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

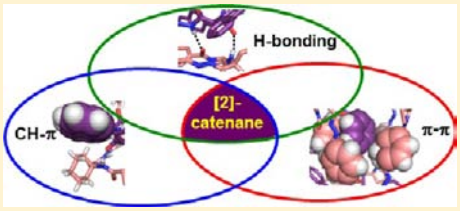
Self-Assembled Multi-Component Catenanes: Structural Insights into an Adaptable Class of Molecular Receptors and [2]-Catenanes

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

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

Supporting Information

ABSTRACT: Under acidic conditions (50 equiv of TFA), combinations of hydrazide A-B monomers self-assemble into octameric [2]-catenanes with high selectivity for $[\mathbf{1}_3\mathbf{2}]_2$, where **1** is a D-Pro-X (X = Aib, Ac_4c, Ac_6c, L-4-Cl-PhGly)-derived monomer and **2** is an L-Pro′-L-arylGly (Pro′ = Pro, *trans*-F-Pro, *trans*-HO-Pro, aryl = naphthyl, phenyl)-derived monomer. Five different combinations of monomers were studied by X-ray crystallography. In each case, the unique aryl glycine unit is located in the core of the structure where the aryl ring templates a CH−π−CH sandwich. Analysis of metrical parameters indicates that this core region is highly conserved, while the more peripheral zones are flexible. ^{1}H NMR spectroscopy indicate that the solid-state structures are largely retained in solution, though several non-C_2-symmetric compounds have a net C_2-symmetry that indicates accessible dynamic processes. Catenane dynamic processes were additionally probed through H/D exchange, with the core being inflexible relative to the peripheral structure. Mass spectrometry was utilized to identify the constitutional isomerism in the minor asymmetric $[\mathbf{1}_5\mathbf{2}_3]$ catenanes.



INTRODUCTION

The self-assembly of complex structures from simple building blocks is a general principle that can lead to significant gains in structure and function.[1] As exemplified by the protein folding problem,[2] global minima are achieved through the interplay of many weak interactions, the quantitation of which is often challenging. Within this context, it is the balance of non-covalent forces (H-bonding, electrostatic, π−π, cation−π, CH−π, and hydrophobic effects) that dictate thermodynamic minima. Model systems have been widely utilized to investigate these contributing non-covalent interactions in relative isolation, with the ease of manipulation and characterization enabling significant insight into the nature and driving force of such interactions.[3] A key limitation, however, is that few such model systems simulate the cooperative folding and binding found in proteins.[4]

We report herein structural studies of a synthetic self-assembling system that bridges the complexity/manipulability gap that exists between small-molecule model systems and complex biosystem such as proteins. The self-assembled structure is generated from dipeptide hydrazide monomers into discrete octameric [2]-catenanes of molecular weight >3 kDa. A structure analysis reveals that the compounds rely on H-bonding, π−π, and CH−π interactions along with biogenic β-turns to achieve their minimum-energy structures. The combination of large size, complexity, and structural manipulability provides the opportunity to gauge how the balance of forces influences the energetics of this self-assembled structure, and thus bridges the gap between small molecules and proteins.

Templated synthesis is a long-established strategy for accessing entropically disfavored compounds and can provide more complex recognition systems than designed host−guest pairs. A dynamic variant has been recently developed by Sanders, Lehn, and others as a means of discovering tight-binding hosts for a broad variety of guests.[5] Termed dynamic combinatorial or constitutional chemistry (DCC), this strategy enables molecular receptors to be selected from dynamic libraries of potential hosts without prior knowledge of the guest's structure. By its very nature, however, the competitive selection masks the recognition elements that create the equilibrium host−guest structure. While much is known about the constitution and binding properties of these host and host−guest assemblies, little is known about their molecular structures.

In the course of our own experiments examining complex libraries assembled from mixtures of dipeptide monomers,[6] we noted that one specific combination of two monomers led to the self-assembly of a [2]-catenane composed of 56-membered macrocyclic tetramers, each of which was composed of a 3:1 mixture of the two monomers (Figure 1).[7] While this [2]-catenane recognizes itself, Sanders has previously demonstrated that catenated dipeptide trimers will bind quat-salts like acetylcholine with sub-micromolar affinities, the host−guest complex of which could be characterized by NMR.[8] The lack of solid-state structural data on these and the other receptors assembled from peptide monomers[8−10] prompted us to utilize X-ray methods in conjunction with solution-phase studies to structurally characterize a number of additional [2]-catenanes.

Received: March 9, 2012
Published: June 11, 2012

   dx.doi.org/10.1021/ja302345n | J. Am. Chem. Soc. 2012, 134, 11415−11429

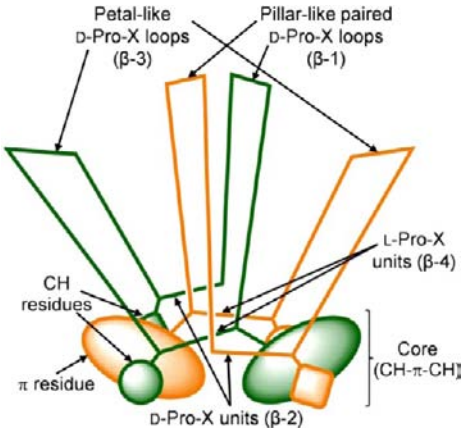


Figure 1. Cartoon showing the flower-like structure of the [2]-catenane composed of two interlocked identical tetramers.

