Investigation of a Catenane with a Responsive Noncovalent Network: Mimicking Long-Range Responses in Proteins

Mee-Kyung Chung, Peter S. White, Stephen J. Lee,[‡] Michel R. Gagné,* and Marcey L. Waters*

Department of Chemistry, CB 3290, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States
[‡]U.S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709, United States

S *Supporting Information*

ABSTRACT: We report a functional synthetic model for studying the noncovalent networks (NCNs) required for complex protein functions. The model [2]-catenane is self-assembled from dipeptide building blocks and contains an extensive network of hydrogen bonds and aromatic interactions. Perturbations to the catenane cause compensating changes in the NCNs structure and dynamics, resulting in long-distance changes reminiscent of a protein. Key findings include the notion that NCNs require regions of negative cooperativity, or "frustrated" noncovalent interactions, as a source of potential energy for driving the response. We refer to this potential energy as *latent free energy* and describe a mechanistic and energetic model for responsive systems.
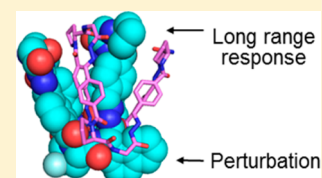
INTRODUCTION

Many protein functions depend on long-distance changes in response to a stimulus. Examples include allostery,[1] signal transduction,[2] and even binding through induced fit or conformational selection mechanisms[3] and catalysis.[4] Moreover, recent work has shown that this response can be structural (e.g., a change in the folded structure) or dynamic (e.g., a change in protein motion).[5] This type of communication through a protein structure is mediated by a noncovalent network (NCN) of interactions inherent in a folded protein.[6] Despite the importance of NCNs in protein function, molecular models for their mode(s) of action are lacking, in part due to the ambiguating complexity of proteins that utilize them for function. Furthermore, while communication through a network is a common mechanism in proteins, there are very few examples in synthetic systems, as small molecule models are generally insufficient to create functional NCNs.[7] While many molecular machines demonstrate switching capabilities between two states in response to a stimulus, these systems generally rely on the on/off switching of one or two dominant interactions rather than a network of weak noncovalent interactions as is typical of a protein. Thus, the ability to mimic, and hence utilize, actuation through a noncovalent network is in its infancy.

Herein, we utilize a synthetic NCN composed of biomimetic noncovalent interactions (H-bonds, $\pi-\pi$, and CH$-\pi$ interactions) of intermediate complexity ($\sim$3000 Da),[8] which combine the molecular features of high manipulability and high-resolution characterizability to systematically study the response of a NCN to mutation (Scheme 1). This study articulates molecular and energetic mechanisms by which synthetic NCNs can respond to a stress and the key notion that "frustrated" noncovalent interactions displaying negative cooperativity are required to provide the driving force for the response. We refer to this driving force as *latent free energy* and describe a mechanistic and energetic model for responsive

Scheme 1. Schematic Representation of the Long-Range Response in the NCN Arising from the Aib to Gly Mutation in the Catenane Model System
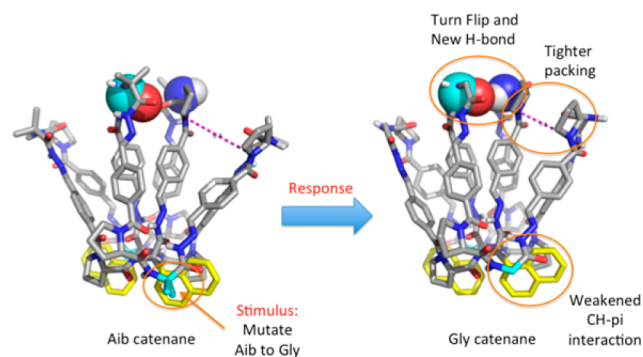
systems. These insights provide a new framework for considering the role of NCNs in complex function while also establishing a basis for designing new artificial stimuli-responsive systems.

RESULTS

System Design. Our model is a self-assembled C_2-symmetric [2]-catenane which we identified by dynamic combinatorial chemistry.[9] The two interlocked macrocyclic rings are each created via hydrazone linkages from four dipeptide monomers with exchangeable (with 50 equiv of CF$_3$COOH only) hydrazide and aldehyde termini (Scheme 2).[8] Once isolated, the catenanes are stable in the absence of TFA and amenable to full characterization.

Each macrocycle contains three D-Pro-X (**X** = Aib, Gly, or *t*-Leu) and one L-Pro-L-NaphGly subunit to form (L-Pro-L-

Scheme 2. Synthetic Scheme for the Self-Assembly of the [2]-Catenane Models with 50 Equiv of TFA
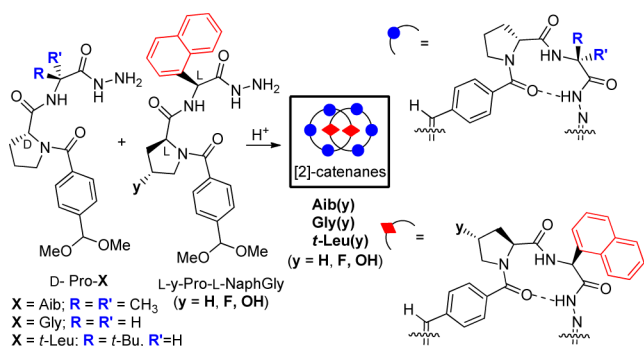
[a]After neutralizing, the catenanes are kinetically stable.

NaphGly)$_2$(D-Pro-**X**)$_6$ catenanes, with the arylGly monomer being required for catenane formation.[10] To facilitate X-ray analysis several L-Pro (**y** = H, OH, F) variants were also synthesized; our analysis shows the influence of **y** to be minimal (see SI Tables 7−9 and SI Figures 3a and 4a). The catenanes herein will be referred to by their **X** residue with the L-Pro substituent **y** noted in parentheses; the parent is thus **Aib(H)** (Scheme 2).

The X-ray structure of **Aib(H)** was previously reported (see ref 8c). Its NCN is assessed in more detail in this report as it provides a baseline for comparison to the mutants, **Gly(H)** and *t*-**Leu(H)**. Structural analysis of **Aib(H)**, which is formed in 63% yield,[8] revealed a highly organized compact structure with little free volume, consisting of 8 β-turns and multiple inter- and intramacrocyclic noncovalent interactions including H bonds, a CH−π-CH sandwich made up of L-Pro, naphthylGly, and a CH$_3$ group from Aib,[11,12] and an aryl barrel consisting of multiple edge−face and offset stacked π−π interactions (Figure 1a and 1b).

