak NOE seen in the 105.4 ppm slice corresponds to the α proton of the D peptide 20Q, which is not seen in a separate structure because the D 21G residue is not labeled with a ^{15}N.

Alanine-Substituted Heterotrimer Analysis. In the case of the ABD triple-helix specificity, it could be argued that the interactions from the leading strand to the middle strand and from the middle strand to the trailing strand are sufficient to set the registration of the triple helix, thus rendering the amide−π interactions from the trailing strand to the leading strand unnecessary. To address this argument, we synthesized a third version of the trailing strand, peptide E, containing alanines in place of arginines or glutamines, as shown in Table 2. As argued in our homotrimer analysis, we do not expect the alanines and phenylalanines to interact significantly with each other and so this peptide should remove the potential for interactions between the trailing and leading chains.

As can be seen in both Table 3 and Figure 3, the E containing ternary heterotrimer loses specificity. The specificity as measured by differential CD melting temperatures drops to 3 °C, which is near the limits of our method's precision. The AB peptide mixture spectrum overlaps almost completely with the spectrum of the ABE peptide mixture. Further, small shoulders in the ABE spectrum line up well with the peak in the AB spectrum and suggest that the mixture of ABE peptides yields a mixture of different triple helices. This means that the formation of the ABE triple helix is not specific to the desired ABD registration. This failure indicates the importance of the amide−π interaction: Because specificity is not achieved when the glutamines are mutated to alanines, this demonstrates that the glutamine−phenylalanine axial interaction is a real and effective amide−π interaction. The amide−π interaction is stabilizing for triple helices and capable of improving the specificity of heterotrimers.

Glutamine−Tyrosine Pairs. The amide−π interaction is not limited exclusively to Q−F pairs. An additional peptide "Y" was prepared based on peptide "A" with all occurrences of phenylalanine replaced with tyrosine (see Table 2). A YBD peptide mixture was prepared and found to have stability and specificity equivalent to the ABD peptide mixture, in contrast to the ABE mixture containing alanine. This demonstrates that the amide−π interaction can be successfully implemented with

https://doi.org/10.1021/acs.biomac.1c00234
Biomacromolecules 2021, 22, 2137−2147

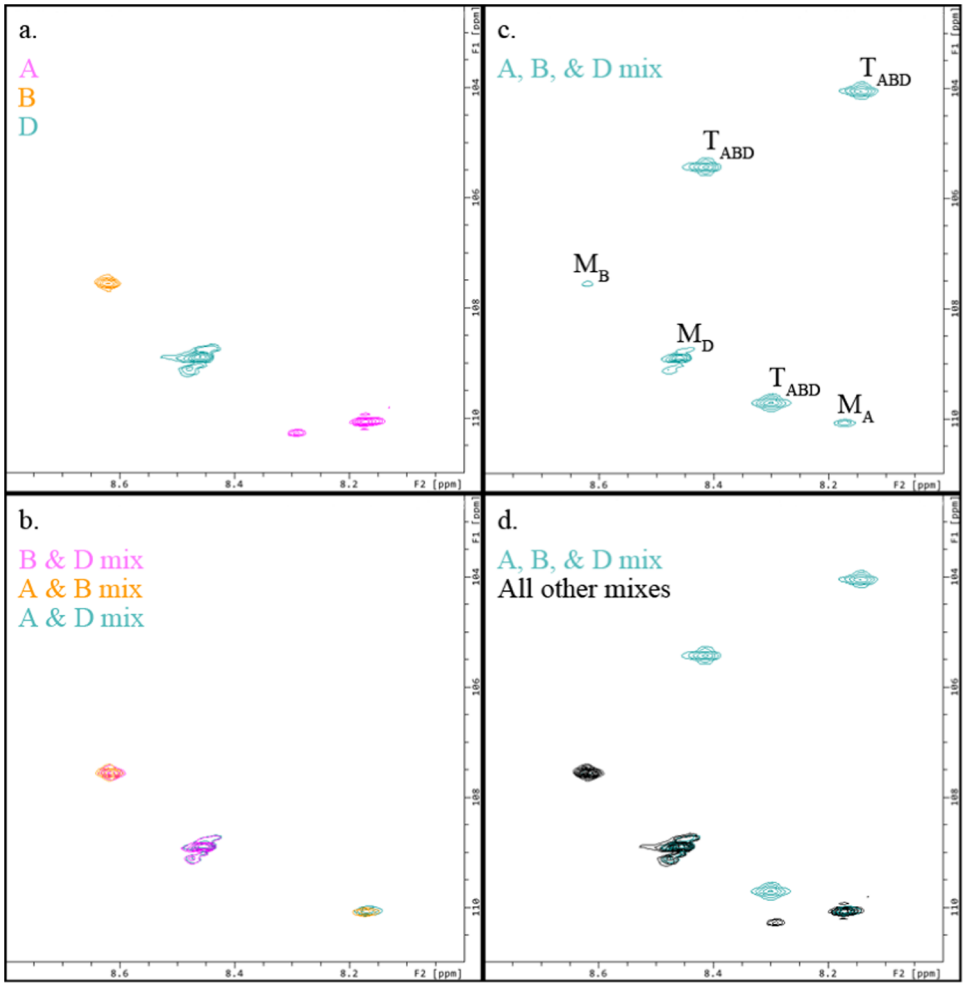


Figure 4. NMR ^{1}H, ^{15}N HSQC plots for the A, B, and D peptides and their mixtures. (a) Unary samples. (b) Binary mixtures. (c) Ternary mixtures. "M" labels monomer peaks with subscripts indicating the peptide. "T" labels triple-helix peaks, with subscripts indicating the composition. (d) Overlapping spectra of all peptide combinations with the ternary ABD combination in cyan and all others in black to emphasize the unique ABD signature. All experiments were run at 25 °C.

tyrosine as the acceptor of the hydrogen bond. A discussion of this helix and its characterization can be found in the supporting information.

