JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

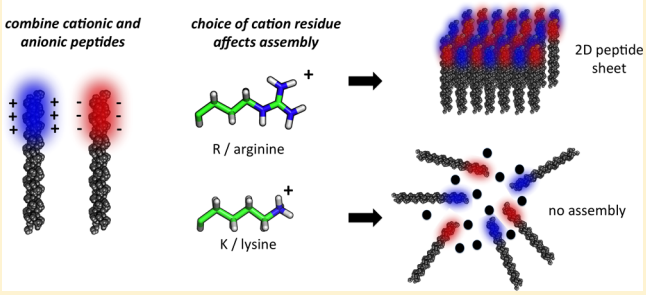
Dissecting Electrostatic Contributions to Folding and Self-Assembly Using Designed Multicomponent Peptide Systems

Avanish S. Parmar,[†,¶] Jose K. James,[‡,¶] Daniel R. Grisham,[‡] Douglas H. Pike,[‡] and Vikas Nanda*[,‡]

[†]Department of Physics, Indian Institute of Technology (Banaras Hindu University), Varanasi 221005, Uttar Pradesh, India
[‡]Center for Advanced Biotechnology and Medicine, Department of Biochemistry and Molecular Biology, Robert Wood Johnson Medical School, Rutgers University, 679 Hoes Lane West, Piscataway, New Jersey 08854, United States

Supporting Information

ABSTRACT: We investigate formation of nano- to microscale peptide fibers and sheets where assembly requires association of two distinct collagen mimetic peptides (CMPs). The multicomponent nature of these designs allows the decoupling of amino acid contributions to peptide folding versus higher-order assembly. While both arginine and lysine containing CMP sequences can favor triple-helix folding, only arginine promotes rapid supramolecular assembly in each of the three two-component systems examined. Unlike lysine, the polyvalent guanidyl group of arginine is capable of both intra- and intermolecular contacts, promoting assembly. This is consistent with the supramolecular diversity of CMP morphologies observed throughout the literature. It also connects CMP self-assembly with a broad range of biomolecular interaction phenomena, providing general principles for modeling and design.



INTRODUCTION

Collagen lacks the extensive hydrophobic core typical of globular proteins, such that electrostatic forces take on particular prominence in specifying structure and supramolecular morphology. This phenomenon has largely been studied in the context of model collagen mimetic peptides (CMPs). Charge–pair interactions can mediate stability[1] and specificity of interchain interactions.[2] Beyond folding, the higher-order association of natural collagens into microfibrils and bundled fibers is governed in large part by long-range electrostatic forces.[3] This process too has been investigated using designed CMPs, where electrostatically driven self-assembly results in diverse nano- and microscale structures from fibers and tapes to highly regular sheets.[4]

Adequate treatment of electrostatic forces continues to be a challenge for computational protein design. The sophistication used to treat electrostatic forces at the atomic scale can range from simple rules, e.g., "opposites attract, like charges repel", which has been extensively applied in rational protein design and in simple scoring models,[5] to sophisticated atomistic calculations.[6] Analytical models for treating electrostatics in polyelectrolytes or macro-ions in solutions of co- and counter-ions are limited to idealized systems and challenging to implement at a detailed atomistic level for proteins. Seemingly innocuous perturbations to a protein sequence, such as switching cationic amino acids arginine and lysine for each other, can significantly impact structure, stability, or function,[7] highlighting the importance of appropriate treatment of electrostatics.

Conservative amino acid substitutions in CMPs can dramatically alter higher-order assembly. Peptides consisting of tandem cationic, neutral, and anionic domains: $(PxG)_a$-$(POG)_b$-$(yOG)_c$ were shown to fold into triple-helices and further self-assemble into higher-order structures through presumed interhelical interactions between oppositely charged domains; P = proline, G = glycine, O = (4*R*)-hydroxyproline. The morphology varied from banded fibers if *x* = arginine/R, *y* = glutamate/E;[4a] fibrous hydrogels if *x* = lysine/K, *y* = aspartate/D;[4b] or rectangular nanosheets if *x* = (4*R*)-amino-proline and *y* = E.[4d,8] Such variations in assembly behavior are difficult to rationalize from a simple view of electrostatics. Effective design of such systems requires a deeper understanding of the role of intermolecular electrostatic forces in peptide assembly.

Given their rich assembly behavior, CMPs are a powerful platform for developing electrostatics-based design rules for controlling self-assembly. Here, we exploit two-component peptide systems that only assemble in the presence of both components. Three systems are presented that are designed to interact through electrostatic forces. The first was derived from previous computational design efforts,[4c] and although an interesting decoupling of triple-helix folding and higher-order assembly was observed, the complexity of the designed sequence was not amenable to developing a mechanism. The subsequent two sets of designs used simple charge patterns, placing cationic and anionic domains on separate components. This allowed us to dissect the contributions of amino acid type and position on self-assembly.