Our communicated X-ray structure of [2]-catenane ^{H}**3a** revealed that the presence of four β-turns in each tetramer results in collapse of the large macrocycle into a conformer that converges functionality and enables self-recognition (Figure 1).[7] Inter- and intra-macrocycle hydrogen bonds, $\pi-\pi$ stacking, and aliphatic CH$-\pi$ interactions all appear to contribute to the high-fidelity recognition of one macrocycle by the other. In addition to being physically interlocked, two β-turns were found to create a clip that binds the aromatic portion of the C_2-related macrocycle, using the C−H bonds of one proline ring and the C−H bonds of one Aib methyl substituent (Aib = 1-aminoisobutyric acid, or α-methyl-alanine) as the teeth of the clip, i.e., a CH$-\pi-$CH sandwich. Due to the C_2-nature of the catenane, this binding interaction is repeated in the symmetry-equivalent position. Because the catenane was found to form only when one monomer contained the phenylglycine or naphthylglycine side chain that forms the CH$-\pi-$CH sandwich, it appeared that this was a critical element required for self-templating of the catenane.

To gain further insight into the relative contribution of each of these non-covalent interactions[11−13] to the tightly constrained self-complementary structure, we designed a series of mutated [2]-catenanes and investigated their structure/stability relationship. This broad study is reported herein in two parts. The first focuses on structure, principally utilizing X-ray methods but also supported by solution characterization tools, including high-resolution NMR spectroscopy to probe the similarity of solid-state and solution structures, and solution-phase H/D exchange to probe structural dynamics and solvent accessibility. These tools are complemented by an in-depth mass spectrometry fragmentation analysis, which enables constitutional isomerism to be deconvoluted. In this first paper, we address the following questions: What interactions are highly conserved? Are there parts of the structure that can accommodate variation? Are the features observed in the solid state maintained in solution? Each of these questions is addressed with regard to the β-turns, hydrogen bonds, and $\pi-\pi$ stacking interactions in the "aryl barrel" and the CH$-\pi-$CH sandwich. Lastly, the solid-state and solution-phase structures of two non-symmetric catenanes made up of two different rings are investigated and compared to the C_2 symmetric catenanes.

The second paper in this series[14] utilizes mutation experiments to determine the relative contribution of each of the non-covalent interactions to the structure and the roles of multivalency and cooperativity in folding and molecular recognition in these catenanes.

These studies provide insight into the effect of a broad variety of multivalent non-covalent interactions on the stability of a self-assembled [2]-catenane in organic solvents. The intermediate size (>3k Da) and complexity of the [2]-catenane help to bridge the gap between small-molecule models and proteins without sacrificing the ability to precisely tune structure, obtain relative stability data, and structurally characterize with high-resolution methods in solution, in the solid state, and in the gas phase (MS).

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 (^{F}**3b**, ^{HO}**3c**, ^{F}**3d**^{Ph}, ^{F}**4b**) were recorded on a Bruker DRX 500 spectrometer or a Bruker 600 Cryoprobe spectrometer and processed using TOPSPIN Bruker NMR software (V. 3.0). High-resolution mass spectra were obtained on an Agilent Accurate LC-TOF mass spectrometer (Agilent Series 6220) operating in positive-ion mode with electrospray ionization source (fragmentor = 175 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. The data were analyzed using Agilent Chemstation. Accurate Mass LC-TOF analyses of DSA reactions were performed on 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 by an Accurate Mass LC/MS TOF mass spectrometer (Agilent Series 6220) in positive-ion mode with an electrospray ionization source (fragmentor = 375V). 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 semi-preparative 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. MS/MS analyses of the isolated [2]-catenanes and DSA reactions were performed on an Accurate Mass LC-QTOF (Agilent Series 6520) equipped with an Agilent Series 1200 LC instrument without/with a column (Halo-C18, 2.1 × 50 mm, 2.7 μm) in positive-ion mode with an electrospray ionization source. X-ray crystallographic analyses of [2]-catenanes ^{F}**3d**^{Ph}, ^{HO}**3c**, ^{F}**3b**, and ^{F}**4b** were performed on a Bruker-AXS SMART APEX-II system equipped with a graphite monochromator. [Please contact Dr. Peter S. White (pwhite@email.unc.edu) for correspondence regarding X-ray analyses.]