Molecular Dynamics of Gln/Asn−Phe Containing Homotrimers. According to our experiments, side-chain amide−π interactions can play an important role in collagen stability and melting temperature. Molecular dynamics (MD) simulation is a robust method to further investigate these interactions and their roles in protein folding and peptide conformations.[72−76] Different force fields, such as CHARMM and AMBER,[68,69] have been used to capture side-chain interactions in collagen and have successfully produced consistent results with experiments in the field.[29,77] Here, we used both CHARMM36 and AMBER-ff14SB force fields to further understand the side-chain interaction between the amide and phenylalanine side chains. CHARMM36 results for **QGFO** are included in the SI, which confirm the conformation suggested by AMBER-ff14SB. The structures analyzed include a $(GPO)_9$ homotrimer (PDB id: 3B0S)[66] and four other homotrimers for each pair analyzed analogous to those analyzed experimentally herein. The starting structures of these homotrimers were developed by mutating residues of the 3B0S crystal structure to match the sequences of each of those homotrimers.

The analyses of the **QGFO** and the **OGFQ** homotrimers corroborate the existence of an amide−π interaction between amino acid side chains exclusively in the axial geometry. Figure 6a shows a snapshot of two axial amide−π interactions in the **QGFO** structure. This snapshot also suggests how these interactions are further stabilized by formation of hydrogen bonds between glutamine side-chain nitrogen and the backbone carbonyl group. Figure 6b shows the same perspective for a **NGFO** snapshot illustrating the lack of amide−π interactions. Figure 6c,d shows the lateral geometries for both of these structures. Figure 6e,f shows the quantification of the distances observed throughout the trajectory of the simulation, where the distance between the amide side-chain nitrogen and the phenylalanine ring center peaks at about 3.8 Å for axially related pairs in **QGFO**, 5.0−6.5 Å for laterally related pairs in both structures, and 8 Å for axially related pairs in **NGFO**. The axial interactions for **QGFO** match the reported values for amide−π interaction distances found in the literature, while the lateral distances are beyond the limit of a viable interaction. Importantly, we performed MD simulations with both AMBER-ff14SB and CHARMM36 force fields, as shown in Figure S22, and the results are consistent between both force fields. Results from both force fields are discussed in the Supporting Information.

https://doi.org/10.1021/acs.biomac.1c00234
Biomacromolecules 2021, 22, 2137−2147

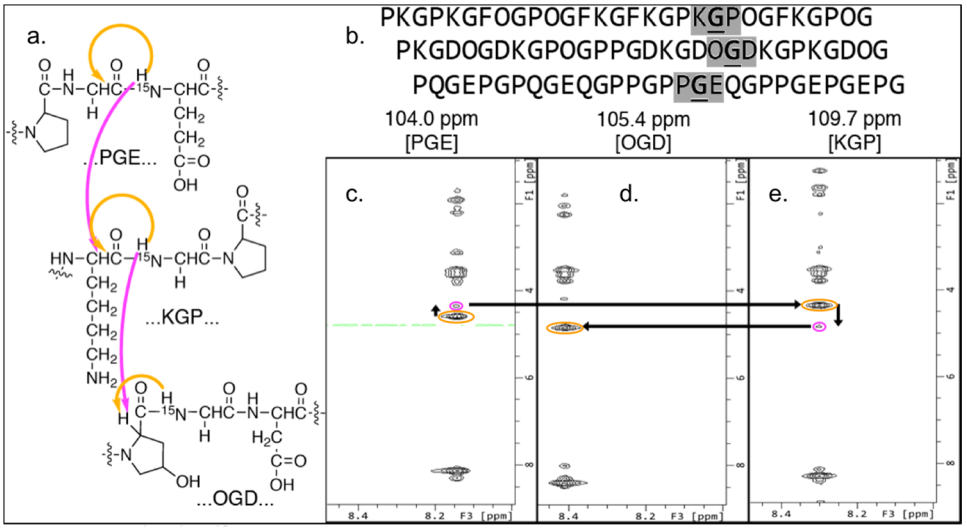


Figure 5. Three-dimensional $^{1}H-^{1}H-^{15}N$ NOESY HSQC analysis of the ABD heterotrimer. (a) Chemical structures of the triplets surrounding the isotopically labeled glycines (highlighted in gray in (b)) and the expected intramolecular (orange) and intermolecular (magenta) NOEs. (b) Sequence and registration of the expected ABD triple helix with isotopically labeled glycines bolded and underlined. (c–e) Three planes of the 3-D NOESY HSQC experiment taken at 104.0, 105.4, and 109.7 ppm in the ^{15}N dimension. The peaks that are used to inform the register assignment are indicated by the arrows. Intermolecular NOEs are indicated by the short vertical arrows and the long horizontal arrows indicate the same proton appearing in multiple planes.