Received: October 1, 2015
Published: March 11, 2016

DOI: 10.1021/jacs.5b10304
J. Am. Chem. Soc. 2016, 138, 4362–4367

RESULTS

Folding/Assembly of Heteromeric CMPs. Previously, three CMPs were designed to assemble heterospecifically through a network of interchain charge−pair interactions.[5d] Of these, charged-peptides B and C (denoted here as cpB-RE and cpC-RE) assembled in a manner dependent on both the stoichiometry and concentration of the two components.[4c] All peptide sequences consisted of one charged residue per three amino acids:

cpB-RE: PEGEOGROGROGPRGPRGPRGROGROGPRG

cpC-RE: PRGROGROGROGPRGEOGPEGPEGEOGEOG

Peptide cpB-RE was rich in arginine (net charge = +6). cpC-RE was composed of equal numbers of of acidic and basic amino acids (no net charge). A persistent hydrogel state was observed when peptides were induced to fold into a triple helix with stoichiometry of 1:2 cpB-RE to cpC-RE, total peptide concentration >0.8 mM. Supramolecular assembly was sensitive to ionic strength, supporting the role of favorable electrostatic interactions in promoting association. Microrheological characterization of the hydrogel state indicated a weak storage modulus, which was speculated to originate from the low stability of the heterotrimeric CMP units (T_m < 20 °C).[4c]

We initially supposed that enhancing intrahelical interactions might also improve hydrogel stability. Interactions involving lysine have been shown to increase CMP stability relative to those involving arginine.[1b,2a,9] Although the cpB-RE + cpC-RE association state was not an intentional outcome of the original design, heterotrimer folding was sensitive to ionic strength, supporting favorable charge−pair interactions promoting this state. To examine whether such interactions could be enhanced, D/K containing variants were synthesized that preserved the original charge pattern.

cpB-KD: PDGDOGKOGKOGPKGPKGPKGKOGKOGPKG

cpC-KD: PKGKOGKOGKOGPKGDOGPDGPDGDOGDOG

Asymmetry in the contributions of K/D vs R/E to cpB and cpC folding was evident: cpB-RE + cpC-KD folded, whereas cpB-KD + cpC-RE did not (Figure S1). Instead of forming a stable triple-helix as expected, cpB-KD + cpC-KD mixtures did not fold; no evidence of triple-helical structure was observed by CD spectroscopy (Figure 1A). Although appropriately presented K/D interactions can confer greater stability than R/E because of stronger charge pairing, R, K, D, and E also have different intrinsic polyproline-II propensities.[1a] The lack of enhanced stability in cpB-KD + cpC-KD mixtures may be due to unfavorable helical propensities of K and D that were not counterbalanced by charge pairs. However, without direct structural information, it is challenging to establish a mechanism.

We also observed decoupling of folding and higher-order assembly contingent on the sequence of the two components. From previous work, it was known that triple-helix folding by cpB-RE+cpC-RE was a prerequisite for supramolecular assembly.[4c] Here, cpB-RE + cpC-KD did fold, but did not self-assemble into fibers despite having the same charge pattern (Figures 1B−E and S1−S3). Charged residues are playing roles in both folding and interactions that drive higher-order assembly in a manner that is sensitive to the choice of lysine versus arginine. The complex charge pattern of the cpB/cpC system makes it difficult to identify specific interactions responsible for these two stages of assembly. Therefore, a multicomponent system was designed that could specifically address the contributions of K, R, D, and E to supramolecular assembly by separating cationic and anionic domains into separate molecules.

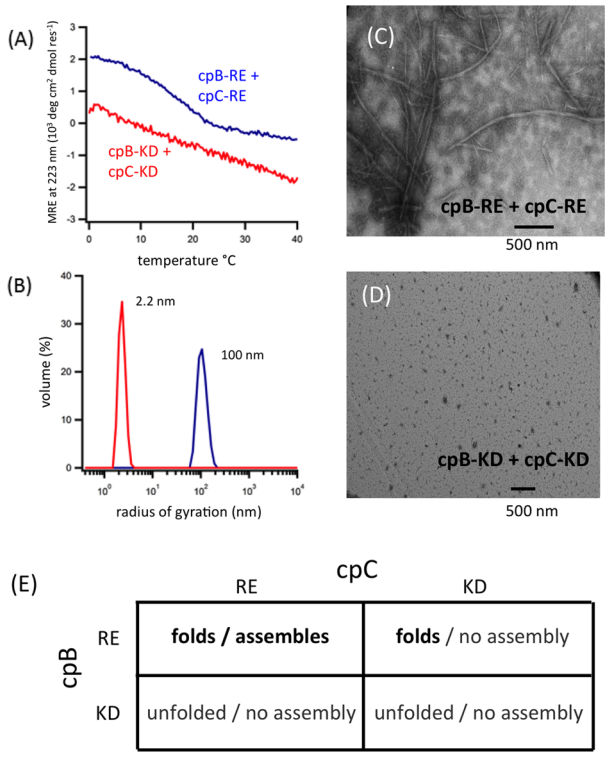


Figure 1. Folding and assembly of the cpB + cpC system. (A) Thermal denaturation monitored by CD shows a cooperative, albeit broad, unfolding transition of the triple-helix for cpB-RE + 2 cPC-RE (blue), but not the K/D containing counterparts (red). Higher-order assembly observed for R/E peptides only by (B) DLS and (C) TEM. (D) No structured aggregates observed for a 1:2 equiv combination of cpB-KD + cpC-KD. (E) Summary of folding and fiber-assembly behavior of cpB + cpC combinations. Additional data in Figure S1.

Folding of Supercharged Peptides. A series of four peptides were designed to fold as homotrimers, each with an N-terminal supercharged domain followed by a neutral, triple-helix promoting domain:

R6: $(RRG)_3(POG)_7$

K6: $(KKG)_3(POG)_7$

E6: $(EEG)_3(POG)_7$

D6: $(DDG)_3(POG)_7$

Although interchain charge repulsions would impact stability of a folded triple-helix, such charged domains were previously shown to be structurally tolerated at the N-terminus.[10] We expected the four homotrimers would be highly soluble alone, and that these could be induced to rapidly assemble by mixed oppositely charged homotrimers. Differences in the nature of multicomponent macro-ion assembly could be used to dissect the relative contributions of these acidic and basic groups.

