Supplemental Information: The Role of Hydrophobicity in the Stability and pH-Switchability of (RXDX)$_4$ β-sheets and Coumarin-(RXDX)$_4$ Conjugates

Ryan Weber[†] and Martin McCullagh[*,‡]

†*Department of Chemistry, Colorado State University, Fort Collins, Colorado 80523, United States*

‡*Department of Chemistry, Oklahoma State University, Stillwater, Oklahoma 74074, United States*

E-mail: martin.mccullagh@okstate.edu

Hydrophobic Residue Choice

We generalize the amino acid sequence to $(RXDX)_4$, where five different hydrophobic amino acids (Ala, Val, Leu, Ile and Phe) are substituted at the X position to investigate the effect of changing the hydrophobic group on both stability and pH-switchability of $(RXDX)_4$ β-sheet sandwich fibers. This set of five residues is chosen due to their incremental change in hydrophobicity, based on water-octanol partition coefficients,[1] without adding additional functional groups which could affect the self-assembled behavior of the material. These values along with β-sheet propensity for these amino acids are provided in Table S1.

Table S1. $(RXDX)_4$ peptide sequences, hydrophobicity, and secondary structure propensity.

Sequence	X	Hydrophobicity (π)[a]	β-sheet Propensity (P_β)[b]
$(RADA)_4$	Ala	0.31	0.75
$(RVDV)_4$	Val	1.22	1.86
$(RLDL)_4$	Leu	1.70	1.32
$(RIDI)_4$	Ile	1.80	1.71
$(RFDF)_4$	Phe	1.79	1.43

[a] Amino Acid hydrophobicity based on water-octanol partition coefficients relative to glycine.[1]
[b] Propensity to occur in β-sheet secondary structures from the PDBselect dataset.[2]

Force Field Verification

To verify our force field we compare the thickness of each each $(RXDX)_4$ fiber at neutral pH calculated from simulation with atomic force microscopy measurements published by Bagrov *et al.*,[3] as shown in Table S2. The fiber in the simulation is aligned such that the principal component vector that is perpendicular to the plane of the 2 β-sheets is parallel to the x-axis of the simulation box. The thickness is calculated as the distance from the end of each central arginine sidechain on one β-sheet to the end of each central arginine on the other β-sheet projected along the x-axis. The center of mass of the NH1 and NH2 atoms is

used as the position of the end of the arginine sidechain. Arginine residues on the outer two β-strands on both ends of each β-sheet and all N-termini arginines are excluded from these calculations as these parts of the β-sheet sandwich are less stable. The thickness is calculated every 100th frame and averaged over the entire 250 ns of all-atom molecular dynamics simulation. The calculated thickness of $(RXDX)_4$ fibers with X=Ala and Leu are in good agreement with experimental results.

Table S2. Height of the RXDX β-sheet sandwich.

System	Height (Å)				
	RADA	RVDV	RLDL	RIDI	RFDF
Calculated	18.357(3)	20.718(3)	23.771(4)	22.170(5)	22.34(2)
Experimental	19(1)[3]	-	23(4)[3]	-	-

Additionally, we simulated 100 coumarins in water for 400 ns. From these simulations the solubility of coumarin in water is calculated to be 0.014 ± 0.006 $\frac{mol}{L}$, in agreement with the solubility value of 0.01706 $\frac{mol}{L}$ reported by Yalkowsky *et al.*[4]

Coulombic Interaction Energies

The coulombic interaction energies broken down into residue type pairs was calculated for each fiber to understand the interactions which lead to the pH-switchability of the $(RXDX)_4$ fibers. Below are the coulombic interaction energies for the $(RVDV)_4$, $(RLDL)_4$, $(RIDI)_4$, and $(RFDF)_4$ fibers.