The necessity of the arylGly monomer in the self-assembly suggested that a conserved pair of CH−π-CH sandwiches was critical to catenane stability (Figure 1a and 1c), a notion that was tested by varying the components of the sandwich. Surprisingly, mutations designed to destabilize the CH component in **Aib(H)** were much less detrimental than expected. For example, deleting the methyl group (Aib to Gly mutation) and thereby removing two Me−π interactions worth approximately 1−1.9 kcal/mol each based on literature values (for a sum of 3−3.8 kcal/mol)[12,13] or increasing the alkyl size (Aib for *t*-Leu) and introducing a steric clash still yielded significant quantities of catenane (54% and 17% yield, respectively, cf. 66%).[14] While we expected these mutations to abolish catenane formation, they were accommodated in some way by the densely packed structure. To validate that yield corresponds to stability, we analyzed the speciation of these libraries and performed a mixed monomer experiment, described below. Our analysis of the equilibrium speciation of the self-assembly indicates that like the parent **Aib(H)** catenane, no unusually stable cyclic dimers, trimers, or tetramers are affecting the catenane equilibrium concentrations in the Gly(H) library. Moreover, a mixed-monomer experiment with both D-Pro-Aib and D-Pro-Gly monomers in combination with L-Pro-NaphGly shows that in addition to predominantly generating catenanes that are a mix of D-Pro-**X** units (by LC-MS), these mixed experiments generate both pure D-Pro-Gly- and D-Pro-Aib-derived catenanes in comparable quantities
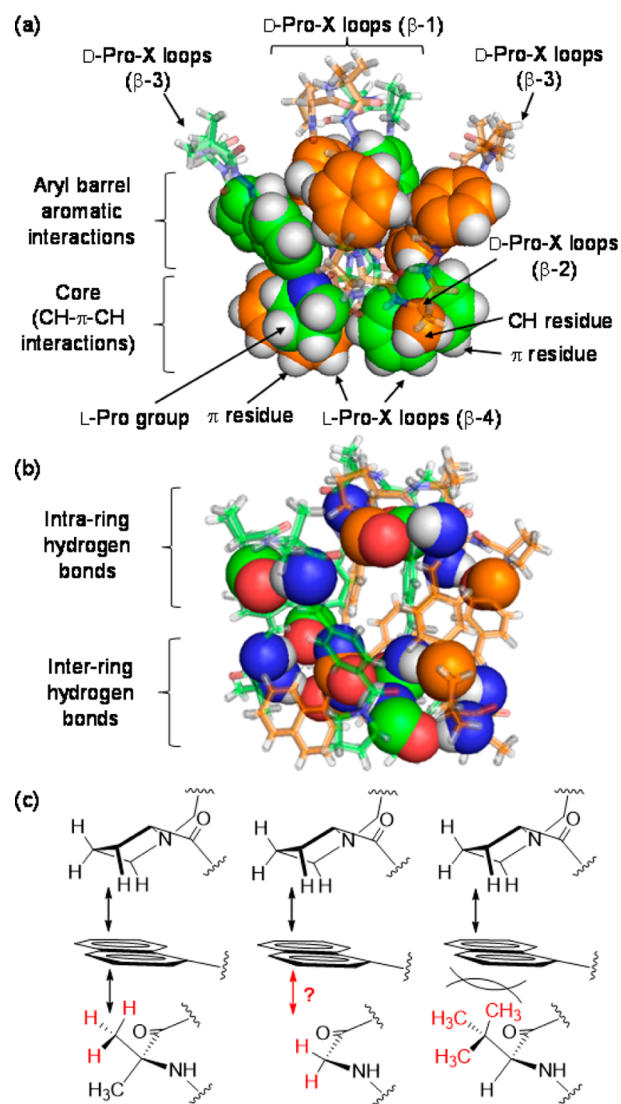

Figure 1. (a) Aryl barrel and the CH−π-CH networks (in space filling) in **Aib(H)**. (b) H-bond network (in space filling) in **Aib(H)**. (c) Schematic showing the Aib to Gly or *t*-Leu mutations. Green and orange represent the carbons in the two interlocking rings.

(within a factor of 2; see SI Figure 1). This experimental approach to validating the hypothesis that the D-Pro-Gly and the D-Pro-Aib catenanes are similar in energy complements numerical simulation approaches that we and others have investigated.[15] These studies indicate that it is inherent stability in the D-Pro-Gly catenane that is responsible for its significant concentration at equilibrium.

Taking yield as a measure of stability, in the **Gly(H)** case, the mutant is only ∼0.3 kcal/mol[8b] less stable than the parent **Aib(H)**.[16] On the basis of a theoretical maximum for the CH−π interaction (1.9 kcal/mol), the loss of two CH−π interactions in the absence of other changes to the structure is expected to destabilize the catenane up to 3.8 kcal/mol. Thus, the NCN provides up to 3.5 kcal/mol of compensating stabilization through its reoptimization. Even with a more modest value for a CH−π interaction of 1 kcal/mol, as measured in a torsion balance in CDCl$_3$,[13] the NCN still provides ∼1.7 kcal/mol in stabilization.

A similar analysis can be performed for *t*-**Leu(H)**. The A value for a *t*-Bu group is 4.7, versus 1.7 for CH$_3$. Thus, two *t*-

Leu groups should destabilize the catenane by 6 kcal/mol in the absence of reoptimization of the NCN. Instead, the observed loss in stability, based on yields, is only about 1.2 kcal/mol. Thus, rebalancing the NCN provides approximately 4.8 kcal/mol in stabilization.

Herein, we investigate the catenane structures and dynamics to gain molecular level insight into how the NCN provided this compensatory stabilization. As discussed below, we propose that this compensating energy can be thought of as latent free energy that is harnessed upon perturbation of the parent structure.

X-ray Structural Analysis. *Differences in β-Turns.* X-ray structures of **Gly(F)** and *t*-**Leu(OH)** were obtained to provide insight into how the NCN responds to the mutations relative to **Aib(H)** (Figure 2).[17] The most dramatic change observed is
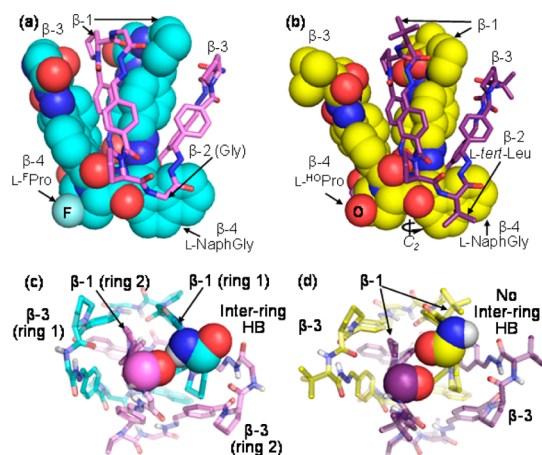

Figure 2. X-ray structures of (a) **Gly(F)** and (b) *t*-**Leu(OH)**; one ring in space filling. (c) New H bond formed upon flipping of the β-1 turns from II′ to I′ in **Gly(F)**. (d) Equivalent region of *t*-**Leu(OH)** (no turn flip, no H bond).

that, in contrast to **Aib(H)** and *t*-**Leu(OH)**, **Gly(F)** has its C_2-symmetry broken by the flipping of one β-1 turn from type II′ in **Aib(H)** to type I′. This turn flip creates an additional H bond between the β-1 and the β-1′ turns (Figure 2c).[18] Furthermore, the 1/1′, 2/2′, and 3/3′ turns in **Gly(F)** all exhibit deviations of more than 10° in one or more dihedral angle, indicating that methyl group loss creates changes across the entire structure. Changes in *t*-**Leu(OH)** are less pronounced with each turn remaining type II′ (Figure 2d) but with some deviations of >10° in turns 1/1′, 2/2′, and 3/3′. In all three catenanes, however, β-turn 4, which contains the L-Pro-arylGly unit in the core, is invariant, despite making up part of the altered CH−π-CH sandwich (see Table 1 and SI Table 2).