R6 and K6 homotrimers were well folded with the relative stabilities consistent with triple-helical propensities of arginine > lysine (Figures 2 and S4). The two acidic species were notably less stable, with E6 only exhibiting marginal structure and a T_m < 20 °C. All species were soluble in buffered solution (10 mM phosphate buffer, pH 7) and showed a hydrodynamic radius of ∼2−3 nm, consistent with that of a rod-like triple-helix;[11] however, the D6 and E6 were likely partially unfolded.

DOI: 10.1021/jacs.5b10304
J. Am. Chem. Soc. 2016, 138, 4362−4367

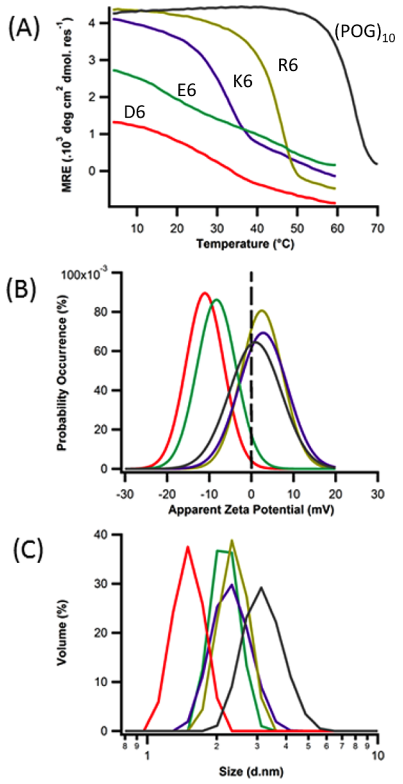


Figure 2. Supercharged peptides. (A) Thermal denaturation profiles show broad range of triple-helix stabilities. Melting temperatures: K6, 31.1 °C; R6, 44.3 °C; D6, 26.5 °C; E6, 17 °C; POG10, 65 °C. (B) Zeta-potentials derived from electrophoretic mobility measured by DLS. (C) Hydrodynamic radii of peptides from 1 to 10 nm consistent with triple-helical rod formation for most peptides with D6 largely unstructured.

Assembly of Supercharged Peptides. If one considers these four peptides as featureless macro-ionic rods, then any combination of a cationic + anionic peptide should result in higher-order assembly based on the premise of opposites attracting. Therefore, it was quite unexpected that of R6 and K6, only R6 assembled in the presence of D6 or E6 (Figures 3 and S5). Association was monitored by dynamic light scattering (DLS), using the total derived count rate as a measure of degree of assembly. R6 assembly with E6 was well described by the Watzky-Finke (WF) minimal model. This model assumes that monomer to aggregate, A → B, proceeds through a generic two-step process of nucleation followed by rapid growth.[12]

$$[\mathrm{B}]_t = [\mathrm{A}]_0 - \frac{\frac{k_1}{k_2} + [\mathrm{A}]_0}{1 + \frac{k_1}{k_2[\mathrm{A}]_0}\exp(k_1 + k_2[\mathrm{A}]_0)t} \quad (1)$$

The nucleation rate constant, k_1 was identical at low and high concentrations; growth rates k_2 proportional to initial concentration. K6 did not assemble with anionic peptides even at the highest tested concentrations after several hours.

Transmission electron micrographs of the R6 + D6 and R6 + E6 aggregates showed extended tens-of-micronscale peptide sheets (Figure 4A,B), sharing many features with two-dimensional morphologies generated previously by metal, electrostatic, hydrophobic, or shape-complementary driven self-assembling CMPs.[4d,8b,13] The K6 containing mixtures showed dense particulates on the order of a few hundred nanometers to 1 μm in size (Figure 4C,D). These particulates were not evident in

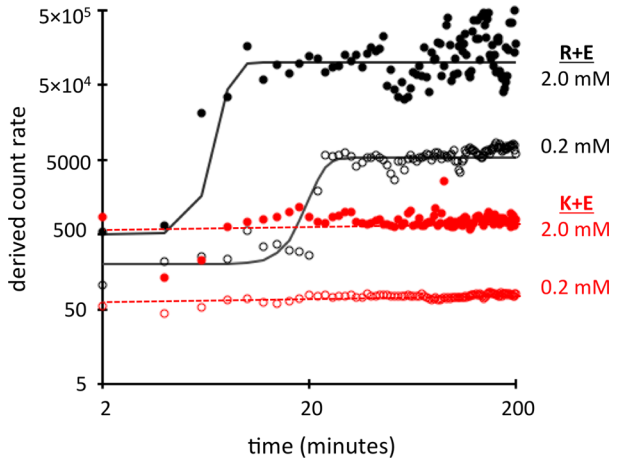


Figure 3. Nucleation and growth of R6 + E6 aggregates monitored by DLS. K6 + E6 peptide mixtures did not show any evidence of assembly over several hours. Solid black lines indicate best model fits to a two-step kinetic scheme.[12] Parameters of the WF model (eq 1) $k_1 = 2.4 \times 10^{-5}$ min^{-1} for both fits, $k_2 = 1 \times 10^{-4}$ M^{-1} min^{-1} for [R + E] = 0.2 mM, and 3×10^{-5} M^{-1} min^{-1} for [R + E] = 2 mM. The derived count rate is the calculated scattering intensity in photon kilo-counts per second.