Generation and Isolation of ^{F}3d^{Ph}, ^{HO}3c, ^{F}3b, and ^{F}4b. For ^{F}**3d**^{Ph}, two 5 mM DSA solutions (D-**1d**/L,L-^{F}**2**^{Ph} = 1:1) were prepared on a 20 mL scale. Each was prepared by dissolving the mixture of D-**1d** (20.2 mg, 50.0 μmol) and L,L-^{F}**2**^{Ph} (22.9 mg, 50.0 μmol) in chloroform (20 mL) and subsequently adding trifluoroacetic acid (50 equiv, 5000 μmol, 371 μL). The solutions were allowed to sit for 6 days, and then triethylamine (50 equiv, 5000 μmol, 0.7 mL) was added to quench the dynamics. To remove the precipitates generated during the DSA process due to the poor solubility of other oligomers in chloroform and after the addition of triethylamine, the solutions were filtered, and the two were combined. The volatiles were removed in vacuo, and chloroform (50 mL) was added. This solution was washed with water (40 mL) and then dried over anhydrous $MgSO_4$. More chloroform (80

dx.doi.org/10.1021/ja302345n | J. Am. Chem. Soc. 2012, 134, 11415−11429

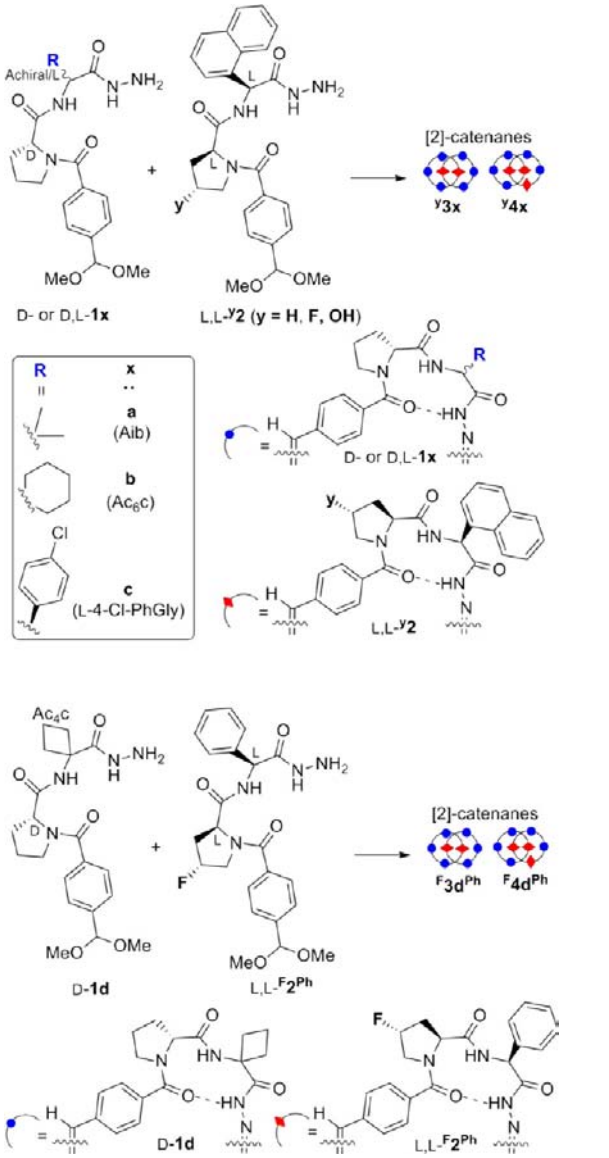
mL) was added, and the solution was heated to 50 °C (to mitigate the poor solubility of ^{F}**3d**^{Ph} in $CHCl_3$) and then filtered to remove $MgSO_4$. The volatiles were removed in vacuo again, and the purity of the crude product was determined by HPLC. Crude ^{F}**3d**^{Ph} contained 4% ^{F}**4d**^{Ph}, and without further purification, it was recrystallized from pyridine by slow evaporation at room temperature to generate crystals suitable for X-ray analysis. For ^{HO}**3c**, two biased monomer mixtures (D-**1c**/L,L-^{HO}**2** = 3:1) were prepared on a 20 mL scale. Each was prepared by dissolving D-**1c** (35.6 mg, 75.0 μmol) and L,L-^{HO}**2** (12.7 mg, 25.0 μmol) in 25% acetonitrile/chloroform (20 mL), to which was subsequently added trifluoroacetic acid (50 equiv, 5000 μmol, 371 μL). The reactions were allowed to sit for 3 days, and then triethylamine (50 equiv, 5000 μmol, 0.7 mL) was added to neutralize the acid. The solutions were combined, and the volatiles were removed in vacuo. Methanol (8 mL) was added to the resulting mixture, which was then filtered. This solution was eluted on the semi-preparative HPLC with 87% methanol/water isocratic solution to isolate ^{HO}**3c**. The collected ^{HO}**3c** was recrystallized from methanol/chloroform by slow evaporation at room temperature to obtain crystals suitable for X-ray analysis. For ^{F}**3b** and ^{F}**4b**, three DSA solutions (D-**1b**/L,L-^{F}**2** = 1:1) were prepared on a 20 mL scale. Each was obtained by dissolving a mixture of D-**1b** (27.0 mg, 50.0 μmol) and L,L-^{F}**2** (19.1 mg, 50.0 μmol) in 25% acetonitrile/chloroform (20 mL) and adding trifluoroacetic acid (50 equiv, 5000 μmol, 371 μL). The reactions were allowed to sit for 9−10 days, and then triethylamine (50 equiv, 5000 μmol, 0.7 mL) was added to quench the reaction. All solutions were combined, and the volatiles were removed in vacuo. Next, 65% methanol/chloroform (15 mL) was added to the resulting mixture, which was then filtered prior to semi-preparative HPLC separation (95% isocratic methanol/water elution). The collected ^{F}**3b** and ^{F}**4b** were recrystallized from methanol/chloroform by slow evaporation at room temperature to obtain crystals suitable for X-ray analysis.