Moreover, the results of the **QGPO** and **OGFO** serve well to justify the occurrence of the weak intermolecular interaction and its substantial effect on the stability of collagen structures observed in **QGFO**. In the **QGPO** structure, it is noted that the glutamine side-chain nitrogen forms a hydrogen bond with a backbone carbonyl group, which is maintained in the **QGFO** and preorganizes the glutamine side-chain atoms in an ideal conformation for the amide–π interaction. The presence of this hydrogen bond in the **QGPO** homotrimer also helps explain the increased thermal stability of this homotrimer in relation to that of the **NGPO** homotrimer in our experiments because the asparagine side chain is not long enough to participate in an analogous hydrogen-bonding interaction. In the **OGFO** structure, it can be noted that the bond angle (~112°) through the phenylalanine α, β, and γ atoms is oriented ideally for the amide–π axial interaction observed in the **QGFO** homotrimer as well. This indicates that the only required degrees of freedom that are not preset for the amide–π interaction are the χ^1 and χ^2 dihedral angles. The low level of perturbation of the side chains needed to form the interaction explains the preference and the stabilization of the interaction while at the same time explaining the absence of interaction in the case involving asparagine.

SCEPTTr Ramifications. The addition of 12 new triple-helix melting temperatures allows for further refinement of SCEPTTr.[64] For an illustration of the importance of this follow-up analysis, Figure 7a shows the fit the SCEPTTr1.0 achieves as previously reported, with computationally predicted values for the Q–F and N–F interactions. SCEPTTr1.0 predicts the singly substituted peptides as well as a few of the others but predicts a few particularly poorly, including the **OGFQ** peptide (experimental, predicted) (24.5, 39.5), the **NGFO** peptide (21.5, 30.6), the ABD triple helix (33.5, 25.6), and the BD triple helix (16.5, 24.8). SCEPTTr's performance is improved in Figure 7b in which all values used by SCEPTTr were subjected to reoptimization to make SCEPTTr1.1. A discussion of the process of the optimization and a table (Table S5) containing the values used by SCEPTTr1.1 can be found in the Supporting Information.

Amide–π Interactions in the Context of Other CMP Interactions. We discussed briefly the comparison of the Q–F interaction and the R–F interaction in the context of the ABC and ABD heterotrimers, but it is important to discuss the amide–π interactions generally compared to other interactions in the context of collagen triple helices. The pairwise interactions that have been successfully used for designing structurally specific heterotrimeric triple helices include (ordered as Y–X pairs) K–D, K–E, R–E, R–F, D–K, E–R, and now Q–F. For the following analysis, all values used are those resulting from the reoptimization of SCEPTTr, as reported in Table S5, Supporting Information. The Q–F and N–F axial interactions are valued at 1.06 and 0.27 °C, respectively, while their respective lateral interactions are −2.06 and −1.00 °C. respectively.

When comparing interactions in collagen, it is important to compare two quantities, stability and specificity. The N–F lateral interaction, though a small effect, is more destabilizing than many of the previously deconvoluted destabilizing interactions; however, paired with such a small stabilizing axial interaction, the N–F pair is not likely to be very helpful in collagen triple-helix design. The Q–F axial interaction is comparable to other previously described moderately stabilizing interactions including the K–D lateral (1.10 °C), R–D axial (1.64 °C), and R–E lateral (0.88 °C) interactions. However, when comparing the specificity of the interactions (the difference between the axial and lateral interactions) for each pair of amino acids, the Q–F (specificity = 3.11) is much more favorable and is similar to the best amino acid pairs such as K–D (specificity = 4.12) and R–F (specificity = 3.43). For this reason, the Q–F interaction exhibits strong potential for use in designing well-folded heterotrimers. However, practically, when designing new heterotrimeric triple helices, interactions will be mixed between different pairs of amino acids rather than strictly competing between lateral and axial geometries of the same pair of amino acids. In this case, these