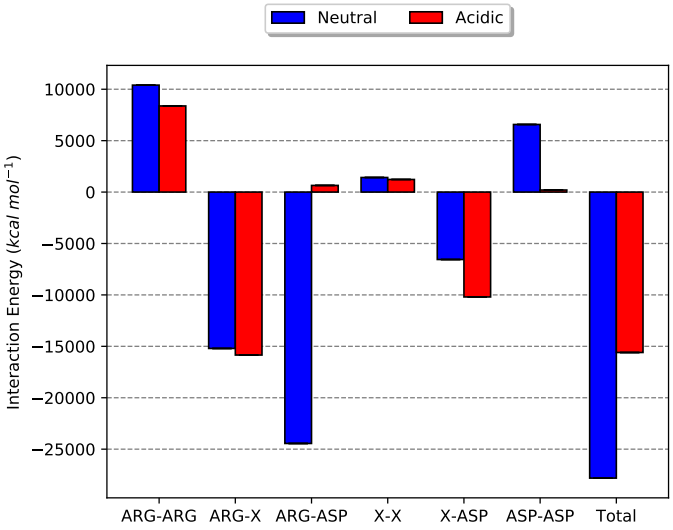

Figure S1. Coulombic interaction energy decomposed into residue type pairs for $(RVDV)_4$ fibers at neutral and acidic pH. The total coulombic interactions are large and attractive stabilizing the fiber. Under acidic conditions there is a 11967 kcal/mol increase in the total coulombic interaction energy.

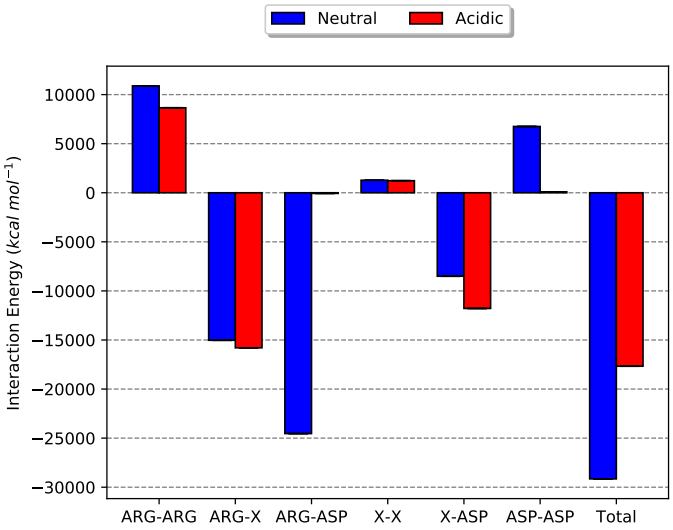

Figure S2. Coulombic interaction energy decomposed into residue type pairs for $(RLDL)_4$ fibers at neutral and acidic pH. The total coulombic interactions are large and attractive stabilizing the fiber. Under acidic conditions there is a 11480 kcal/mol increase in the total coulombic interaction energy.

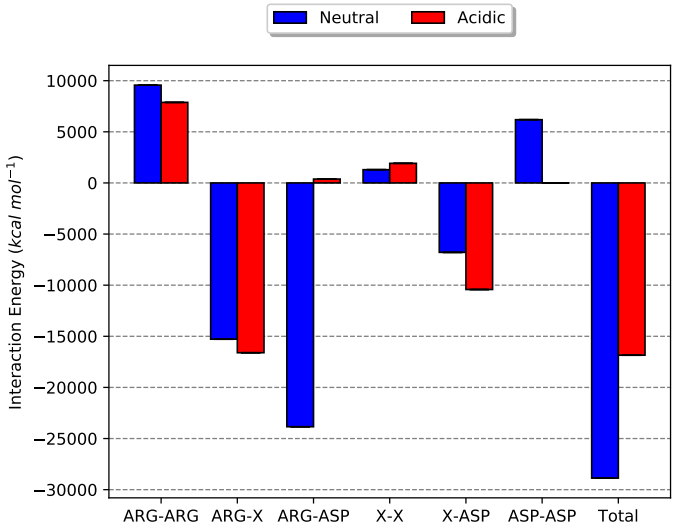

Figure S3. Coulombic interaction energy decomposed into residue type pairs for $(RIDI)_4$ fibers at neutral and acidic pH. The total coulombic interactions are large and attractive stabilizing the fiber. Under acidic conditions there is a 12023 kcal/mol increase in the total coulombic interaction energy.