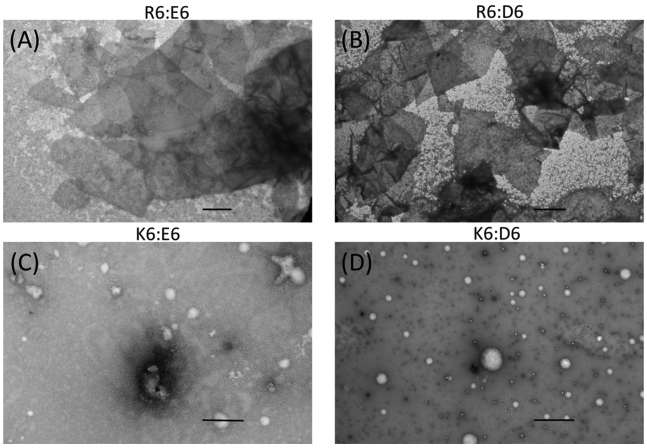


Figure 4. TEM of supercharged peptide aggregates. R6 combined with (A) E6 and (B) D6 form tens-of-microscale sheets (scale bar = 2 μm). K6 mixed with (C) E6 or (D) D6 formed particulates structures potentially induced by aggregation during drying.

solution DLS measurements indicating they may have formed during deposition and drying on the EM grid.

Long-Range Electrostatic Forces. For the solution composition and temperature used in this work, the Debye length over which long-range electrostatic forces can mediate intermolecular interactions would be approximately 1.2 nm. Electrophoretic mobilities of the peptides were determined by DLS and modeled as apparent zeta potentials (ζ) using the Smoluchowski equation, representing the electrostatic potential at the slip-plane of the double-layer surrounding the peptide macro-ion. R6 and K6 both showed $\zeta > 0$ mV with nominal differences in their relative magnitudes (Figure 2B). The anionic D6 and E6 peptides showed large negative ζ values, and $\zeta \approx 0$ mV for the neutral $(\mathrm{POG})_{10}$ peptide. These observations indicate that long-range electrostatic phenomena on nanometer length scales are unlikely to be responsible for the difference in the aggregation behavior of R6 and K6.

Effect of Charge Position on Folding and Assembly. It has been demonstrated in multiple host−guest studies on CMPs

DOI: 10.1021/jacs.5b10304
J. Am. Chem. Soc. 2016, 138, 4362−4367

that an amino acid has different physical properties at the *x* versus *y* position in the Gly-*x*-*y* triplet repeat.[1a,d,e,14] In supercharged CMPs, one cannot readily determine which positions are promoting assembly. This is further complicated by the poor structure and stability of D6 and E6. Therefore, an additional set of peptides were constructed comprising eight variations on D, E, R, and K at *x* or *y* positions.

PRG3: $(PRG)_3(POG)_7$; **ROG3**: $(ROG)_3(POG)_7$

PKG3: $(PKG)_3(POG)_7$; **KOG3**: $(KOG)_3(POG)_7$

PEG3: $(PEG)_3(POG)_7$; **EOG3**: $(EOG)_3(POG)_7$

PDG3: $(PDG)_3(POG)_7$; **DOG3**: $(DOG)_3(POG)_7$

All eight peptides formed well folded homotrimers as shown by a clear cooperative unfolding transition monitored by CD (Figure S6). The lowest stability species was PDG3 with a T_m = 31 °C. Significant position-specific effects on stability were observed; e.g., T_m of KOG3 = 55 °C versus PKG3 = 40 °C.

Sixteen combinations of cationic and anionic components were assessed for ability to form higher-order structure. In nearly all cases no assembly was observed. The only exceptions were PRG3 + PEG3 and ROG3 + PEG3 (Figure 5). Unlike the R6 + D6/E6 interaction, where assembly occurred in under an hour at total peptide concentrations from 0.2 to 2 mM, the PEG3 + PRG3/ROG3 assembly was only observed after incubating mixtures for several days at 4 °C, and at 2 mM total concentration. Instead of two-dimensional sheets, these peptides form a colloidal aggregate (Figure S7).

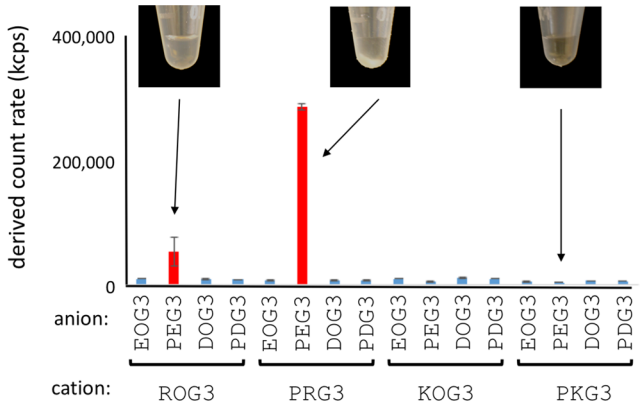


Figure 5. DLS of cation−anion peptide mixtures containing charged residues at *x* or *y* positions. Assembly is only observed in ROG3 + PEG3 and PRG3 + PEG3 mixtures. Visual inspection of ROG3 + PEG3 shows a cloudy suspension. PRG3 + PEG3 settles as an opaque aggregate. For comparison, PKG3+ PEG3 solution is clear. Images of all mixtures in Figure S8.

DISCUSSION

From the perspective of the cation, only arginine-containing peptides participate in extensive assembly in all three two-component systems. Compared to ROG3, PRG3 is the more stable peptide and the more commonly occurring triplet in fibrillar animal collagens. In the case of the anion, both amino acid type and position-specific effects are observed. PEG3 is less stable than EOG3, and as a triplet is less abundant in animal collagen sequences. It had been previously shown that many CMPs aggregate close to their melting temperature, suggesting that assembly proceeds through a molten, partially unfolded triple-helix transition state.[11] This model does not describe electrostatically driven assembly of these components where the high-stability PRG3 assembles most readily with the low-stability PEG3. In other systems, a repeating EOG triplet is used for the anionic domains,[4a,d,8] resulting in structured assemblies. These peptides have four or more repeats and contain both cationic and anionic domains in the same peptide, resulting in a greater number of potential interhelical interactions per helix.