Analysis of CH−π-CH Interactions. The X-ray structures show intact CH−π-CH sandwiches, though they are maintained to different degrees in response to the variation in one CH component (Figure 3). The non-C_2-symmetric **Gly(F)** exhibits two slightly different weak Gly−naphthyl interactions. As shown in Figure 3a and 3b for one of these interactions, the Gly residue is displaced from both the center of the naphthyl ring and its plane, with Gly Cα distances to the naphthyl centroid lengthening to 4.85 and 4.93 Å (cf. 3.48 Å for the Aib methyl to centroid distance, see Figure 3c), at the extreme limit of what might be described as a CH−π interaction. The (expected) diminishment/loss of the Gly CH−π interaction is, however, counterbalanced by a tighter prolyl−π interaction (3.72 and 3.74 Å for the F-Pro groups versus 4.08 Å for **Aib(H)**).[19]

The *t*-**Leu(OH)** catenane also displays significant plasticity, as it accommodates the bulky *tert*-butyl group by shifting the entire *t*-Leu-containing β-turn away from the naphthyl ring to place a CH₃ in the same position as the Aib CH₃ (CH₃−π = 3.28 Å; Figure 3d). This shift pushes the proline away from the center of the naphthyl ring (angle between the proline centroid

Table 1. Direct Comparison of Metrical β-Turn Parameters of Aib(H), Gly(F), and *t*-Leu(OH) (dihedral angle differences > 10° in red bold italics)

		Dihedral Angle (°) of **Aib(H)**, (turn type)	Δ**Gly(F)** − **Aib(H)** (°)[a,b], (turn type)				Δ*t*-**Leu(OH)** − **Aib(H)** (°)[a,b], (turn type)	
			Ring 1		Ring 2			
β-1 (loop)	Φ_{i+1}	60.1	0.9		-2.0		-2.1	
	Ψ_{i+1}	-145.9	*176.9*	(I′)	*11.7*	(II′)	5.0	(II′)
	Φ_{i+2}	-52.3	*162.3*		*-23.7*		*-13.0*	
	Ψ_{i+2}	-35.9	9.9		*26.9*		*11.7*	
β-2 (core)	Φ_{i+1}	58.7	7.4		6.4		*12.2*	
	Ψ_{i+1}	-128.9	1.6	(II′)	6.4	(II′)	1.3	(II′)
	Φ_{i+2}	-62.7	*-13.3*		*-20.3*		-9.1	
	Ψ_{i+2}	-13.0	*15.0*		*16.0*		-1.2	
β-3 (loop)	Φ_{i+1}	58.5	*10.4*		3.5		*10.2*	
	Ψ_{i+1}	-126.6	-5.5	(II′)	*13.6*	(II′)	4.3	(II′)
	Φ_{i+2}	-62.5	*-20.5*		*-13.5*		-8.8	
	Ψ_{i+2}	-21.9	*27.9*		*14.9*		4.9	
β-4 (core)	Φ_{i+1}	-66.1	-5.4		-6.5		-2.5	
	Ψ_{i+1}	-30.7	-2.5	(VIII)	-1.6	(VIII)	3.2	(VIII)
	Φ_{i+2}	-159.0	-1.0		6.1		4.6	
	Ψ_{i+2}	83.0	-8.2		1.0		-4.1	

[a]The corresponding dihedral angle differences in **Aib(H)** vs **Gly(F)** and *t*-**Leu(OH)**. Two values are reported for **Gly(F)** as this compound is non-C_2 symmetric in the solid state. The ring containing a flipped β-1 turn (type I′) is assigned ring 1 (see Figure 2c). [b]The RMSD metrics of structural similarity to the parent structure (**Aib(H)**) obtained by PyMOL pair fitting using 48 arbitrarily chosen backbone atom pairs for **Gly(F)** vs **Aib(H)** and *t*-**Leu(OH)** vs **Aib(H)** are 0.811 and 0.642 Å, respectively.
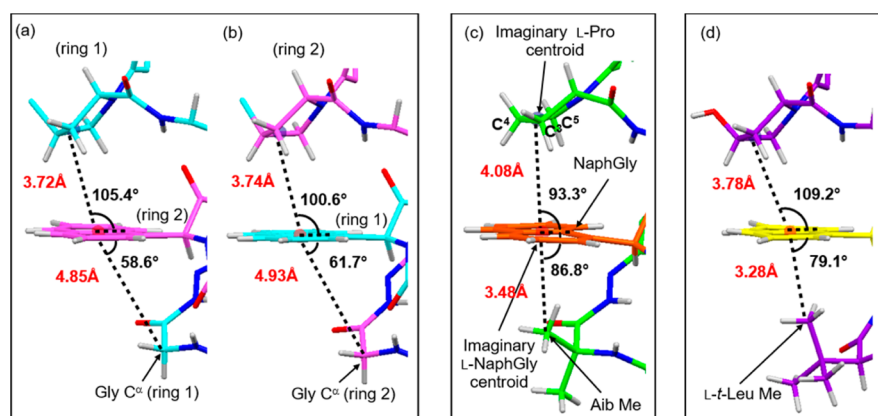
Figure 3. Direct comparison of the distances (red bold, Angstroms) and angles (bold, degrees) characterizing CH−π interactions (computed from the imaginary proline and L-NaphGly centroids and from the imaginary L-NaphGly centroid and Gly CH₂, Aib Me, and L-t-Leu Me, respectively) in (a and b) **Gly(F)** (The ring containing a flipped β-1 turn (type I′) is assigned ring 1), (c) **Aib(H)**, and (d) **t-Leu(OH)**; distance cutoff is X---M ≤ 5.0 Å (see Shanthi, V.; Ramamathan, K.; Sethumadhavan, R. *Appl. Biochem. Biotechnol.* **2010**, *160*, 1473−1483). The imaginary proline centroid was generated from three carbons of an L-Pro (C^3, C^4, C^5) and an imaginary L-NaphGly centroid computed from 10 carbons of the L-Naph ring.