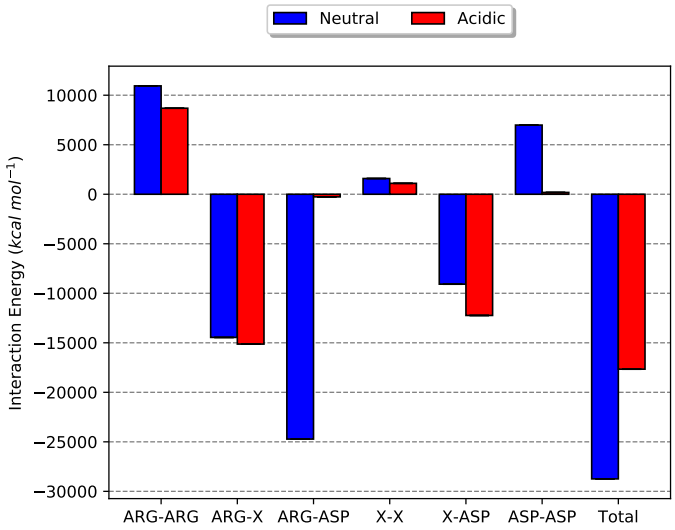

Figure S4. Coulombic interaction energy decomposed into residue type pairs for $(RFDF)_4$ fibers at neutral and acidic pH. The total coulombic interactions are large and attractive stabilizing the fiber. Under acidic conditions there is a 11077 kcal/mol increase in the total coulombic interaction energy.

Coumarin Ordering

The joint probability density as a function of both the coumarin-coumarin separation distance and the angle between the dipole moments of the two coumarin is used to measure the organization of coumarin within coumarin-$(RXDX)_4$ fibers. Below are the joint probability densities for coumarin-$(RADA)_4$ and coumarin-$(RVDV)_4$ fibers.

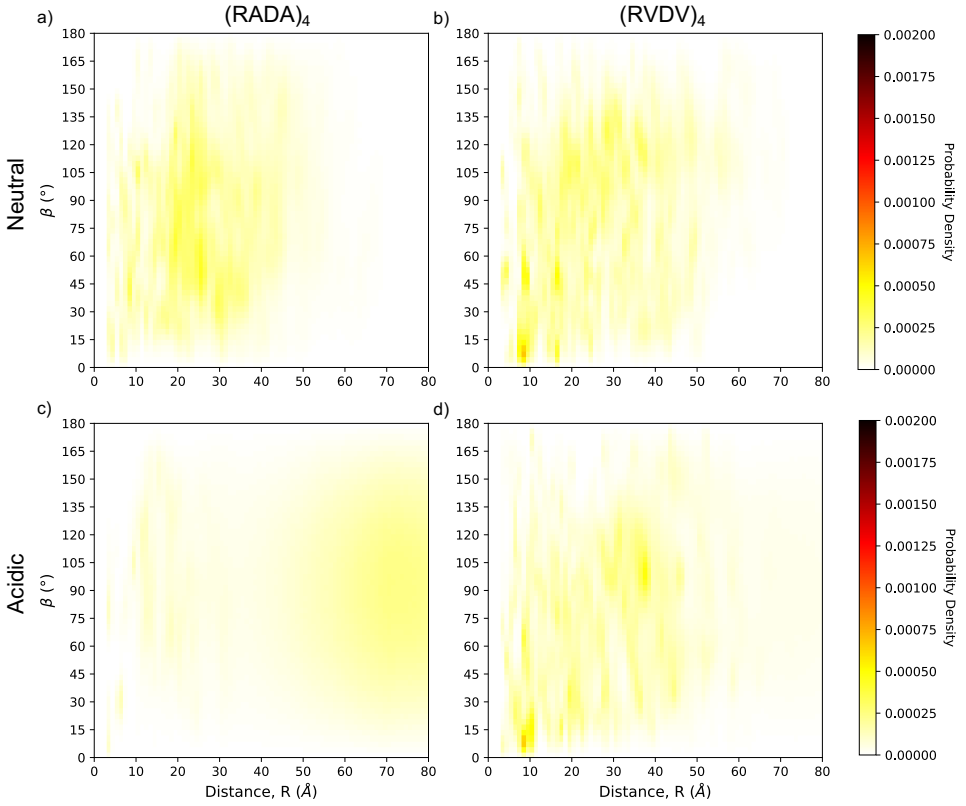

Figure S5. Joint probability distributions of all coumarin sidechain dimers as a function of separation distance and angle between their respective dipole moments for (a,c) $(RADA)_4$ and (b,d) $(RVDV)_4$ at neutral and acidic pH, respectively.