https://doi.org/10.1021/acs.biomac.1c00234
Biomacromolecules 2021, 22, 2137–2147

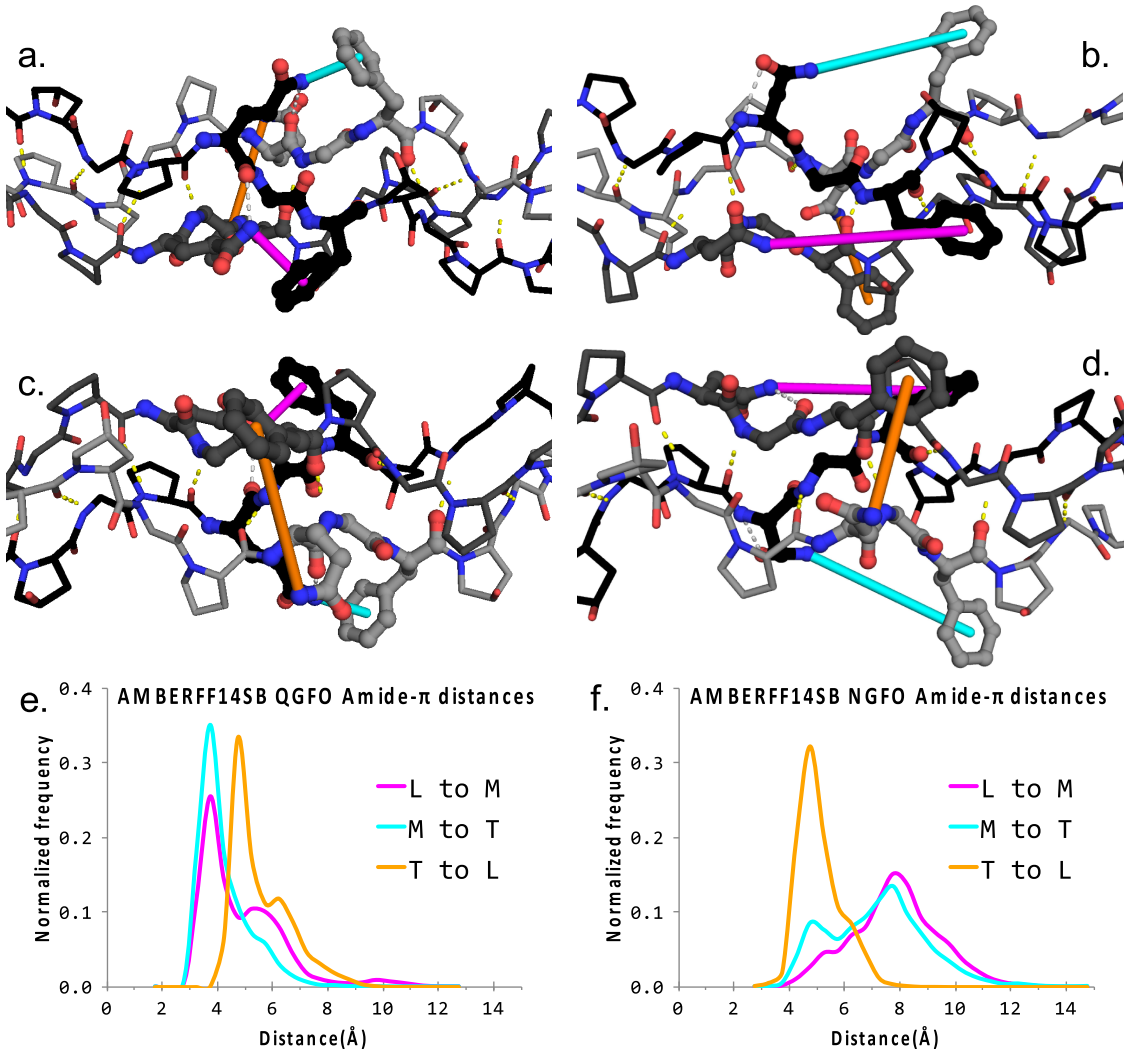


Figure 6. Results from molecular dynamics (MD) simulations. (a) Snapshot of **QGFO** MD highlighting the two axial amide−π interactions. (b) Snapshot of **NGFO** MD indicating the lack of axial amide−π interactions. (c) Snapshot highlighting the **QGFO** lateral geometry. (d) Snapshot highlighting the **NGFO** lateral geometry. The white dotted lines highlight the hydrogen bonds between the glutamine side chains and backbone carbonyls and the yellow dotted lines indicate backbone hydrogen bonds. (e) Plotted frequencies of distances between the glutamine nitrogen and the phenylalanine ring center for each pair of interacting amino acids in **QGFO**, as calculated by AMBER-ff14SB. (f) Plotted distance frequencies for **NGFO**. The distance distribution functions were normalized by the area under each curve. L, M, and T refer to the leading, middle, and trailing strands, respectively. The solid lines colored in (a)−(d) indicate the same interactions as indicated by the colored key used in (e) and (f).

interactions will be even more useful: If a phenylalanine has the opportunity to form an axial interaction with an arginine in one registration or a lateral interaction with a glutamine in another registration, the first will form with higher specificity (5.11 °C) than is accessible to either pair of amino acids alone. This illustrates the desirability of a diverse set of interactions. As a final point of comparison, it is worth reiterating that the Q−F amide−π interaction described here is the first example of a charge-free stabilizing amino acid pairwise interaction for collagen. For charged amino acids, pH and salt concentration become important considerations depending on the intended setting of the application. Additionally, highly charged peptides can be toxic to cells and thus interfere with cell studies.[78−81] These disadvantages can potentially be avoided with charge-free amino acids.

CONCLUSIONS

In this report, we have tested and determined the utility of novel supramolecular interactions in the context of collagen triple helices. Amide−π cooperative interactions were tested and deconvoluted experimentally after which the glutamine−phenylalanine pair was employed in the design of a heterotrimer and proved effective in controlling the triple-helix composition and register. Finally, this interaction was incorporated into the previously reported SCEPTTr algorithm to update it to SCEPTTr1.1. This report adds to the list of those that promote the utility of the amide−π interaction and is the first report of the engineered use of amide−π interactions to design and control a structure. Furthermore, these interactions are important as this is the first report of amide−π pairwise interactions in the context of collagen triple helices, but more so because it is the first report of charge-free pairwise interactions used for controlling triple-helical arrangement. The addition of these interactions to the toolbox of collagen research will allow for both better control of these secondary and tertiary structures through expanded available complimentary interactions and easier application to problems in which charged peptides may be disadvantageous.