Table 2. Direct Comparison of the Distances (Angstroms) between O and the N Atoms of the Main Chain C=O and NH groups forming Intraturn and Inter-Ring Hydrogen Bonds in Aib(H), Gly(F), and t-Leu(OH)

		N-O distance (Å) in Aib(H)	ΔGly(F) − Aib(H)[a]		Δt-Leu(OH) − Aib(H)[a]
			ring 1[b]	ring 2[b]	
Intra-turn hydrogen bonds of **Gly(F)**	β-1	3.14	-0.27	-0.29	-0.20
	β-2	2.96	-0.01	0.01	0.07
	β-3	2.91	-0.02	0.20	0.10
Inter-ring hydrogen bonds of **Gly(F)**	N (amide) in β-2 to O in β-4	2.99	-0.10	-0.15	-0.11
	N (hydrazone) in β-4 to O in β-2	2.95	-0.00	0.01	-0.02
	N (amide) in β-4 to O in β-4	2.86	0.03	-0.02	0.02
	N (amide) in β-1 to O in β-1′	none	2.99	none	none

[a]The differences in the corresponding H-bond distances forming intraturn and inter-ring hydrogen bonds in **Aib(H)** vs **Gly(F)** and **Aib(H)** vs **t-Leu(OH)**. The negative values represent the shorter distance compared to those in **Aib(H)**. [b]The ring containing the N atom of an NH group forming inter-ring hydrogen bonds. The ring contained a flipped β-1 turn (type I′) is assigned as ring 1 (see Figure 2c).

and the naphthyl ring =109.2° versus 93.3° for **Aib(H)**) but 0.3 Å closer than for **Aib(H)** (Figure 3d and SI Table 3). Thus, in both **Gly(F)** and **t-Leu(OH)** the NCN has flexed to optimize the critical CH−π-CH sandwich.

Analysis of the H-Bond Network. Compensatory changes also appear as strengthened H-bond contributors to the NCN. **Aib(H)** contains six pairs of symmetry-related H bonds with N−O distances of 2.8−3.1 Å. As mentioned above, the turn flip at β-1 in **Gly(F)** creates a new H bond with a N−O distance of 3 Å (Figure 2c and Table 2). Also shortened are the β-1 intraring H bonds (by 0.3 Å) and the inter-ring H bonds in the core between β-2 and β-4 (by 0.1 Å).[20] Except for the new β-1/

β-1′ H bond, these same H-bonds are also strengthened in **t-Leu(OH)** (see Table 2), implying that the NCN can reoptimize specific local and distant noncovalent interactions to compensate for the mutations.

Changes in the Aryl Barrel. Changes to the CH−π-CH sandwich and H bonds are also accompanied by constriction in the aryl barrel in both the Gly and the t-Leu catenanes relative to the Aib catenane (see SI Tables 5 and 6 for metrical parameters), suggesting plasticity in the aryl barrel that accommodates changes elsewhere in the structure. *Together this structural analysis indicates that both mutated catenanes*
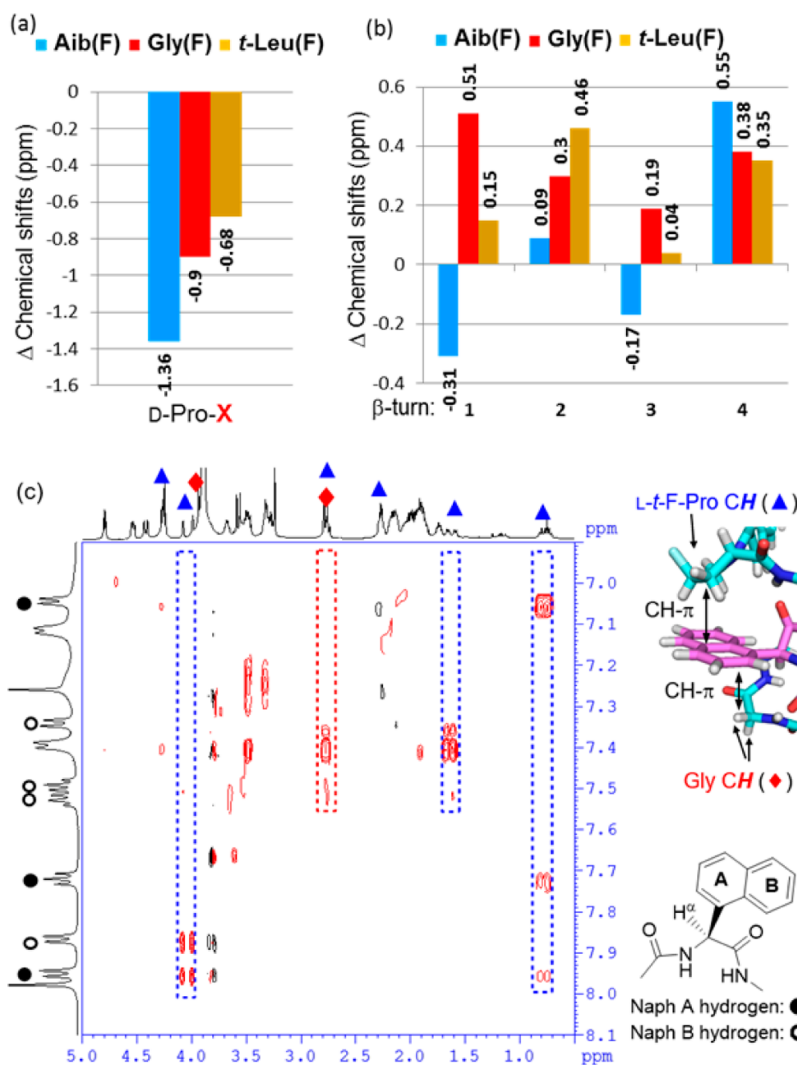
Figure 4. Direct chemical shift comparisons of (a) catenane D-Pro-X residue X signal (β-2 turn) and (b) catenane amide NHs. Series of catenanes are **Aib(F)**, **Gly(F)**, and ***t*-Leu(F)**. (c) Portion of the 2D 600.1 MHz ROESY spectrum of **Gly(F)** (F1 δ = 6.9−8.1 ppm; F2 δ = 0.0−5.0 ppm; MeOH-d_4/5CDCl$_3$, 25 °C) showing CH−π interactions between L-*trans*-F-Pro or a Gly CH$_2$ group and the NaphGly residue, respectively.

accommodate the perturbation of the CH−π interaction via reoptimization of the entire NCN rather than just local changes.

NMR Structural Analysis of Gly- and *t*-Leu Catenanes. NMR analysis of the *trans*-F-Pro version of all three catenanes was also undertaken to provide additional structural and dynamic information.[21] At RT, the catenanes all display C_2-symmetric ^{1}H NMR spectra, despite the asymmetric X-ray structure of **Gly(F)** (but see VT NMR of **Gly(F)** vide infra).