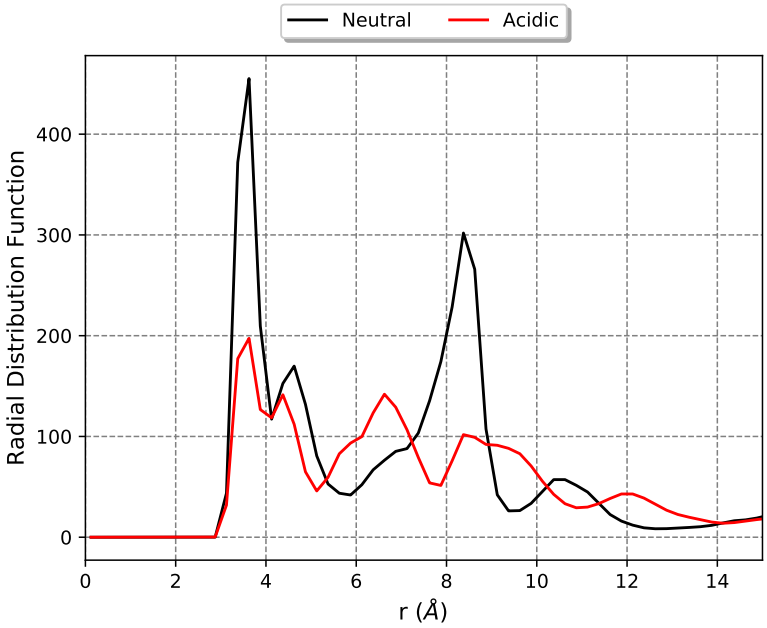

Figure S6. Radial distribution function of coumarin sidechain dimers as a function of separation distance at neutral and acidic pH.

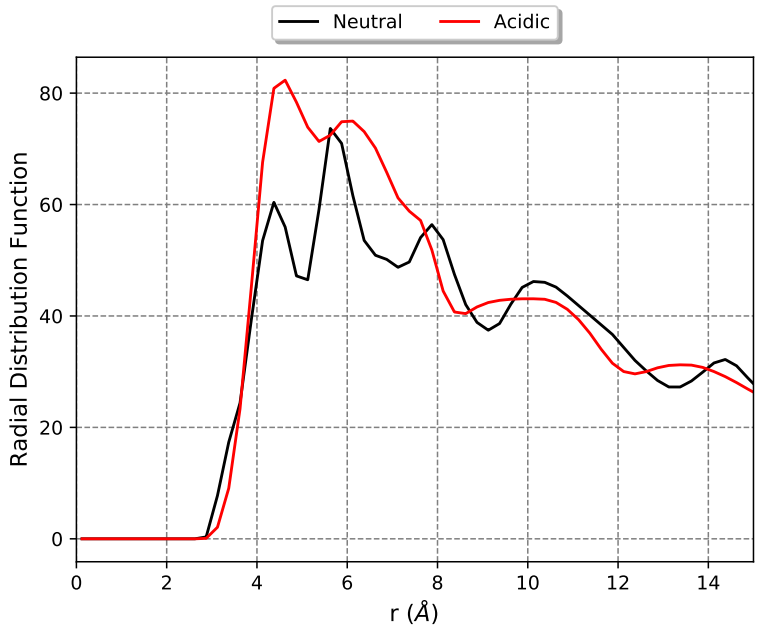

Figure S7. Radial distribution function of coumarin and phenylalanine sidechains as a function of separation distance at neutral and acidic pH.