https://doi.org/10.1021/acs.biomac.1c00234
Biomacromolecules 2021, 22, 2137−2147

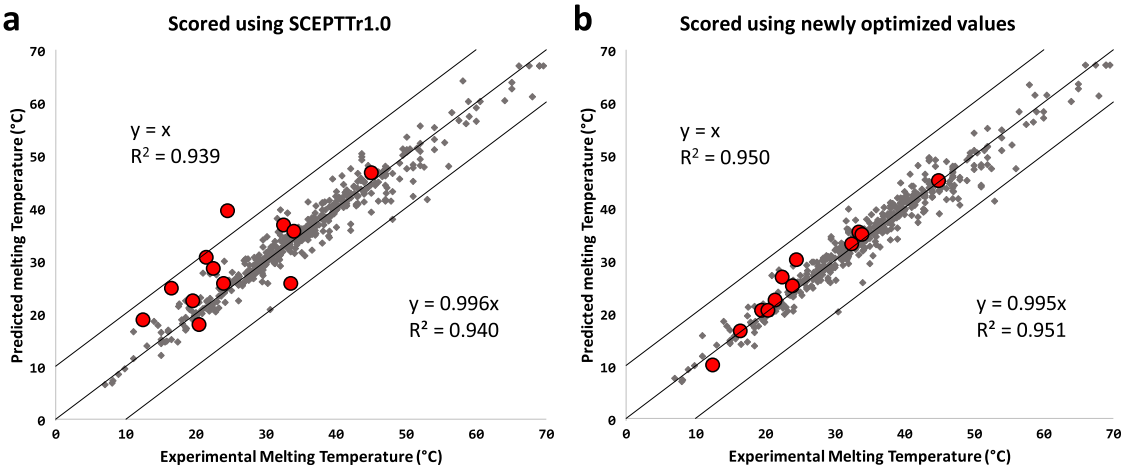


Figure 7. Results of SCEPTTr's analysis of the library of published triple helices. Triple helices first reported here emphasized as the red circles. Each panel illustrates SCEPTTr's performance when using different values in the algorithm. (a) Previously reported SCEPTTr1.0.[64] (b) Algorithm was reoptimized using data from the peptides reported here to generate SCEPTTr1.1.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.biomac.1c00234.

Additional characterization and description of methods (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

Douglas R. Walker − *Department of Chemistry, Rice University, Houston, Texas 77005, United States*

Ali A. Alizadehmojarad − *Department of Chemistry, Rice University, Houston, Texas 77005, United States;* orcid.org/0000-0001-6806-5415

Anatoly B. Kolomeisky − *Department of Chemistry, Department of Chemical and Biomolecular Engineering, and Department of Physics and Astronomy, Rice University, Houston, Texas 77005, United States;* orcid.org/0000-0001-5677-6690

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.biomac.1c00234

Author Contributions

D.R.W. formulated the project idea, designed experiments, performed laboratory experiments, and co-wrote the manuscript. A.A.A. ran molecular dynamics simulations and co-wrote the MD discussion. A.B.K. oversaw molecular dynamics experiments. J.D.H. designed experiments, oversaw the entire project, and co-wrote the manuscript. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was funded in part by the National Science Foundation (CHE 1709631) and the Robert A. Welch Foundation (C1557 and C1559).

■ ABBREVIATIONS

CMP, collagen mimetic peptide; NT, no transition; NMR, nuclear magnetic resonance; HSQC, heteronuclear single quantum coherence; NOESY, nuclear Overhauser effect spectroscopy; SCEPTTr, scoring function for collagen emulating peptides' temperature of transition

■ REFERENCES

(1) Kleinman, H. K.; McGoodwin, E. B.; Martin, G. R.; Klebe, R. J.; Fietzek, P. P.; Wooley, D. E. Localization of the Binding Site for Cell Attachment in the al(I) Chain of Collagen. *J. Biol. Chem.* **1978**, *253*, 5642−5646.

(2) Shah, N. K.; Ramshaw, J. A.; Kirkpatrick, A.; Shah, C.; Brodsky, B. A Host-Guest Set of Triple-Helical Peptides: Stability of Gly-X-Y Triplets Containing Common Nonpolar Residues. *Biochemistry* **1996**, *35*, 10262−10268.

(3) Venugopal, M. G.; Ramshaw, J. A. M.; Braswell, E.; Zhu, D.; Brodsky, B. Electrostatic Interactions in Collagen-like Triple-Helical Peptides. *Biochemistry* **1994**, *33*, 7948−7956.

(4) Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrick, A.; Brodsky, B. Amino Acid Propensities for the Collagen Triple-Helix. *Biochemistry* **2000**, *39*, 14960−14967.

(5) Lauer-Fields, J. L.; Tuzinski, K. A.; Shimokawa, K.; Nagase, H.; Fields, G. B. Hydrolysis of Triple-helical Collagen Peptide Models by Matrix Metalloproteinases. *J. Biol. Chem.* **2000**, *275*, 13282−13290.

(6) 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.

(7) Kim, J. K.; Xu, Y.; Xu, X.; Keene, D. R.; Gurusiddappa, S.; Liang, X.; Wary, K. K.; Hook, M. A novel binding site in collagen type III for integrins alpha1beta1 and alpha2beta1. *J. Biol. Chem.* **2005**, *280*, 32512−32520.

(8) Gauba, V.; Hartgerink, J. D. Synthetic Collagen Heterotrimers: Structural Mimics of Wild-Type and Mutant Collagen Type I. *J. Am. Chem. Soc.* **2008**, *130*, 7509−7515.

(9) 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.