Insight into the mechanism underlying the complex dependence of assembly on amino acid type and position may be gained by considering the larger problem of what drives protein−protein interactions. In early work by Janin and Chothia, it was observed that arginine is the most abundant charged residue promoting interactions between protein subunits.[15] This analysis was extended by Jones and Thornton to 50 high-resolution structures of protein−protein interfaces in the Protein Data Bank (PDB),[16] where arginine was statistically overrepresented at interfaces (Figure 6A). In the same analysis, lysine and aspartate were the two most unfavorable interfacial residues of all 20 amino acids. The aggregation preferences of arginine and glutamate over lysine and aspartate across CMPs are fully consistent with the relative frequencies of these amino acids at protein interfaces.

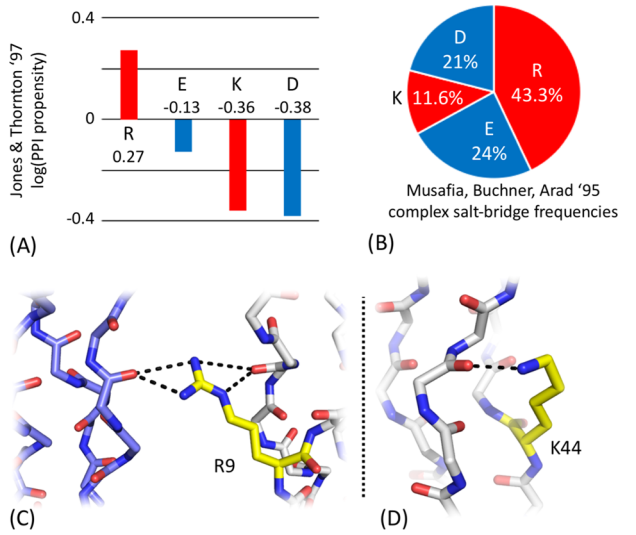


Figure 6. (A) Interfacial propensities of charged amino acids from a survey of high-resolution structures of protein−protein interfaces (PPIs). Positive values represent favorable interface propensity. Data excerpted from ref 16. (B) Frequencies of charged amino acids connecting two other groups in a complex salt-bridge. Data excerpted from ref 18. (C) Arginine (PDB ID: 4GYX[20]) involved intra- and interhelical hydrogen bonds with backbone carbonyls. (D) Lysine (PDB ID: 1QSU[21]) involved in only an intrahelical hydrogen bond. Other examples are shown in Figure S9.

The assembly prone nature of arginine may be due to the capacity of the polyvalent side chain to form intermolecular contacts without necessarily sacrificing intramolecular interactions.[17] Analysis of approximately 1000 charge pairs in 94 proteins by Musafia et al. showed that arginine was four-fold more likely than lysine to participate in a complex network of electrostatic interactions involving three or more residues[18] (Figure 6B). In these networks, arginine was at the center making multiple hydrogen bonds through its three guanidyl nitrogens. Only arginine and glutamate were found in cases where one residue was simultaneously contacting three others. A more recent analysis of charged residue networks with a larger and

DOI: 10.1021/jacs.5b10304
J. Am. Chem. Soc. 2016, 138, 4362−4367

higher quality data set[19] confirms the observations of the Musafia et al.[18] study.

Existing studies of the PDB indicate arginine has a higher propensity than lysine to be found at protein−protein interfaces and to participate in networks of electrostatic interactions. The question then arises whether such networks occur at interfaces between triple-helices. There are no available high-resolution structures of self-assembled charged CMPs in their fiber or sheet form. Instead, a handful (seven) CMP structures determined by X-ray crystallography exist that include arginine or lysine. Crystal lattice packing contacts involving arginine or lysine may be considered as surrogates for contacts formed during aggregation and reveal how such residues control assembly. Each of the four arginine-containing CMP structures showed interhelical interactions across crystal lattice interfaces. A common feature of such interfaces was the simultaneous participation of the polyvalent guanidyl moiety of arginine in networks of hydrogen bonds with its helix and with neighboring helices in the crystal lattice (Figure 6C and S9). In contrast, the three unique lysine containing structures showed that the side chain amine either participated in intra or interhelical contacts with backbone carbonyls, but never both for the same side chain (Figure 6D and S9). The strongest arginine interactions as measured by donor−acceptor heavy atom separations <3.0 Å were with other protein groups. In contrast, the tightest lysine-mediated contacts were with bound waters, consistent with a barrier to desolvation of the side chain amine at protein interfaces.

In this context, we can begin to understand the diverse supramolecular morphologies of $(PxG)_a$-$(POG)_b$-$(yOG)_c$ CMPs described in the literature. When x = D and y = K, the lysine side chain prefers intrahelical contacts, promoting a sticky-ended triple-helix.[4b,e] Further assembly into an infinite triple-helix is mediated by backbone hydrogen bonds. For y = R, the nanosheet[4d,8] or banded fiber[4a] morphology arises from arginine participating in both intrahelical and interhelical interactions. Structural plasticity is seen outside of CMPs in other self-assembling systems, such as recently described α-helical nanotubes.[7b] In this system, replacing a single critical arginine with lysine dramatically alters nanotube morphology. High-resolution cryo-EM structures show this arginine participates in a complex network of supramolecular interactions with the carboxy termini of neighboring α-helices.