MS/MS Analysis of Isolated ^{F}4b and DSA Reaction Mixtures Containing ^{F}4b and ^{H}4a. For the isolated ^{F}**4b**, the sample was prepared by first dissolving in methanol and then eluting with an isocratic eluent (50% methanol/water containing 0.1% formic acid) at a flow rate of 0.2 mL/min for 2 min. Injection volume was 1 μL. The MS source nebulizer gas and Vcap parameters were set to 35 psig and 3500 V, respectively. The fragmentor parameter was adjusted to either 175.0 or 375.0 V, depending on the experiments. The other MS source parameters were default settings. The ramped collision energy was adjusted by slope (typically 2) and offset (0 to 30) values. The MS1 and MS2 data were collected in the range of 100−3200 (m/z) with a reference ion (m/z = 922.009798). With a 175.0 V fragmentor setting, the targeted MS1 ion was the abundant doubly charged ion $[^{F}$***4b*** + 2H$]^{2+}$ (m/z = 1588.7219) under ESI conditions. In the experiment with a 375.0 V fragmentor setting, the targeted MS1 ion was the singly charged ion $[\boldsymbol{1b}_2{}^{F}\mathbf{2}_2 + H]^+$ (m/z = 1626.7100). For DSA reaction mixtures containing ^{F}**4b** and ^{H}**4a**, 0.1 mL of a DSA solution was diluted with 0.1 mL of 25% acetonitrile/chloroform solution prior to the injection, and then the mixture was eluted with a gradient profile (methanol/water containing 0.1% formic acid) at a flow rate of 0.5 mL/min and at 50 °C. The injection volume was typically 2 μL. MS source parameters, the ramped collision energy, and the MS1 and MS2 data collection ranges were set in the same way. For the DSA solutions generating ^{F}**4b**, the fragmentor parameters and targeted MS1 ions were the same as those for the isolated ^{F}**4b**. For the DSA solutions providing ^{H}**4a**, 175.0 and 420.0 V were chosen as the fragmentor settings. At 175.0 V, the targeted MS1 ion was the doubly charged $[^{H}$***4a*** + 2H$]^{2+}$ (m/z = 1461.1533), while at 420.0 V, the targeted MS1 ion was the singly charged $[\boldsymbol{1a}_2{}^{H}\mathbf{2}_2 + H]^+$ (m/z = 1510.6581).

RESULTS

1. Self-Assembly and Isolation of [2]-Catenanes. As shown in Scheme 1, the combination of the D-Pro-X monomer [(D- or D,L-**1x**; X = Aib, 1-aminocyclobutanecarboxylic acid (Ac_4c), 1-aminocyclohexanecarboxylic acid (Ac_6c), L-(4-chlorophenyl)glycine (L-4-Cl-PhGly)] and L-^{y}Pro-L-arylGly [L,L-^{y}**2** or ^{y}**2**^{Ph}; y = H, F or OH; aryl = Naph or Ph,

Scheme 1. [2]-Catenanes Self-Assembled from Pairs of Hydrazide/Acetal Monomers



respectively] under the acidic conditions necessary for promoting hydrazone exchange yields a variety of cyclic oligomers, with dimers, trimers, and tetramers predominating (traces of pentamers and hexamers are also detected by HPLC and LC-TOF analyses) (SI1 Figure 1).

Also appearing under these dynamic self-assembly (DSA) reaction conditions are significantly less polar species that steadily grow to a maximum after 3−5 days. LC-TOF analyses showed them to be catenated octamers as judged by the octamer to tetramer fragmentation pattern, which is characteristic of an interlocked structure (vide infra) (SI1 Figures 2−4).[15] Two [2]-catenanes dominated, a cyclic octamer (**3**) made up of two rings having identical mass, i.e. $[\mathbf{1}_3\mathbf{2}_1]_2$, and an asymmetric octamer (**4**) composed of interlocked $[\mathbf{1}_3\mathbf{2}_1]$ and $[\mathbf{1}_2\mathbf{2}_2]$ rings (Figure 2).