Analysis of CH−π-CH Interactions. Inspection of several key chemical shifts relative to their monomers gives molecular level insights into the response of the NCN to individual perturbations of the CH−π interaction. The chemical shifts of the CH−π participating methyl and CH resonances in **Aib(F)** and **Gly(F)**, respectively, show a weaker (less populated) CH−π interaction for **Gly(F)** than for **Aib(F)** (normalizing for the number of hydrogens, 0.9 ppm and 1.36 × 3 = 4.08 ppm upfield shifted for **Gly(F)** and **Aib(F)** relative to monomer, respectively, see Figure 4a). Despite the smaller shift for **Gly(F)**, the upfield shifting still suggests that the CH is near the face of the aromatic group, in agreement with the X-ray structure (see Figure 3a and 3b). The CH$_3$ groups of the *t*-Bu residue in ***t*-Leu(F)** are upfield shifted by 0.68 ppm. This is an

average shift for the three equilibrating methyl groups, thus providing a net upfield shifting of 0.68 × 9 = 6.12 ppm when normalized for the number of hydrogens to compare to **Gly(F)** above. While this may be a slight overestimate, this large shift is in agreement with the close contact observed in the X-ray structure (see Figure 3c). Thus, the apparent order of CH−π interactions between the mutated residue and the naphthyl group is ***t*-Leu(F)** > **Aib(F)** > **Gly(F)**. The variation in the proline chemical shifts that participates in the Pro−π interaction is much smaller (see SI Figure 3), suggesting that the interaction is of similar magnitude across all three catenanes.

Analysis of the H-Bond Network. Comparison of the amide NH chemical shifts also revealed changes in the H-bonding network (see Figure 4b). The β-1 amide, which is in the central "stem" of the catenane (Table 2), is 0.8 ppm farther downfield in **Gly(F)** relative to **Aib(F)**, consistent with the former's new inter-ring H bond (X-ray, Figure 2c). The β-2 amide NH in the core also becomes increasingly downfield shifted on going from Aib to Gly to *t*-Leu, suggesting that this inter-ring H bond is stronger (more populated) in the mutants. Conversely, the β-4 amide NH (which is buried in the core of the catenane and

forms an inter-ring H-bond) is 0.2 ppm less downfield shifted in **Gly(F)** and **t-Leu(F)**, suggesting a slightly weaker (less populated) H bond in the core relative to **Aib(F)**. *Taken together these H-bond shifts suggest that in addition to the already noted changes in the CH−π interactions, Gly(F) and t-Leu(F) also compensate for their mutations by reoptimizing their entire H-bond networks.*

Changes in the Aryl Barrel. Analysis of the aryl rings in the aryl barrel indicate that only minor changes occur on mutation of Aib to Gly or *t*-Leu (see SI Table 8). In the *t*-Leu case, there is no significant change in the chemical shifts of the aryl protons (<0.1 ppm in all positions). In the **Gly(F)** catenane, most protons exhibit upfield shifting of ∼0.5 ppm relative to **Aib(H)**, and one proton is shifted upfield by ∼0.2 ppm, suggesting a modestly more favorable edge−face interaction. *In sum, this demonstrates that while the aryl barrel is able to transmit a signal (mutation) from one end of the catenane to the other, it is quite malleable and does not appear to directly contribute significant stability to the mutants through new or optimized noncovalent interactions.*

Dynamics by H/D Exchange. H/D exchange was used to extract dynamics information from the various catenanes, as has been done in proteins such as streptavidin.[22] While noncyclized monomers undergo complete H/D exchange in <5 min, amide exchange times in these closed structures is much slower and structurally variable. For example, the exchange half-lives for **Aib(F)** vary from minutes in the periphery to days in the core β-4 amide positions (Figure 5). By contrast, the amide
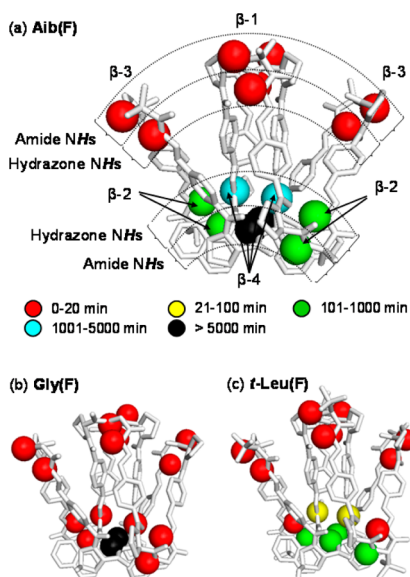

Figure 5. Half-life for position-dependent H/D exchange in (a) **Aib(F)**, (b) **Gly(F)**, and (c) **t-Leu(F)** determined by NMR spectroscopy. Reaction conditions: MeOH-d_4/5CDCl$_3$, 20 °C.

exchange behavior of **Gly(F)** is characterized by slow exchange at β-4 and fast exchange at all other positions (<20 min). Interestingly, the exchange rates for **t-Leu(F)** exhibit a different pattern in which exchange is faster than **Aib(F)** and **Gly(F)** at the most buried β-4 amide (factor of 20) but similar to **Aib(F)** at the slightly more exposed β-2 amide positions (factor of 3) (Figure 5c and SI Table 10).

VT NMR of Gly(F). VT NMR indicates that **Gly(F)** displays more significant dynamics than the other catenanes (see Figure 6 and SI Figures 17 and 18). While its RT ^{1}H NMR displays
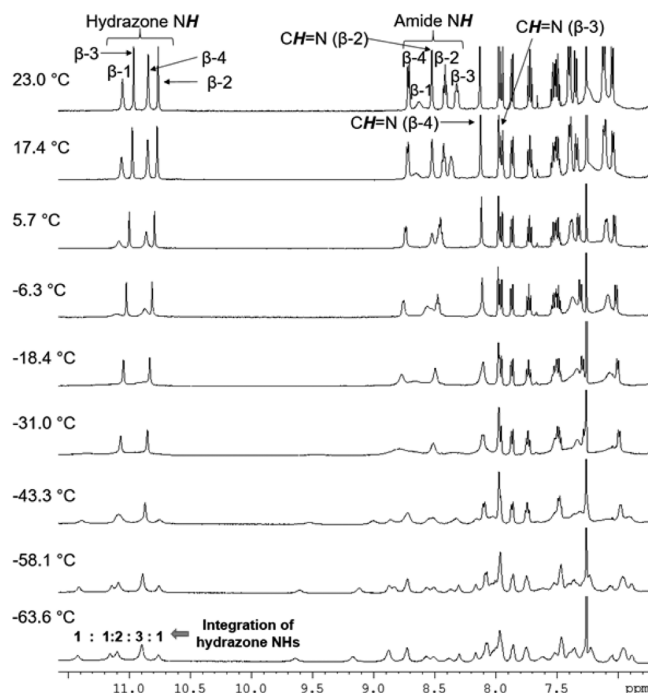