Pressure Equilibration

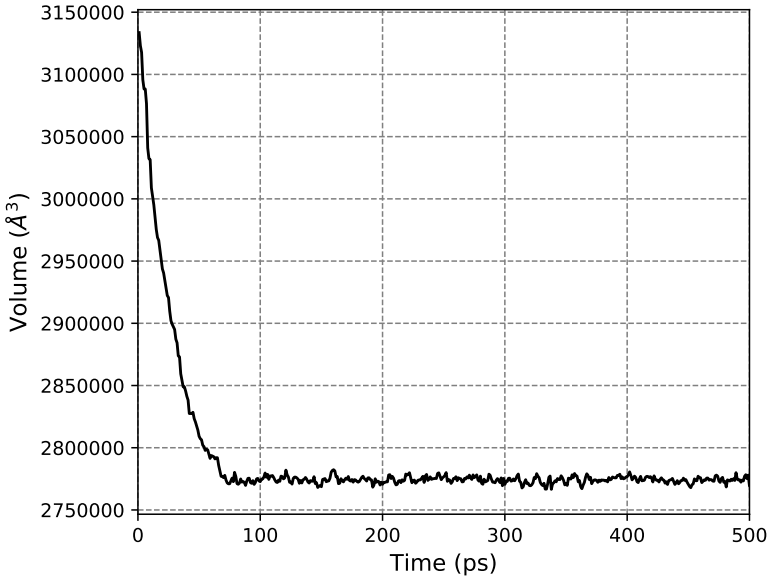

Figure S8. Timetrace of the volume of the simulation box for the coumarin-$(RADA)_4$ system (replica 1). The volume and, therefore, pressure become equilibrated after 80 ps of simulation following the heating step.

Atom Parameters

Below are the atom and bonded parameters used for the unnatural coumarin amino acid. The bonds, angles, and dihedrals shown below are the parameters used for interacting atoms that are not within the same forcefield. All other parameters are provided by the ff15ipq and GAFF forcefields.

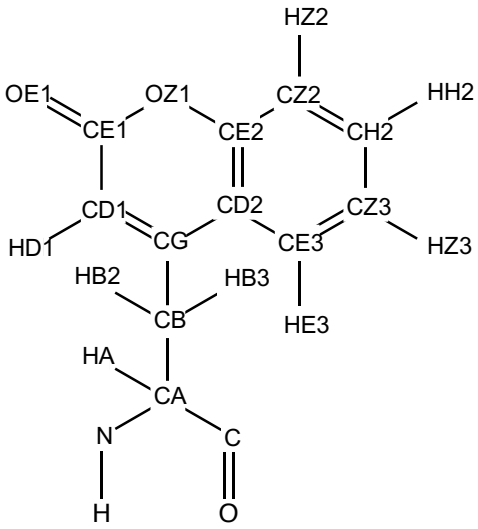

Figure S9. Atom names of unnatural coumarin amino acid

Table S3. Parameters for unnatural coumarin amino acid atoms.

Name	Type	Forcefield	Charge
O	OD	ff15ipq	-0.610490
C	C	ff15ipq	0.701560
CA	CX	ff15ipq	-0.146170
N	N	ff15ipq	-0.417020
H	H	ff15ipq	0.316160
HA	H1	ff15ipq	0.156000
CB	TA	ff15ipq	-0.240850
HB2	HC	ff15ipq	0.138730
HB3	HC	ff15ipq	0.138730
CG	c2	GAFF	0.239000
CD1	c2	GAFF	-0.520650
HD1	ha	GAFF	0.200630
CE1	c2	GAFF	0.895610
OE1	o	GAFF	-0.674080
OZ1	os	GAFF	-0.60182
CE2	ca	GAFF	0.473750
CZ2	ca	GAFF	-0.322010
HZ2	ha	GAFF	0.197470
CH2	ca	GAFF	-0.074560
HH2	ha	GAFF	0.316160
CZ3	ca	GAFF	-0.171280
HZ3	ha	GAFF	0.161480
CE3	ca	GAFF	-0.160930
HE3	ha	GAFF	0.159950
CD2	ca	GAFF	-0.155440

Table S4. Bond parameters for unnatural coumarin amino acid. The parameters below are used in an AMBER frcmod file.