(10) 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.
(11) Schwarz, M., Jr.; Poland, D. Statistical Thermodynamics of Triple-Helix Unzippering for the Collagen Model (Gly-Pro-Pro), and Implications for Natural Collagen. *Biopolymers* **1974**, *13*, 687−701.
(12) Pires, M. M.; Chmielewski, J. Self-assembly of Collagen Peptides into Microflorettes via Metal Coordination. *J. Am. Chem. Soc.* **2009**, *131*, 2706−2712.
(13) Przybyla, D. E.; Chmielewski, J. Metal-Triggered Collagen Peptide Disk Formation. *J. Am. Chem. Soc.* **2010**, *132*, 7866−7867.
(14) O'Leary, L. E. R.; Fallas, J. A.; Bakota, E. L.; Kang, M. K.; Hartgerink, J. D. Multi-hierarchical self-assembly of a collagen mimetic peptide from triple helix to nanofibre and hydrogel. *Nat. Chem.* **2011**, *3*, 821−828.
(15) Jiang, T.; Xu, C.; Liu, Y.; Liu, Z.; Wall, J. S.; Zuo, X.; Lian, T.; Salaita, K.; Ni, C.; Pochan, D.; Conticello, V. P. Structurally defined nanoscale sheets from self-assembly of collagen-mimetic peptides. *J. Am. Chem. Soc.* **2014**, *136*, 4300−4308.
(16) Tanrikulu, I. C.; Forticaux, A.; Jin, S.; Raines, R. T. Peptide tessellation yields micrometre-scale collagen triple helices. *Nat. Chem.* **2016**, *8*, 1008−1014.
(17) Lovejoy, B.; Choe, S.; Cascio, D.; McRorie, D. K.; DeGrado, W. F.; Eisenberg, D. Crystal Structure of a Synthetic Triple-Stranded α-Helical Bundle. *Science* **1993**, *259*, 1288−1293.
(18) Kiehna, S. E.; Waters, M. L. Sequence dependence of beta-hairpin structure: comparison of a salt bridge and an aromatic interaction. *Protein Sci.* **2003**, *12*, 2657−2667.
(19) Zimenkov, Y.; Conticello, V. P.; Guo, L.; Thiyagarajan, P. Rational design of a nanoscale helical scaffold derived from self-assembly of a dimeric coiled coil motif. *Tetrahedron* **2004**, *60*, 7237−7246.
(20) Zimenkov, Y.; Dublin, S. N.; Ni, R.; Tu, R. S.; Breedveld, V.; Apkarian, R. P.; Conticello, V. P. Rational Design of a Reversible pH-Responsive Switch for Peptide Self-Assembly. *J. Am. Chem. Soc.* **2006**, *128*, 6770−6771.
(21) Moutevelis, E.; Woolfson, D. N. A periodic table of coiled-coil protein structures. *J. Mol. Biol.* **2009**, *385*, 726−732.
(22) Aulisa, L.; Dong, H.; Hartgerink, J. D. Self-Assembly of Multidomain Peptides- Sequence Variation Allows Control over Cross-Linking and Viscoelasticity. *Biomacromolecules* **2009**, *10*, 2694−2698.
(23) Fallas, J. A.; Ueda, G.; Sheffler, W.; Nguyen, V.; McNamara, D. E.; Sankaran, B.; Pereira, J. H.; Parmeggiani, F.; Brunette, T. J.; Cascio, D.; Yeates, T. R.; Zwart, P.; Baker, D. Computational design of self-assembling cyclic protein homo-oligomers. *Nat. Chem.* **2017**, *9*, 353−360.
(24) Rhys, G. G.; Wood, C. W.; Lang, E. J. M.; Mulholland, A. J.; Brady, R. L.; Thomson, A. R.; Woolfson, D. N. Maintaining and breaking symmetry in homomeric coiled-coil assemblies. *Nat. Commun.* **2018**, *9*, No. 4132.
(25) Koepnick, B.; Flatten, J.; Husain, T.; Ford, A.; Silva, D. A.; Bick, M. J.; Bauer, A.; Liu, G.; Ishida, Y.; Boykov, A.; Estep, R. D.; Kleinfelter, S.; Norgard-Solano, T.; Wei, L.; Players, F.; Montelione, G. T.; DiMaio, F.; Popovic, Z.; Khatib, F.; Cooper, S.; Baker, D. De novo protein design by citizen scientists. *Nature* **2019**, *570*, 390−394.
(26) Fallas, J. A.; Gauba, V.; Hartgerink, J. D. Solution structure of an ABC collagen heterotrimer reveals a single-register helix stabilized by electrostatic interactions. *J. Biol. Chem.* **2009**, *284*, 26851−26859.
(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) Chen, C.-C.; Hsu, W.; Hwang, K. C.; Hwu, J. R.; Lin, C. C.; Horng, J.-C. Contributions of cation-π interactions to the collagen triple helix stability. *Arch. Biochem. Biophys.* **2011**, *508*, 46−53.
(29) Zheng, H.; Lu, C.; Lan, J.; Fan, S.; Nanda, V.; Xu, F. How electrostatic networks modulate specificity and stability of collagen. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 6207−6212.