In some cases, other octamers having a mass of $[\mathbf{1}_4\mathbf{2}_4]$ as well as constitutional isomers of both **3** and **4** were observed in low quantities (<10%, Figure 2b). While the amount of catenane **3** was usually improved through the use of biased component mixtures (i.e., **1**:**2** = 3:1), the yield of **4** was typically not

dx.doi.org/10.1021/ja302345n | J. Am. Chem. Soc. 2012, 134, 11415−11429

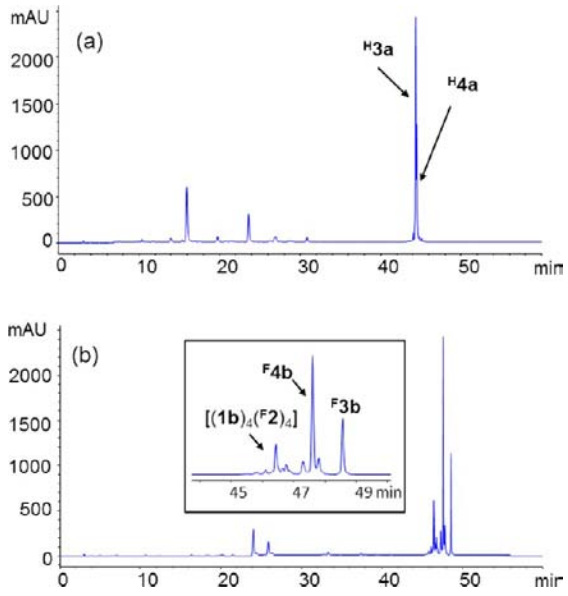


Figure 2. (a) HPLC-UV trace (289 nm) at day 7 of a DSA reaction with D-**1a** and L,L-^{H}**2** (1:1, 5 mM in MeCN/$CHCl_3$ = 1:3, 50 equiv of TFA). (b) HPLC-UV trace (289 nm) at day 7 of a DSA reaction with D-**1b** and L,L-^{F}**2** (1:1, 5 mM in MeCN/$CHCl_3$ = 1:3, 50 equiv of TFA). Inset: [2]-catenane region.

sensitive to the feed ratio. The choice of the reaction solvent, however, strongly affects the [2]-catenane speciation, with pure chloroform giving **3** exclusively while 25% MeCN/$CHCl_3$ maximized the amount of **4**. Using these conditions, a series of [2]-catenanes were self-assembled and isolated after neutralizing the reaction solution to halt hydrazone exchange; additional purification was achieved either by flash column chromatography or by semi-preparative HPLC. Slow concentration of these products from MeOH/$CHCl_3$ or pyridine led directly to X-ray-quality crystals.

For the purpose of this and the following paper, **3** corresponds to the C_2-symmetric [2]-catenane that is made up of two interlocked macrocycles assembled from three **1** monomers and one **2** monomer. The **Ph** superscript designation refers to the phenyl glycine version of **2** (cf. naphthylglycine), while the **x** and **y** designations correspond to the D-monomer and naphthylglycine variants described in Scheme 1, respectively.

2. Structural Analysis of [2]-Catenanes Having Two Identical Rings (^{y}3x). *2.1. Solid-State Analysis (X-ray).* Previously we communicated the high-resolution solid-state structural features of ^{H}**3a**.[7] Its most striking feature was its compact structure and interlocked C_2-symmetric nature (Figure 3). The structure can be loosely described as flower-like with a dense core containing the two unique L-proline-L-naphthyl glycine (L-Pro-L-NaphGly) units, from which radiate four loops that terminate in β-turns. One loop from each macrocycle is splayed out in a petal-like arrangement, while the second is tightly associated with its C_2-related pair in a pillar-like central projection (Figure 3a).

Like the petal-loops, the pillar-like loops terminate in a proline-induced β-turn. Each of the individual cyclic tetramers adopts a roughly saddle-shaped conformation, with each of the four Pro-X sequences adopting a β-turn structure that reverses the direction of the chain. The fold of the tetramers is such that the β-turns create a clip-like arrangement that positions a proline and an Aib methyl above and below the naphthyl ring

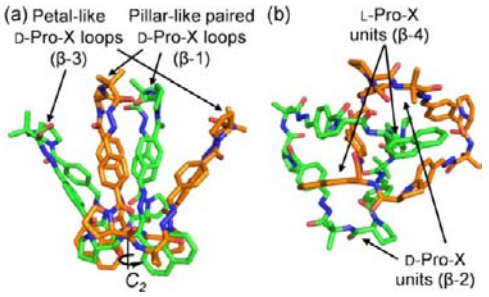


Figure 3. (a) X-ray structure of ^{H}**3a** displaying two identical interlocked tetramers. (b) View from the bottom of (a) along the C_2 axis showing a dense core. All hydrogen atoms are omitted for clarity.

of the other macrocycle. The metrical parameters of these units suggests a nearly perfect arrangement for a beneficial CH−π interaction.[16]

Similar features were observed in a series of new [2]-catenanes (^{F}**3b**, ^{HO}**3c**, ^{F}**3d**^{Ph}, Figure 4) self-assembled from structurally modified monomer units (Scheme 1). Each is flower-like with a dense core containing the aryl glycine units sandwiched between L-proline and a variable unit that includes 1-aminocyclobutanecarboxylic acid (Ac_4c), 1-aminocyclohexanecarboxylic acid (Ac_6c), and L-(4-chlorophenyl)glycine (L-4-Cl-PhGly). Like ^{H}**3a**, each tetrameric ring is saddle-shaped, with four β-turns connected by a planar, extended 1,4-hydrazone/benzamide linkage.