The nonequivalency of lysine and arginine extends to numerous other electrostatic phenomena in biology. Oligo-arginine peptides interact more tightly with nucleic acids than oligo-lysine.[22] Rather than aggregate in the presence of anions, oligo-lysine will often form a coacervate,[23] a liquid phase separation of the cationic and anionic species from bulk solution where the desolvation barrier prevents further assembly. The particulate structures observed by TEM in the case of K6 + D6 or E6 peptide mixtures may be the result of such coacervates drying out and solidifying on the EM grid. A number of PDB analysis studies have shown different propensities for arginine and lysine at interfaces.[16,24] Lysine to arginine substitutions on the surfaces of proteins can be used as a strategy for promoting protein crystallization for structure solution.[25] Arginine peptides and proteins are better able to insert and penetrate lipid bilayers than lysine,[26] a property that has been exploited in the design of cell-penetrating peptides and antimicrobial peptides.[27]

Charge-complementarity has been used from the earliest peptide designs to current supramolecular systems. Computational methods are now being developed to automate supramolecular peptide design[28] and explicit consideration of valence and hydrophobicity of arginine and lysine in coarse-grained[29] and atomistic potentials will be critical. The multicomponent nature of these peptide systems provide powerful platforms for dissecting intermolecular forces and developing modeling tools.

EXPERIMENTAL SECTION

Peptide Synthesis. Peptides were synthesized using solid-phase FMOC chemistry at the Tufts University Core Facility, Boston, MA. N- and C-termini were acetylated and amidated, respectively. Peptides were purified to 95% purity by reverse-phase high-performance liquid chromatography (HPLC), and products were verified by mass spectrometry. All peptide solutions were prepared in 10 mM phosphate buffer pH = 7.0. Peptide concentrations in solution were estimated by obtaining the absorbance at 214 nm using ε_{214} = 2200 M^{-1} cm^{-1}. After mixtures were prepared at room temperature, they were allowed to fold at 4 °C for at least 48 h.

Circular Dichroism Spectroscopy. CD measurements were conducted using the Aviv Model 420SF Spectrophotometer equipped with a Peltier temperature controller. For wavelength spectra, measurements were made at every 0.5 nm step with an averaging time of 10 s at each wavelength. Scans were conducted from 190 to 260 nm at 5 °C. Observed ellipticity was converted to molar ellipticity by dividing raw values by the peptide concentration, number of residues, and cell path length. For temperature induced denaturation, ellipticity was measured at 225 nm for $(POG)_{10}$ or at 223 nm for charged peptides. Total peptide concentrations used were 0.2 mM. CD temperature denaturation measurements were performed at 0.33 °C/step and 2 min equilibration time.

Light Scattering. Dynamic light scattering (DLS) measurements were performed using a Zetasizer Nano ZS (Malvern Instruments, U.K.). Data was collected with a 3 mW He−Ne laser at a 633 nm wavelength. This unit collects back scattered light at an angle of 173° and contains a built-in Peltier element temperature controller. Autocorrelation functions were determined with an acquisition time of 2 min per correlation function. For kinetics measurement, intensity autocorrelation functions of scattered light were collected continuously using acquisition times of 120 s per correlation function for at least 12 h for each sample. Derived count rates were calculated in units of kilocounts per second (kcps) assuming no attenuation of the laser light source by the instrument. Reported viscosity values[30] were used for the hydrodynamic radius calculation. Correlation plots are in Figure S10. A detailed description of DLS can be found elsewhere.[31]

Zeta Potential. Zeta potential measurements were conducted with the same Zetasizer Nano ZS (Malvern Instruments, U.K.). The instrument uses a PALS (Phase Analysis Light Scattering) method to measure the electrophoretic mobility of particles in solution. The Smoluchowshi equation is used to calculate the zeta potential of the particles in solution. Measurements were done in disposable folded capillary cells. Each measurement was the average of 100 runs 10 s each. Five measurements were conducted for each sample, and the final result is the average of these measurements.

Transmission Electron Microscopy. Samples were gently resuspended and placed onto copper grids, then negatively stained with 1% phosphotungstic acid and imaged with a Philip 420 Electron Microscope.

Structure Analysis. Relevant structures from the PDB were analyzed for crystal lattice contacts: 1BKV,[32] 1Q7D,[33] 3DMW,[34] 4GYX,[20] 1QSU,[21] 3PON,[35] 3POD.[35] Neighboring asymmetric units were constructed using the symmetry transformations provided with the structure. Potential hydrogen bonding and electrostatic interactions were identified in the pyMol structure modeling platform if interatomic distances ≤4 Å.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b10304.

DOI: 10.1021/jacs.5b10304
J. Am. Chem. Soc. 2016, 138, 4362−4367

Detailed description of methods and supporting data (PDF)

AUTHOR INFORMATION

Corresponding Author

*nanda@cabm.rutgers.edu

Author Contributions

¶A.S.P. and J.K.J. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank Ife Aridegbe for assistance in analyzing collagen peptide structures in the PDB. We gratefully acknowledge support from NSF DMR-0907273, NIH DP2-OD-006478-1 to carry out this work. J.K.J. was supported by Rutgers Biotechnology Training Program (NIH Grant Number 5T32GM008339-20). A.S.P. acknowledges support from Department of Science & Technology (DST-SERB) - SB/FTP/PS-073/2014, India.

REFERENCES

(1) (a) Persikov, A. V.; Ramshaw, J. A.; Brodsky, B. *J. Biol. Chem.* **2005**, *280*, 19343. (b) Persikov, A. V.; Ramshaw, J. A.; Kirkpatrick, A.; Brodsky, B. *Biochemistry* **2005**, *44*, 1414. (c) Persikov, A. V.; Ramshaw, J. A.; Kirkpatrick, A.; Brodsky, B. *J. Mol. Biol.* **2002**, *316*, 385. (d) Ramshaw, J. A.; Shah, N. K.; Brodsky, B. *J. Struct. Biol.* **1998**, *122*, 86. (e) Chan, V. C.; Ramshaw, J. A.; Kirkpatrick, A.; Beck, K.; Brodsky, B. *J. Biol. Chem.* **1997**, *272*, 31441. (f) Yang, W.; Chan, V. C.; Kirkpatrick, A.; Ramshaw, J. A.; Brodsky, B. *J. Biol. Chem.* **1997**, *272*, 28837. (g) Venugopal, M. G.; Ramshaw, J. A.; Braswell, E.; Zhu, D.; Brodsky, B. *Biochemistry* **1994**, *33*, 7948. (h) Hulmes, D. J.; Miller, A.; Parry, D. A.; Piez, K. A.; Woodhead-Galloway, J. *J. Mol. Biol.* **1973**, *79*, 137.