Bond Atom Types	Force Constant ($\frac{kcal}{mol \cdot A^2}$)	Bond Length ($\mathring{A}$)
TA-c2	17.0000	1.5100

Table S5. Angle parameters for unnatural coumarin amino acid. The parameters below are used in an AMBER frcmod file.

Angle Atom Types	Force Constant ($\frac{kcal}{mol \cdot rad^2}$)	Equilibrium Angle (°)
CX-TA-c2	63.0000	114.00
TA-c2-c2	70.0000	120.00
TA-c2-ca	70.0000	120.00
HC-TA-c2	50.0000	109.50
o -c2-os	76.662	118.370
os-c2-o	76.662	118.370
c2-TA-HC	50.0000	109.50

Table S6. Dihedral parameters for unnatural coumarin amino acid. The parameters below are used in an AMBER frcmod file.

Dihedral Atom Types	Barrier Height	Phase Shift Angle(°)	Periodicity
C -CX-TA-c2	-0.22262	0.0	-4.0
C -CX-TA-c2	-1.19867	0.0	-3.0
C -CX-TA-c2	-0.03841	0.0	-2.0
C -CX-TA-c2	-0.17107	0.0	1.0
CX-TA-CA-ca	0.00000	0.0	-4.0
CX-TA-CA-ca	0.11558	0.0	-3.0
CX-TA-CA-ca	-0.39107	0.0	-2.0
CX-TA-CA-ca	-0.50022	0.0	1.0
CX-TA-CA-c2	0.00000	0.0	-4.0
CX-TA-CA-c2	0.11558	0.0	-3.0
CX-TA-CA-c2	-0.39107	0.0	-2.0
CX-TA-CA-c2	-0.50022	0.0	1.0
N -CX-TA-c2	-0.22082	0.0	-4.0
N -CX-TA-c2	0.13006	0.0	-3.0
N -CX-TA-c2	0.01776	0.0	-2.0
N -CX-TA-c2	-0.62013	0.0	1.0
H1-CX-TA-c2	0.38033	0.0	3.0
HC-TA-c2-c2	-0.46048	0.0	2.0
HC-TA-c2-ca	-0.46048	0.0	2.0
CX-TA-c2-c2	0.00000	0.0	-4.0
CX-TA-c2-c2	0.11558	0.0	-3.0
CX-TA-c2-c2	-0.39107	0.0	-2.0
CX-TA-c2-c2	-0.50022	0.0	1.0
CX-TA-c2-ca	0.00000	0.0	-4.0
CX-TA-c2-ca	0.11558	0.0	-3.0
CX-TA-c2-ca	-0.39107	0.0	-2.0
CX-TA-c2-ca	-0.50022	0.0	1.0

Table S7. Dihedral parameters for unnatural coumarin amino acid. The parameters below are used in an AMBER frcmod file.

Dihedral Atom Types	Barrier Height (÷2)	Phase Shift Angle(°)	Periodicity
TA-c2-c2-ca	1.1	180.0	2.0
c2-c2-c2-ha	1.1	180.0	2.0
c2-o -c2-os	1.1	180.0	2.0
ca-ca-ca-os	1.1	180.0	2.0
ca-ca-ca-ha	1.1	180.0	2.0

References

1. Fauchere, J. L.; Charton, M.; Kier, L. B.; Verloop, A.; Plliska, V. Amino acid side chain parameters for correlation studies in biology and pharmacology. *Int. J. Pept. Protein Res.* **1988**, *32*, 269–278.
2. Costantini, S.; Colonna, G.; Facchiano, A. M. Amino acid propensities for secondary structures are influenced by the protein structural class. *Biochem. Biophys. Res. Commun.* **2006**, *342*, 441–451.
3. Bagrov, D.; Gazizova, Y.; Podgorsky, V.; Udovichenko, I.; Danilkovich, A.; Prusakov, K.; Klinov, D. Morphology and aggregation of RADA-16-I peptide Studied by AFM, NMR and molecular dynamics simulations. *Biopolymers* **2016**, *106*, 72–81.
4. Yalkowsky, S. H.; He, Y.; Jain, P. *Handbook of Aqueous Solubility Data*, 2nd ed.; Taylor and Francis Group, LLC: Boca Raton, FL, 2010; p 558.