2.1.1. β-Turn Characterization. While proline is well known to nucleate turn structures in peptides and proteins,[10] the structure adopted by the Pro-X segments (X = variable unit) in these small-molecule receptors has not previously been established. In each of the structurally characterized catenanes, however, it is clear from metrical parameters (Φ, Ψ analysis, SI1 Table 1) that the Pro-X segment adopts a standard suite of β-turns. For the purpose of this discussion, the β-turns are identified as β-1−β-4, with β-2 (D-Pro-X unit) and β-4 (L-Pro-L-arylGly unit) making up the core, and β-1 and β-3 (D-Pro-X units) creating the turns at the periphery of the central pillar-like and petal-like structures, respectively (Figure 4). In general β-2 and β-4 were invariant, adopting a type II′ and type VIII turn, respectively.[17] Depending on the structural modifications, the peripheral β-1 and β-3 turns were more variable, with type II′ or type I′ being accessible.[17]

A comparative analysis of the β-turn metrical parameters of the four structurally characterized catenanes was informative (Table 1). Most striking was the invariability of the aryl glycine unit in the central core (β-4), as almost no dihedral angle differences (<9°) were noted across the series of structures. The D-Pro-X unit, which provides the second half of the molecular clip in the core (β-2), was also tightly constrained, though the L-4-Cl-PhGly case (^{HO}**3c**), which positions the 4-Cl-phenyl unit below the naphthyl core and is non-C_2-symmetric, displayed larger deviations. By contrast, the peripheral β-turns (β-1, β-3) were considerably more flexible, and were responsive to changes in the D-Pro-X unit. For example, the Ac_6c group in ^{F}**3b** generated high torsional angle deviation in β-3, while the Ac_4c group in ^{F}**3d**^{Ph} and the L-4-Cl-PhGly group in ^{HO}**3c** caused deviations in β-3 and, more noticeably, in β-1 where the two loops interact. This analysis shows the turn structure in the central core, i.e., the point where the CH−π−CH "clip" interactions occur, to be highly conserved across the series of

dx.doi.org/10.1021/ja302345n | J. Am. Chem. Soc. 2012, 134, 11415−11429

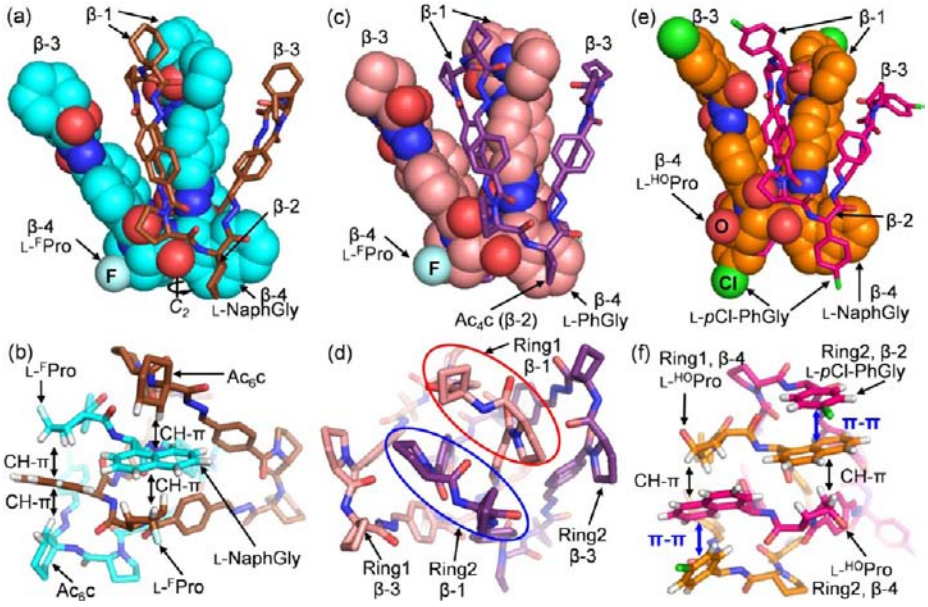


Figure 4. X-ray structure representations with one ring displayed as space filling models and the second interlocked ring as stick figures. (a) Side view of ^{F}**3b** and (b) bottom view looking up onto the CH−π−CH sandwiches. (c) Side view of ^{F}**3d**^{Ph} and (d) top view of (c) highlighting the two different β-1 turns (type I′ for ring 1 and type II′ for ring 2). (e) Side view of ^{HO}**3c** showing two interlocked tetramers and (f) bottom view of (e) highlighting the CH−π interactions between L-NaphGly and L-^{HO}Pro residues, and the two different π−π stacking interactions between L-NaphGly and L-4-Cl-PhGly moieties. All hydrogen atoms except those participating in CH−π−CH and CH−π−π sandwiches are omitted for clarity.