Figure 6. Portion of the ^{1}H variable-temperature (VT) NMR spectra of **Gly(F)** (δ = 7.4−11.6 ppm, 500.1 MHz, MeOH-d_4/5CDCl$_3$) showing the decoalescence of hydrazone and amide NHs.

net C_2 symmetry, cooling desymmetrizes the spectra into two sets of resonances (<−26 °C) that are consistent with the asymmetry observed in the X-ray structure. This process does not occur in **Aib(F)** or **t-Leu(F)** even as low as −80 °C. These millisecond dynamics combined with enhanced rates of H/D exchange show **Gly(F)** to be far more dynamic than **Aib(F)**, while **t-Leu(F)** is intermediate. *As discussed below, one intriguing implication is that part of the unexpected stability in Gly(F) is due to enhanced dynamics.*

■ DISCUSSION

These studies describe the molecular level response of a NCN to a mutation-induced stress. In the mutation of Aib to Gly, the stress is a loss of a key CH$_3$−π interaction, whereas in the Aib to t-Leu mutation, a steric clash is introduced. In the absence of any changes in the catenane structure, these mutations would be expected to abolish the thermodynamically controlled self-assembly. However, due to long-range restructuring of the NCN in combination with changes in dynamics, these mutants still provide stable, well-folded catenanes. Our detailed structural (X-ray and NMR) and dynamics studies provide sufficient resolution to describe this phenomenon in terms of specific molecular features.

NMR and X-ray data indicate that the CH−π interaction is in fact weakened or lost in the **Gly(X)** catenane, while the Pro−π interaction is maintained. These structural studies show a number of short- and long-range changes to the baseline interactions that together enthalpically compensate for this loss. Most prominent is the long-distance turn flip that introduces a new H bond between the β-1 and the β-1′ rings. Other H-bonds in turns 2 and 3 are more populated as well, based on NH chemical shifts. Structural changes are also observed in the aryl barrel, although with minimal changes to chemical shifts. Interestingly, this system appears to employ a seesaw type of mechanism, where a large perturbation at one end of the NCN

(CH−π interaction) induces a large change at a distal site (turn flip and new H bond), suggesting that the plasticity of the aryl barrel may facilitate communication from one end of the catenane to the other. *Thus, while the responses to the mutations can be assessed on an individual residue level, they are best viewed as a network of perturbations that propagate through the NCN of hydrogen bonds and aromatic noncovalent interactions.*

Entropy also contributes to the stabilization of **Gly(X)**, including a millisecond time scale conformational switching term ($R\cdot$ln 2 = 0.4 kcal/mol)[23] as well as longer time scale (min +) structure-dependent (**Gly(F)** vs **Aib(F)** vs **t-Leu(F)**) and position-dependent H/D exchange processes. Importantly, entropy and enthalpy are not always opposed in these systems, as some amides became more dynamic while simultaneously exhibiting stronger H-bonds based on chemical shifts.

The *t*-Leu catenane also displays both enthalpic and entropic stabilization throughout the NCN. The CH−π-CH sandwich shifts to accommodate the t-Leu, resulting in a CH−π interaction that is stronger relative to Aib, based on CH$_3$−π distances and CH$_3$ chemical shifts. As with **Gly(F)**, **t-Leu(F)** also exhibits several more populated H bonds based on amide NH chemical shifts, while its dynamics, based on H/D exchange, are intermediate to those of **Gly(F)** and **Aib(F)**. In this case, however, no turn flip was induced, and so no additional H-bond was gained. *Thus, each mutant utilizes a unique rebalancing of the NCN to maintain stability.*

These two mutants exhibit a number of protein-like qualities. The Aib-to-*t*-Leu mutant is reminiscent of a nonlethal protein mutation which is accommodated by many small changes throughout the folded structure. The Aib-to-Gly mutant, which induces the aforementioned long-distance turn flip and the creation of a new H bond, is a model for actuation through a network. Signal transduction through a membrane relies on the propagation of a signal (structural perturbation) at one end of the protein or protein complex into a commensurate structural change over the long distances characteristic of membrane spanning proteins. Furthermore, the opposite mutation from Gly to Aib (i.e., introducing a methyl group) exhibits characteristics of binding by conformational selection.[24] The Gly catenane exhibits significant conformational flexibility, but introduction of the methyl group in the Aib catenane causes one conformation to be stabilized over the others. In this case, the conformational selection is the result of mutation rather than binding.

While the stimulus-response mechanisms observed in these catenanes are comparable to those observed in proteins, it is significantly different than typical designed stimuli-responsive systems, such as molecular shuttles or designed allosteric systems.[25] Most designed stimulus-responsive systems utilize simple "on−off" interactions rather than networked interactions as observed here (and utilized by proteins).

Analysis of the changes in the NCN in these systems leads to a novel energetic model regarding how NCNs are able to respond to a stimulus as described below. First, consider a case in which every individual noncovalent interaction in a NCN is fully optimized. Introduction of a stimulus (in this case a mutation that results in loss of a favorable interaction) would necessarily destabilize the network by the same amount. In this scenario, the network would be unable to respond. However, if the NCN is in its lowest energy accessible state but *not every individual* noncovalent interaction can access its optimal geometry, perturbation at one position in the network *could be compensated for by improvements in different positions*, i.e. loss

of one interaction allows for formation/optimization of others. This reoptimization of the network can result in long-distance structural and dynamic changes, as are required for complex protein functions such as allostery and signal transduction. Stated another way, *some interactions in the initial NCN must exhibit negative cooperativity for the NCN to be responsive to a stimulus.* In a broad sense, such a *responsive NCN* represents an example of systems chemistry and is conceptually equivalent to dynamic combinatorial chemistry, in which a system finds a new lowest energy state in response to a stimulus.[26]

In this description, a *responsive NCN* is therefore one that stores potential energy—the potential energy of the non-optimal noncovalent interactions. We refer to this stored energy as latent free energy (LFE). In the Aib-to-Gly mutation, for example, latent free energy is released in the form of new and strengthened H-bonds and increased entropy, the combination of which compensate for the loss of the CH−π interactions. While this concept of latent free energy is related to cooperativity, the concepts describe different phenomena. Cooperativity defines the energetic relationship between individual noncovalent interactions within a network relative to an unnetworked reference state, whereas latent free energy is associated with the ability of a cooperative network of interactions to re-equilibrate in response to a stimulus. The concept of a responsive NCN being one that stores latent free energy may be related to the observation that functional sites in proteins (both binding and catalytic sites and allosteric sites) contain a higher degree of "frustration", as described by Wolynes.[27] Frustration at such sites may correlate with the storage/buildup of latent free energy and hence responsiveness of the network. We suggest that this conceptual framework connecting molecular level changes to an energetic model of latent free energy will be valuable for developing new synthetic systems with more complex function.