(2) (a) Gauba, V.; Hartgerink, J. D. *J. Am. Chem. Soc.* **2007**, *129*, 15034. (b) Gauba, V.; Hartgerink, J. D. *J. Am. Chem. Soc.* **2007**, *129*, 2683.

(3) (a) Leikin, S.; Rau, D. C.; Parsegian, V. A. *Nat. Struct. Biol.* **1995**, *2*, 205. (b) Leikin, S.; Rau, D. C.; Parsegian, V. A. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 276. (c) Freudenberg, U.; Behrens, S. H.; Welzel, P. B.; Muller, M.; Grimmer, M.; Salchert, K.; Taeger, T.; Schmidt, K.; Pompe, W.; Werner, C. *Biophys. J.* **2007**, *92*, 2108.

(4) (a) Rele, S.; Song, Y.; Apkarian, R. P.; Qu, Z.; Conticello, V. P.; Chaikof, E. L. *J. Am. Chem. Soc.* **2007**, *129*, 14780. (b) O'Leary, L. E.; Fallas, J. A.; Bakota, E. L.; Kang, M. K.; Hartgerink, J. D. *Nat. Chem.* **2011**, *3*, 821. (c) Xu, F.; Li, J.; Jain, V.; Tu, R. S.; Huang, Q.; Nanda, V. *J. Am. Chem. Soc.* **2012**, *134*, 47. (d) Jiang, T.; Xu, C.; Liu, Y.; Liu, Z.; Wall, J. S.; Zuo, X.; Lian, T.; Salaita, K.; Ni, C.; Pochan, D.; Conticello, V. P. *J. Am. Chem. Soc.* **2014**, *136*, 4300. (e) Jalan, A. A.; Jochim, K. A.; Hartgerink, J. D. *J. Am. Chem. Soc.* **2014**, *136*, 7535. (f) Sarkar, B.; O'Leary, L. E.; Hartgerink, J. D. *J. Am. Chem. Soc.* **2014**, *136*, 14417. (g) Krishna, O. D.; Kiick, K. L. *Biomacromolecules* **2009**, *10*, 2626.

(5) (a) Summa, C. M.; Rosenblatt, M. M.; Hong, J. K.; Lear, J. D.; DeGrado, W. F. *J. Mol. Biol.* **2002**, *321*, 923. (b) Nautiyal, S.; Woolfson, D. N.; King, D. S.; Alber, T. *Biochemistry* **1995**, *34*, 11645. (c) Xu, F.; Zahid, S.; Silva, T.; Nanda, V. *J. Am. Chem. Soc.* **2011**, *133*, 15260. (d) Xu, F.; Zhang, L.; Koder, R. L.; Nanda, V. *Biochemistry* **2010**, *49*, 2307. (e) Graddis, T. J.; Myszka, D. G.; Chaiken, I. M. *Biochemistry* **1993**, *32*, 12664.

(6) (a) Vizcarra, C. L.; Mayo, S. L. *Curr. Opin. Chem. Biol.* **2005**, *9*, 622. (b) Jaramillo, A.; Wodak, S. J. *Biophys. J.* **2005**, *88*, 156.

(7) (a) Mrabet, N. T.; Van den Broeck, A.; Van den brande, I.; Stanssens, P.; Laroche, Y.; Lambeir, A. M.; Matthijssens, G.; Jenkins, J.; Chiadmi, M.; van Tilbeurgh, H.; Rey, F.; Janin, J.; Quax, W. J.; Lasters, I.; De Maeyer, M.; Wodak, S. J. *Biochemistry* **1992**, *31*, 2239. (b) Egelman, E. H.; Xu, C.; DiMaio, F.; Magnotti, E.; Modlin, C.; Yu, X.; Wright, E.; Baker, D.; Conticello, V. P. *Structure* **2015**, *23*, 280.

(8) (a) Jiang, T.; Vail, O. A.; Jiang, Z.; Zuo, X.; Conticello, V. P. *J. Am. Chem. Soc.* **2015**, *137*, 7793. (b) Jiang, T.; Xu, C.; Zuo, X.; Conticello, V. P. *Angew. Chem., Int. Ed.* **2014**, *53*, 8367.

(9) (a) Keshwani, N.; Banerjee, S.; Brodsky, B.; Makhatadze, G. I. *Biophys. J.* **2013**, *105*, 1681. (b) Gurry, T.; Nerenberg, P. S.; Stultz, C. M. *Biophys. J.* **2010**, *98*, 2634.

(10) Parmar, A. S.; Joshi, M.; Nosker, P. L.; Hasan, N. F.; Nanda, V. *Biomolecules* **2013**, *3*, 986.

(11) Kar, K.; Amin, P.; Bryan, M. A.; Persikov, A. V.; Mohs, A.; Wang, Y. H.; Brodsky, B. *J. Biol. Chem.* **2006**, *281*, 33283.

(12) Morris, A. M.; Watzky, M. A.; Agar, J. N.; Finke, R. G. *Biochemistry* **2008**, *47*, 2413.