Table 1. Direct Comparison of Metrical β-Turn Parameters of ^{H}3a, ^{F}3b, ^{F}3d^{Ph}, and ^{HO}3c

				Δ^{F}3d^{Ph} − ^{H}3a (°)[a]		Δ^{HO}3c − ^{H}3a (°)[a]	
		dihedral angle (deg) of ^{H}3a	Δ^{F}3b − ^{H}3a (°)[a]	ring 1	ring 2	ring 1	ring 2
β-1 (loop)	Φ_{i+1}	60.1	−4.6	−4.1	2.7	1.3	−5.6
	Ψ_{i+1}	−145.9	3.4	***196.9***	***12.6***	6.4	***12.9***
	Φ_{i+2}	−52.3	0.1	***147.3***	***−24.7***	***−20.1***	***−10.4***
	Ψ_{i+2}	−35.9	−2.2	−2.1	***15.9***	***25.2***	7.5
β-2 (core)	Φ_{i+1}	58.7	6.0	0.7	3.4	−0.7	7.1
	Ψ_{i+1}	−128.9	−3.2	−0.4	−0.4	−5.6	7.6
	Φ_{i+2}	−62.7	4.7	−3.6	−2.4	***−17.9***	−6.7
	Ψ_{i+2}	−13.0	***−12.6***	−4.0	−6.0	***20.2***	1.6
β-3 (loop)	Φ_{i+1}	58.5	−6.0	−5.5	−0.5	2.2	3.3
	Ψ_{i+1}	−126.6	***−11.2***	−1.8	−1.4	−6.0	−8.9
	Φ_{i+2}	−62.5	−6.6	−6.5	−9.5	***−10.7***	−5.8
	Ψ_{i+2}	−21.9	***17.6***	7.9	***15.9***	7.1	6.6
β-4 (core)	Φ_{i+1}	−66.1	1.2	−3.4	−5.3	−5.8	−6.0
	Ψ_{i+1}	−30.7	−3.2	−1.0	−1.8	−4.8	0.8
	Φ_{i+2}	−159.0	−3.5	−1.4	0.9	−1.1	−1.7
	Ψ_{i+2}	83.0	−0.2	−6.8	−8.6	3.2	−4.2

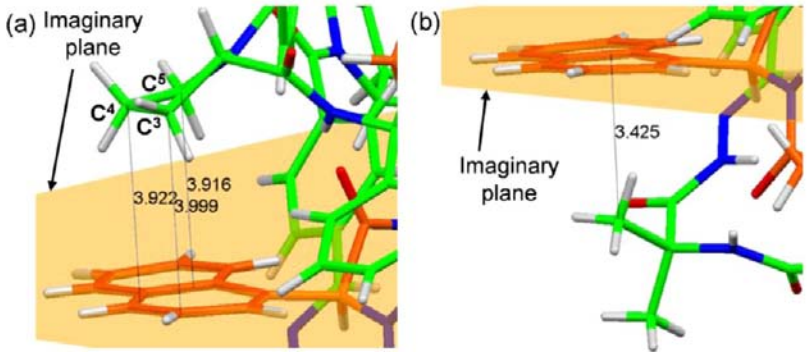
[a]The corresponding dihedral angle differences in ^{H}**3a** vs ^{F}**3b**, ^{F}**3d**^{Ph}, and ^{HO}**3c**. Dihedral angle differences of >10° are in bold italics. The rmsd metric of structural similarity to the parent structure (^{H}**3a**) obtained by PyMOL pair fitting using 48 arbitrarily chosen backbone atom pairs for ^{F}**3b** vs ^{H}**3a**, ^{F}**3d**^{Ph} vs ^{H}**3a**, and ^{HO}**3c** vs ^{H}**3a** are 0.246, 0.791, and 1.067 Å, respectively. See details in SI1 Table 2.

[2]-catenanes. The tight packing of this region reinforces the notion of a highly optimized three-dimensional structural motif which provides significant structural stabilization.

2.1.2. Hydrogen-Bonding Interactions.[18] In the peripheral turns (β-1 and β-3) of the [2]-catenane, the expected i/i+3 intra-turn hydrogen bonds (HBs) are observed though there are no inter-macrocycle HBs (SI1 Table 3). In contrast, the β-turns in the core are distorted to enable inter-ring HBs for β-2 and β-4, with only β-2 containing a typical i/i+3 HB; β-4 has its i+3 residue engaged in a conserved inter-ring HB. The series is, therefore, characterized by a total of six intra-turn hydrogen bonds for β-1, β-2, and β-3, and six inter-ring hydrogen bonds (β-2(N)→β-4′(O), β-4(N)→β-2′(O), β-2′(N)→β-4(O), β-4′(N)→β-2(O), β-4(N)→β-4′(O), and β-4′(N)→β-4(O)).[16] Except for catenane ^{F}**3d**^{Ph}, the O_i---N_{i+3} distances for inter-ring hydrogen bonds only marginally deviate from the parent catenane ^{H}**3a** (≤0.063 Å) (SI1 Table 4), indicating that the skeleton of the [2]-catenane is conserved across the series. As

dx.doi.org/10.1021/ja302345n | J. Am. Chem. Soc. 2012, 134, 11415−11429

Table 2. Direct Comparison of the Distances (Å) Characterizing CH−π or π−π Stacking Interactions in ^{H}3a, ^{F}3b, ^{F}3d^{Ph}, and ^{HO}3c