■ CONCLUSIONS

Our model demonstrates that a stimulus-triggered rebalancing of attractive and repulsive forces in a NCN can drive complex responses and thereby link molecular observables to an energetic framework. The outcome of our studies on a responsive NCN provides a strategy for mimicking similar, albeit more complex, protein functions involving actuation. If general, a corollary of our LFE model of NCN properties is that de novo designed proteins may be overdesigned for stability rather than for the latent free energy required to drive function.[28] Moreover, this work provides a conceptual basis for incorporating complex NCN-coupled functions in novel abiotic systems.

■ EXPERIMENTAL SECTION

General Methods. Chemicals were purchased from Aldrich, Fisher Scientific, and Chem-Impex International, Inc. and used as received. ^{1}H and ^{13}C NMR spectra of monomers and ^{1}H, ^{13}C, gradient TOCSY, ROESY, and HSQC NMR spectra of [2]-catenanes were recorded on a Bruker DRX 500 spectrometer or a Bruker 600 Cryoprobe spectrometer and processed using TOPSPIN Bruker NMR software (V. 3.1) or MNOVA software (V.9.0). High-resolution mass spectra (HRMS) for [2]-catenanes were obtained on an Agilent Accurate LC-TOF Mass Spectrometer (Agilent Series 6220) operating in positive ion mode with an electrospray ionization source (fragmentor = 375 V). The data were analyzed using Agilent MassHunter Workstation Software, Qualitative Analysis (V. B.02.00). HPLC analysis was performed on a Hewlett-Packard Series 1100 instrument using a Halo-C18 column (4.6 × 150 mm, 2.7 μm) with gradient elution

(methanol/water) at a flow rate of 0.45 mL/min and at 55 °C. The injection volume for a 5 mM dynamic self-assembly (DSA) reaction was typically 3.5 μL. UV absorbance chromatograms were recorded at wavelengths of 220 and 289 nm, and the data were analyzed using Agilent Chemstation software. DSA reaction analyses were also performed on an Accurate Mass LC-TOF instrument equipped with an Agilent Series 1200 LC instrument using a Halo-C18 column (2.1 × 50 mm, 2.7 μm) with gradient elution (methanol/water containing 0.1% formic acid) at a flow rate of 0.5 mL/min and at 50 °C. The eluent was analyzed in positive ion mode with an electrospray ionization source (fragmentor = 375 V). The data were analyzed using Agilent MassHunter Workstation Software, Qualitative Analysis (V. B.02.00). The isolation of [2]-catenanes was performed on a modified semipreparative HPLC (Agilent Series 1200 LC instrument) using an Agilent Zorbax Eclipse XDB-C18 PrepHT Cartridge column (21.2 × 250 mm, 7 μm) with isocratic methanol/water elution (no additive) at a flow rate of 9 mL/min and at 55 °C. The eluent was monitored by UV absorbance chromatogram at a wavelength of 289 nm, and the corresponding [2]-catenanes were collected with a fraction collector. X-ray crystallographic analyses of the [2]-catenanes **Gly(F)** and *t*-**Leu(OH)** were performed on a Bruker-AXS SMART APEX-II CCD′ system equipped with a graphite monochromator: please contact Dr. Peter S. White (pwhite@email.unc.edu) for correspondence regarding X-ray analyses.

Synthesis of Catenanes and Generation of Gly(F) and *t*-Leu(OH) Crystals for X-ray Analysis. D-Pro-Gly, D-Pro-L-*t*-Leu, and L-**y**-Pro-L-NaphGly were prepared using previously reported methods. Several 5 mM dynamic self-assembly solutions (D-Pro-Gly or D-Pro-L-*t*-Leu/L-**y**-Pro-L-NaphGly = 3:1) were prepared on a 20 mL scale. Each was prepared by dissolving the mixture of D-Pro-Gly or D-Pro-L-*t*-Leu (75.0 μmol) and L-**y**-Pro-L-NaphGly (25.0 μmol) in 25% acetonitrile/chloroform (20 mL) and subsequently adding trifluoroacetic acid (50 equiv, 5000 μmol). The solutions were allowed to sit for several days until a steady state was reached (typically 7−10 days) prior to LC analyses. Once steady state had been reached, triethylamine (50 equiv, 5000 μmol) was added to neutralize the acid. The solutions were combined, and the volatiles were removed in vacuo. The minimum amount of 60% methanol/chloroform solution was added to the resulting mixture and then filtered. This solution was eluted on the semipreparative HPLC with various methanol/water isocratic solutions to isolate the catenanes. Recrystallization of *t*-**Leu(OH)** from methanol/chloroform by slow vaporization at RT for several weeks offered crystals suitable for X-ray analysis. For the catenane **Gly(F)**, the crude mixture obtained after neutralization by trimethylamine was directly used without further purification. The recrystallization of this crude mixture from methanol/chloroform at low temperature (<−20 °C) over a year offered crystals suitable for X-ray analysis.

H/D Exchange Experiments of Catenane Amide and Hydrazone NHs. The isolated catenanes purified by semipreparative HPLC were recrystallized from methanol/chloroform/pentanes several times and then dried at 70 °C in vacuo prior to use for NMR analysis. For the H/D exchange experiments, the catenane (1 μmol) was mixed with the premade MeOH-d_4/5CDCl$_3$ solution (0. 55 mL) and spectra immediately recorded at different times from the mixing. For turn assignments of the amide and hydrazone NHs, 2D NMR (TOCSY and HSQC) and 1D selective gradient ROESY NMR experiments were conducted by adding the catenane to premade CD$_3$OH/5CDCl$_3$ solution (see SI Figures 9−12). To compute the half-life ($t_{1/2}$) values of the catenane amide and hydrazone NHs, their integrated peak areas were calibrated to an internal imine proton signal integration area (CH═N; set as 1) in the ^{1}H NMR spectra and then plotted versus the time from mixing (see SI Figure 15).

ASSOCIATED CONTENT

⑤ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b07833.

Synthetic and H/D exchange procedures; NMR characterization data; LC-TOF data for mixed monomer thermodynamic determination; VT NMR data, X-ray numerical analysis data (PDF)
pdb file for **Aib(H)** (PDB)
pdb file for **Gly(F)** (PDB)
pdb file for *t*-**Leu(OH)** (PDB)

AUTHOR INFORMATION

Corresponding Authors
*mgagne@unc.edu.
*mlwaters@unc.edu.

Notes
The authors declare no competing financial interest.