(30) Wulf, O. R.; Liddel, U.; Hendricks. The Effect of Ortho Substitution on the Absorption of the OH Group of Phenol in the Infrared. *J. Am. Chem. Soc.* **1936**, *58*, 2287−2293.
(31) Burley, S. K.; Petsko, G. A. Amino-aromatic interactions in proteins. *FEBS Lett.* **1986**, *203*, 139−143.
(32) Levitt, M.; Perutz, M. F. Aromatic Rings Act as Hydrogen Bond Acceptors. *J. Mol. Biol.* **1988**, *201*, 751−754.
(33) Rodham, D. A.; Suzuki, S.; Suenram, R. D.; Lovas, F. J.; Dasgupta, S.; Goddard, W. A., III; Blake, G. A. Hydrogen bonding in the benzene-ammonia dimer. *Nature* **1993**, *362*, 735−737.
(34) van der Spoel, D.; van Buuren, A. R.; Tieleman, D. P.; Berendsen, H. J. C. Molecular dynamics simulations of peptides from BPTI- A closer look at amide-aromatic interactions. *J. Biomol. NMR* **1996**, *8*, 229−238.
(35) Duan, G.; Smith, V. H., Jr.; Weaver, D. F. An ab initio and data mining study on aromatic−amide interactions. *Chem. Phys. Lett.* **1999**, *310*, 323−332.
(36) Duan, G.; Smith, V. H., Jr.; Weaver, D. F. A data mining and ab initio study of the interaction between the aromatic and backbone amide groups in proteins. *Int. J. Quantum Chem.* **2000**, *80*, 44−60.
(37) Duan, G.; Smith, V. H., Jr.; Weaver, D. F. Characterization of Aromatic−Amide(Side-Chain) Interactions in Proteins through Systematic ab Initio Calculations and Data Mining Analyses. *J. Phys. Chem. A* **2000**, *104*, 4521−4532.
(38) Tóth, G.; Murphy, R. F.; Lovas, S. Investigation of Aromatic-Backbone Amide Interactions in the Model Peptide Acetyl-Phe-Gly-Gly-N-Methyl Amide Using Molecular Dynamics Simulations and Protein Database Search. *J. Am. Chem. Soc.* **2001**, *123*, 11782−11790.
(39) Biot, C.; Buisine, E.; Rooman, M. Free-Energy Calculations of Protein−Ligand Cation−π and Amino−π Interactions- From Vacuum to Proteinlike Environments. *J. Am. Chem. Soc.* **2003**, *125*, 13988−13994.
(40) Toth, G.; Kover, K. E.; Murphy, R. F.; Lovas, S. Aromatic−Backbone Interactions in α-Helices. *J. Phys. Chem. B* **2004**, *108*, 9287−9296.
(41) Chin, W.; Mons, M.; Dognon, J.; Mirasol, R.; Chass, G.; Dimicoli, I.; Piuzzi, F.; Butz, P.; Tardivel, B.; Compagnon, I.; von Helden, G.; Meijer, G. The Gas-Phase Dipeptide Analogue Acetyl-phenylalanyl-amide- A Model for the Study of Side Chain:Backbone Interactions in Proteins. *J. Phys. Chem. A* **2005**, *109*, 5281−5288.
(42) Vaupel, S.; Brutschy, B.; Tarakeshwar, P.; Kim, K. S. Characterization of Weak NH−π Intermolecular Interactions of Ammonia with Various Substituted π-Systems. *J. Am. Chem. Soc.* **2006**, *128*, 5416−5426.
(43) Ottiger, P.; Pfaffen, C.; Leist, R.; Leutwyler, S.; et al. Strong N-H···π Hydrogen Bonding in Amide-Benzene Interactions. *J. Phys. Chem. B* **2009**, *113*, 2937−2943.
(44) Trachsel, M. A.; Ottiger, P.; Frey, H. M.; Pfaffen, C.; Bihlmeier, A.; Klopper, W.; Leutwyler, S. Modeling the Histidine-Phenylalanine Interaction: The NH...π Hydrogen Bond of Imidazole.Benzene. *J. Phys. Chem. B* **2015**, *119*, 7778−7790.
(45) Sohn, W. Y.; Brenner, V.; Gloaguen, E.; Mons, M. Local NH-π interactions involving aromatic residues of proteins: influence of backbone conformation and $\pi\pi^*$ excitation on the π H-bond strength, as revealed from studies of isolated model peptides. *Phys. Chem. Chem. Phys.* **2016**, *18*, 29969−29978.
(46) Cheng, J.; Kang, C.; Zhu, W.; Luo, X.; Puah, C. M.; Chen, K.; Shen, J.; Jiang, H. N-Methylformamide-Benzene Complex as a Prototypical Peptide N-H,,,π Hydrogen-Bonded System- Density Functional Theory and MP2 Studies. *J. Org. Chem.* **2003**, *68*, 7490−7495.
(47) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. Origin of the Attraction and Directionality of the NH:π Interaction-Comparison with OH:π and CH:π Interactions. *J. Am. Chem. Soc.* **2000**, *122*, 11450−11458.
(48) Suresh, C. H.; Mohan, N.; Vijayalakshmi, K. P.; George, R.; Mathew, J. M. Typical aromatic noncovalent interactions in proteins: A theoretical study using phenylalanine. *J. Comput. Chem.* **2009**, *30*, 1392−1404.