(13) (a) Przybyla, D. E.; Rubert Perez, C. M.; Gleaton, J.; Nandwana, V.; Chmielewski, J. *J. Am. Chem. Soc.* **2013**, *135*, 3418. (b) Przybyla, D. E.; Chmielewski, J. *J. Am. Chem. Soc.* **2010**, *132*, 7866. (c) McGuinness, K.; Khan, I. J.; Nanda, V. *ACS Nano* **2014**, *8*, 12514. (d) Xu, F.; Khan, I. J.; McGuinness, K.; Parmar, A. S.; Silva, T.; Murthy, N. S.; Nanda, V. *J. Am. Chem. Soc.* **2013**, *135*, 18762.

(14) Persikov, A. V.; Ramshaw, J. A.; Kirkpatrick, A.; Brodsky, B. *Biochemistry* **2000**, *39*, 14960.

(15) Janin, J.; Miller, S.; Chothia, C. *J. Mol. Biol.* **1988**, *204*, 155.

(16) Jones, S.; Thornton, J. M. *J. Mol. Biol.* **1997**, *272*, 121.

(17) Borders, C. L., Jr.; Broadwater, J. A.; Bekeny, P. A.; Salmon, J. E.; Lee, A. S.; Eldridge, A. M.; Pett, V. B. *Protein Sci.* **1994**, *3*, 541.

(18) Musafia, B.; Buchner, V.; Arad, D. *J. Mol. Biol.* **1995**, *254*, 761.

(19) Donald, J. E.; Kulp, D. W.; DeGrado, W. F. *Proteins: Struct., Funct., Bioinf.* **2011**, *79*, 898.

(20) Boudko, S. P.; Bachinger, H. P. *J. Biol. Chem.* **2012**, *287*, 44536.

(21) Kramer, R. Z.; Venugopal, M. G.; Bella, J.; Mayville, P.; Brodsky, B.; Berman, H. M. *J. Mol. Biol.* **2000**, *301*, 1191.

(22) (a) DeRouchey, J.; Hoover, B.; Rau, D. C. *Biochemistry* **2013**, *52*, 3000. (b) Mascotti, D. P.; Lohman, T. M. *Biochemistry* **1997**, *36*, 7272.

(23) (a) Dora Tang, T. Y.; Rohaida Che Hak, C.; Thompson, A. J.; Kuimova, M. K.; Williams, D. S.; Perriman, A. W.; Mann, S. *Nat. Chem.* **2014**, *6*, 527. (b) Priftis, D.; Laugel, N.; Tirrell, M. *Langmuir* **2012**, *28*, 15947.

(24) (a) Pednekar, D.; Tendulkar, A.; Durani, S. *Proteins: Struct., Funct., Genet.* **2009**, *74*, 155. (b) Bogan, A. A.; Thorn, K. S. *J. Mol. Biol.* **1998**, *280*, 1. (c) Ben-Shimon, A.; Eisenstein, M. *J. Mol. Biol.* **2010**, *402*, 259. (d) Jones, S.; Thornton, J. M. *J. Mol. Biol.* **1997**, *272*, 133. (e) Jones, S.; Thornton, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 13.

(25) Czepas, J.; Devedjiev, Y.; Krowarsch, D.; Derewenda, U.; Otlewski, J.; Derewenda, Z. S. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 275.

(26) (a) Han, K.; Jeon, M. J.; Kim, S. H.; Ki, D.; Bahn, J. H.; Lee, K. S.; Park, J.; Choi, S. Y. *Mol. Cells* **2001**, *12*, 267. (b) Sakai, N.; Matile, S. *J. Am. Chem. Soc.* **2003**, *125*, 14348.

(27) Veiga, A. S.; Sinthuvanich, C.; Gaspar, D.; Franquelim, H. G.; Castanho, M. A.; Schneider, J. P. *Biomaterials* **2012**, *33*, 8907.

(28) (a) Frederix, P. W.; Scott, G. G.; Abul-Haija, Y. M.; Kalafatovic, D.; Pappas, C. G.; Javid, N.; Hunt, N. T.; Ulijn, R. V.; Tuttle, T. *Nat. Chem.* **2015**, *7*, 30. (b) Smadbeck, J.; Chan, K. H.; Khoury, G. A.; Xue, B.; Robinson, R. C.; Hauser, C. A.; Floudas, C. A. *PLoS Comput. Biol.* **2014**, *10*, e1003718.

(29) de Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schafer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. *J. Chem. Theory Comput.* **2013**, *9*, 687.

(30) Parmar, A. S.; Muschol, M. *J. Colloid Interface Sci.* **2009**, *339*, 243.

(31) (a) Parmar, A. S.; Gottschall, P. E.; Muschol, M. *Biophys. Chem.* **2007**, *129*, 224. (b) Parmar, A. S.; Muschol, M. *Biophys. J.* **2009**, *97*, 590.

(32) Kramer, R. Z.; Bella, J.; Mayville, P.; Brodsky, B.; Berman, H. M. *Nat. Struct. Biol.* **1999**, *6*, 454.

(33) Emsley, J.; Knight, C. G.; Farndale, R. W.; Barnes, M. J. *J. Mol. Biol.* **2004**, *335*, 1019.

(34) Boudko, S. P.; Engel, J.; Okuyama, K.; Mizuno, K.; Bachinger, H. P.; Schumacher, M. A. *J. Biol. Chem.* **2008**, *283*, 32580.

(35) Gingras, A. R.; Girija, U. V.; Keeble, A. H.; Panchal, R.; Mitchell, D. A.; Moody, P. C.; Wallis, R. *Structure* **2011**, *19*, 1635.

DOI: 10.1021/jacs.5b10304
J. Am. Chem. Soc. 2016, 138, 4362−4367