REFERENCES

(1) (a) Williams, D. H.; Zhou, M.; Stephens, E. *J. Mol. Biol.* **2006**, 355, 760−767. (b) del Sol, A.; Tsai, C. J.; Ma, B. Y.; Nussinov, R. *Structure* **2009**, 17, 1042−1050.
(2) Grigoryan, G.; Moore, D. T.; DeGrado, W. F. *Annu. Rev. Biochem.* **2011**, 80, 211−237.
(3) Csermely, P.; Palotai, R.; Nussinov, R. *Trends Biochem. Sci.* **2010**, 35, 539−546.
(4) Williams, D. H.; Stephens, E.; O'Brien, D. P.; Zhou, M. *Angew. Chem., Int. Ed.* **2004**, 43, 6596−6616.
(5) (a) Wand, A. J. *Curr. Opin. Struct. Biol.* **2013**, 23, 75−81. (b) Whitley, M. J.; Lee, A. L. *Curr. Protein Pept. Sci.* **2009**, 10, 116−127. (c) Tzeng, S. R.; Kalodimos, C. G. *Curr. Opin. Struct. Biol.* **2011**, 21, 62−67.
(6) Creighton, T. E. *Proteins: Structures and Molecular Properties*, 2nd ed.; W. H. Freeman, 1992.
(7) (a) Otto, S. *Dalton Trans.* **2006**, 2861−2864. (b) Rodriguez-Docampo, Z.; Pascu, S. I.; Kubik, S.; Otto, S. *J. Am. Chem. Soc.* **2006**, 128, 11206−11210. (c) Rodriguez-Docampo, Z.; Eugenieva-Ilieva, E.; Reyheller, C.; Belenguer, A. M.; Kubik, S.; Otto, S. *Chem. Commun.* **2011**, 47, 9798−9800. (d) Gunasekara, R. W.; Zhao, Y. *J. Am. Chem. Soc.* **2015**, 137, 843−849. (e) Carrillo, R.; Feher-Voelger, A.; Martin, T. *Angew. Chem., Int. Ed.* **2011**, 50, 10616−10620. (f) Zhao, C.; Li, P.; Smith, M. D.; Pellechia, P. J.; Shimizu, K. D. *Org. Lett.* **2014**, 16, 3520−3523.
(8) (a) Chung, M.-K.; White, P. S.; Lee, S. J.; Waters, M. L.; Gagne, M. R. *J. Am. Chem. Soc.* **2012**, 134, 11415−11429. (b) Chung, M.-K.; Lee, S. J.; Waters, M. L.; Gagne, M. R. *J. Am. Chem. Soc.* **2012**, 134, 11430−11443. (c) Chung, M.-K.; White, P. S.; Lee, S. J.; Gagne, M. R. *Angew. Chem., Int. Ed.* **2009**, 48, 8683−8686.
(9) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S. *Chem. Rev.* **2006**, 106, 3652−3711.
(10) In addition to the 2:6 catenane, some 3:5 stoichiometry catenane is formed. The total catenane yields for the **Aib(H)**, **Gly(F)**, and *t*-**Leu(OH)** syntheses are 66%, 53%, and 20%, respectively.
(11) Zondlo, N. J. *Acc. Chem. Res.* **2013**, 46, 1039−1049.
(12) Nishio, M. *Phys. Chem. Chem. Phys.* **2011**, 13, 13873−13900.
(13) Carroll, W. R.; Zhao, C.; Smith, M. D.; Pellechia, P. J.; Shimizu, K. D. *Org. Lett.* **2011**, 13, 4320−4323.
(14) Although not discussed herein, numerous (but not all) mutations completely inhibit catenane formation (see ref 8b).
(15) (a) Severin, K. *Chem. - Eur. J.* **2004**, 10, 2565−2580. (b) Chung, M.-K.; Hebling, C. M.; Jorgenson, J. W.; Severin, K.; Lee, S. J.; Gagne, M. R. *J. Am. Chem. Soc.* **2008**, 130, 11819−11827. (c) Otto, S. *Acc. Chem. Res.* **2012**, 45, 2200−2210 and references therein. (d) Orrillo, A.

G.; Furlan, R. L. E. *J. Org. Chem.* **2010**, *75*, 211−214. (e) Misuraca, M. C.; Moulin, E.; Ruff, Y.; Giuseppone, N. *New J. Chem.* **2014**, *38*, 3336−3349 and references therein.

(16) Three monomer self-assembly experiments were used to directly complete the two catenanes (and numerous mixed structures) and verify the similar stability of **Aib(H)** and **Gly(F)** and the slightly lowered stability of *t*-**Leu(OH)** vs **Aib(H)** (see SI Figures 1 and 2).

(17) By NMR these catenane variants are symmetric in solution (see SI Tables 7−9 and SI Figures 3 and 4).

(18) A similar turn flip was observed in a catenane with D-Pro-cyclobutylglycine in which the core CH−π-CH interaction was weakened (see refs 8a and b). To the best of our knowledge, a type I′ turn has not previously been observed in a D-Pro-Gly containing natural peptide, suggesting that it is the context of the catenane not the inherent turn preference that induced the turn flip. (a) Gopi, H. N.; Roy, R. S.; Raghothama, S. R.; Karle, I. L.; Balaram, P. *Helv. Chim. Acta* **2002**, *85*, 3313−3330. (b) Wilmot, C. M.; Thornton, J. M. *J. Mol. Biol.* **1988**, *203*, 221−232.

(19) F-Pro, HO-Pro, and H-Pro each provide similar quantities of the Gly-catenanes at equilibrium, and their NMR behavior indicates that they are equally well packed on the arene face.

(20) These two H bonds are similarly shortened in *t*-**Leu(OH)** (see Table 2 and SI Table 4).

(21) The HO-Pro analogs were also analyzed by ^{1}H NMR spectroscopy and found to behave similarly (see SI Tables 7−9 and SI Figures 3 and 4).

(22) Williams, D. H.; Stephens, E.; Zhou, M. *J. Mol. Biol.* **2003**, *329*, 389−399.

(23) Csermely, P.; Palotai, R.; Nussinov, R. *Trends Biochem. Sci.* **2010**, *35*, 539−546.

(24) Eliel, E. L.; Wilen, S. H. *Stereochemistry of Organic Compounds*; John Wiley and Sons: New York, 1994; pp 597−602.

(25) Erbas-Cakmak, S.; Leigh, D. A.; McTernan, C. T.; Nussbaumer, A. L. *Chem. Rev.* **2015**, *115*, 10081−10206.

(26) (a) Li, J.; Nowak, P.; Otto, S. *J. Am. Chem. Soc.* **2013**, *135*, 9222−9239. (b) Mattia, E.; Otto, S. *Nat. Nanotechnol.* **2015**, *10*, 111−119 and references therein..

(27) (a) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. *Proteins: Struct., Funct., Genet.* **1995**, *21*, 167−195. (b) Ferreiro, D. U.; Hegler, J. A.; Komives, E. A.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 19819−19824. (c) Ferreiro, D. U.; Hegler, J. A.; Komives, E. A.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 3499−3503. (d) Jenik, M.; Parra, R. G.; Radusky, L. G.; Turjanski, A.; Wolynes, P. G.; Ferreiro, D. U. *Nucleic Acids Res.* **2012**, *40*, W348−W351.

(28) (a) Shoichet, B. K.; Baase, W. A.; Kuroki, R.; Matthews, B. W. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 452. (b) Korendovych, I. V.; DeGrado, W. F. *Curr. Opin. Struct. Biol.* **2014**, *27*, 113−121.