https://doi.org/10.1021/acs.biomac.1c00234
Biomacromolecules 2021, 22, 2137−2147

(49) Liao, S.-M.; Du, Q.-S.; Meng, J.-Z.; Pang, Z.-W.; Huang, R.-B. The multiple roles of histidine in protein interactions. *Chem. Cent. J.* **2013**, *7*, No. 44.
(50) Du, Q.-S.; Wang, Q.-Y.; Du, L.-Q.; Huang, R.-B. Theoretical study on the polar hydrogen-π (Hp-π) interactions between protein side chains. *Chem. Cent. J.* **2013**, *7*, No. 92.
(51) Du, Q. S.; Chen, D.; Xie, N. Z.; Huang, R. B.; Chou, K. C. Insight into a molecular interaction force supporting peptide backbones and its implication to protein loops and folding. *J. Biomol. Struct. Dyn.* **2015**, *33*, 1957−1972.
(52) Braga, D.; Grepioni, F.; Tedesco, E. X-H—π (X) O, N, C) Hydrogen Bonds in Organometallic Crystals. *Organometallics* **1998**, *17*, 2669−2672.
(53) Steiner, T.; Koellner, G. Hydrogen bonds with pi-acceptors in proteins: frequencies and role in stabilizing local 3D structures. *J. Mol. Biol.* **2001**, *305*, 535−557.
(54) Tóth, G.; Watts, C. R.; Murphy, R. F.; Lovas, S. Significance of aromatic-backbone amide interactions in protein structure. *Proteins* **2001**, *43*, 373−381.
(55) Anbarasu, A.; Anand, S.; Sethumadhavan, R. NH...π Interactions: Investigations on the Evidence and Consequences in RNA Binding Proteins. *Open Struct. Biol. J.* **2008**, *2*, 33−42.
(56) Mohapatra, S. R.; Ramanathan, K.; Shanthi, V.; Srivastava, S.; Sethumadhavan, R. Computational Investigation Of N-H...π Interactions In The Structural Stability of Transmembrane Proteins. *Int. J. Pharm. Pharm. Sci.* **2011**, *3*, 106−111.
(57) Shanthi, V.; Sethumadhavan, R. Computational Perspective in the Structural Stability of 'All-Alpha' Proteins: The NH...π Interactions. *Int. J. Pharm. Pharm. Sci.* **2011**, *3*, 138−144.
(58) Mooibroek, T. J.; Gamez, P. How directional are D−H···phenyl interactions in the solid state (D = C, N, O)? *CrystEngComm* **2012**, *14*, 8462−8467.
(59) Lavanya, P.; Ramaiah, S.; Anbarasu, A. Computational analysis of N-H-π interactions and its impact on the structural stability of beta-lactamases. *Comput. Biol. Med.* **2014**, *46*, 22−28.
(60) Kerr, J. R.; Trembleau, L.; Storey, J. M.; Wardell, J. L.; Harrison, W. T. Different N-H−π interactions in two indole derivatives. *Acta Crystallogr., Sect. E: Crystallogr. Commun.* **2016**, *72*, 699−703.
(61) Armstrong, K. M.; Fairman, R.; Baldwin, R. L. The (i,i+4) Phe-His Interaction Studied in an Alanine-based α-Helix. *J. Mol. Biol.* **1993**, *230*, 284−291.
(62) Parkinson, G.; Gunasekera, A.; Vojtechovsky, J.; Zhang, X.; Kunkel, T. A.; Berman, H.; Ebright, R. H. Aromatic hydrogen bond in sequence specific protein DNA recognition. *Nat. Struct. Mol. Biol.* **1996**, *3*, 837−841.
(63) Hughes, R. M.; Waters, M. L. Effects of Lysine Acetylation in beta-Hairpin Peptide Comparison of an Amide-π and a Cation-π Interaction. *J. Am. Chem. Soc.* **2006**, *128*, 13586−13591.
(64) Walker, D. R.; Hulgan, S. A. H.; Peterson, C. M.; Li, I.-C.; Gonzalez, K. J.; Hartgerink, J. D. Predicting the Stability of Homo- and Hetero-trimeric Collagen Helices. *Nat. Chem.* **2021**, *13*, 260−269.
(65) Li, I.-C.; Hulgan, 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.
(66) Okuyama, K.; Miyama, K.; Mizuno, K.; Bachinger, H. P. Crystal structure of (Gly-Pro-Hyp)(9): implications for the collagen molecular model. *Biopolymers* **2012**, *97*, 607−616.
(67) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(68) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; Mackerell, A. D., Jr. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257−3273.
(69) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.
(70) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: AnN·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(71) Humphrey, W.; Dalke, A.; Schulten, K. VMD Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(72) Taddese, B.; Garnier, A.; Abdi, H.; Henrion, D.; Chabbert, M. Deciphering collaborative sidechain motions in proteins during molecular dynamics simulations. *Sci. Rep.* **2020**, *10*, No. 15901.
(73) Cousin, S. F.; Kaderavek, P.; Bolik-Coulon, N.; Gu, Y.; Charlier, C.; Carlier, L.; Bruschweiler-Li, L.; Marquardsen, T.; Tyburn, J. M.; Bruschweiler, R.; Ferrage, F. Time-Resolved Protein Side-Chain Motions Unraveled by High-Resolution Relaxometry and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2018**, *140*, 13456−13465.
(74) Trbovic, N.; Cho, J.-H.; Abel, R.; Friesner, R. A.; Rance, M.; Palmer, A. G., III Protein Side-Chain Dynamics and Residual Conformational Entropy. *J. Am. Chem. Soc.* **2009**, *131*, 615−622.
(75) Guvench, O.; Mallajosyula, S. S.; Raman, E. P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T. J.; Jamison, F. W., 2nd; Mackerell, A. D., Jr. CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling. *J. Chem. Theory Comput.* **2011**, *7*, 3162−3180.
(76) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78*, 1950−1958.
(77) Gurry, T.; Nerenberg, P. S.; Stultz, C. M. The contribution of interchain salt bridges to triple-helical stability in collagen. *Biophys. J.* **2010**, *98*, 2634−2643.
(78) Saar, K.; Lindgren, M.; Hansen, M.; Eiriksdottir, E.; Jiang, Y.; Rosenthal-Aizman, K.; Sassian, M.; Langel, U. Cell-penetrating peptides: a comparative membrane toxicity study. *Anal. Biochem.* **2005**, *345*, 55−65.
(79) Marr, A. K.; Gooderham, W. J.; Hancock, R. E. Antibacterial peptides for therapeutic use: obstacles and realistic outlook. *Curr. Opin. Pharmacol.* **2006**, *6*, 468−472.
(80) Gallardo-Godoy, A.; Muldoon, C.; Becker, B.; Elliott, A. G.; Lash, L. H.; Huang, J. X.; Butler, M. S.; Pelingon, R.; Kavanagh, A. M.; Ramu, S.; Phetsang, W.; Blaskovich, M. A.; Cooper, M. A. Activity and Predicted Nephrotoxicity of Synthetic Antibiotics Based on Polymyxin B. *J. Med. Chem.* **2016**, *59*, 1068−1077.
(81) Cardozo, A. K.; Buchillier, V.; Mathieu, M.; Chen, J.; Ortis, F.; Ladriere, L.; Allaman-Pillet, N.; Poirot, O.; Kellenberger, S.; Beckmann, J. S.; Eizirik, D. L.; Bonny, C.; Maurer, F. Cell-permeable peptides induce dose- and length-dependent cytotoxic effects. *Biochim. Biophys. Acta, Biomembr.* **2007**, *1768*, 2222−2234.

https://doi.org/10.1021/acs.biomac.1c00234
Biomacromolecules 2021, 22, 2137−2147