X-ray structural representations of ^{H}3a showing the distances (a) between three carbons of L-Pro CHs (C^3H, C^4H and C^5H) and an imaginary naphthyl plane of the L-NaphGly unit and (b) between the naphthyl plane and an Aib methyl group carbon atom.

				$\Delta^{F}3d^{Ph} - {}^{H}3a$ (Å)^{b,c}		$\Delta^{HO}3c - {}^{H}3a$ (Å)^{b}	
		^{H}3a (Å)^{a}	$\Delta^{F}3b - {}^{H}3a$ (Å)^{b}	ring 1	ring 2	ring 1	ring 2
L-proline and aryl	L-Pro C^3	3.999	−0.291	−0.297	−0.260	−0.215	−0.315
	L-Pro C^4	3.922	−0.311	−0.289	−0.450	−0.144	−0.374
	L-Pro C^5	3.916	−0.163	−0.074	−0.317	−0.136	−0.310
X and aryl		X = Aib	X = Ac_6c	X = Ac_4c		X = L-4-Cl-PhGly	
		3.425	0.089	0.133	0.167	0.120	0.078
			0.629^{d}				

^{a}The distances between an imaginary L-NaphGly plane and the corresponding carbons of three L-Pro CH's, and an Aib methyl carbon and the L-NaphGly plane in ^{H}3a. ^{b}The differences in the corresponding distances indicating CH−π or π−π stacking interactions in ^{H}3a vs ^{F}3b, ^{F}3d^{Ph}, and ^{HO}3c. The negative values represent the shorter distance compared to those in ^{H}3a, implying stronger interactions. For the ring assignment, see Figure 4d,f. ^{c}An imaginary phenyl plane of the L-PhGly unit was used for computing the distances in ^{F}3d^{Ph}. ^{d}Following the distance cutoff (X···M < 4.3 Å [Steiner, T.; Koellner, G. *J. Mol. Biol.* **2001**, *305*, 535−557], the distance between the L-NaphGly plane and the carbon of the second CH_2 group (C^2H_2) in Ac_6c moiety was included (see SI1 Table 5) and compared with the distance between an Aib methyl carbon and the NaphGly plane in ^{H}3a.

will be discussed in more depth in the following paper,[14] the O_i---N_{i+3} distances for two inter-ring hydrogen bonds in ^{F}3d^{Ph} (β-2(N)→β-4′(O) and β-2′(N)→β-4(O)) are shorter than those in the rest of the catenanes. The trigger for the shortening of these key inter-ring HBs is the replacement of a naphthyl group for a phenyl group, which initiates a series of changes to the network of non-covalent interactions.

2.1.3. CH−π Interactions in the Core. CH−π interactions are well-known stabilizing molecular interactions.[19,20] Since an aryl group in the L,L-**2** and **2**^{Ph} monomers is required for catenane formation and the catenane adopts a fold that sandwiches this arene between an L-proline unit and the X-moiety of the other macrocycle, CH−π interactions are reasonably inferred as an important, if not critical, stabilizing interaction. Metrical parameters from ^{H}3a indicate a constrained geometry that maximizes the CH−π interaction between naphthyl and L-Pro on one side (3.9−4.0 Å) and Aib on the opposite (3.4 Å) (Table 2). Within the new series of catenanes, ^{F}3d^{Ph} and ^{F}3b follow a similar CH−π arrangement, with a geminally substituted cyclohexane replacing the Aib methyl group of ^{H}3a below the naphthyl ring in ^{F}3b, and a cyclobutyl group replacing it in ^{F}3d^{Ph} (see SI1 Table 5 for key metrical parameters).

In the case of catenane ^{HO}3c, the naphthyl ring is pinched between L-Pro and a 4-Cl-phenyl residue, i.e., a CH−π−π molecular clip. The non-C_2-symmetric nature of ^{HO}3c is expressed in the differential π−π interaction between naphthyl and the 4-Cl-phenyl unit, one being biased toward a nearly eclipsed face-to-face interaction, while the other is better described as intermediate between an edge-to-face and face-to-face (~33°, SI1 Table 5). NMR experiments (vide infra) indicate that the complex is fluxional, as only a single C_2-symmetric structure is observed in ^{1}H NMR spectra.

A comparison of the metrical parameters for these catenanes to ^{H}3a was informative (Table 2). In ^{F}3b and ^{HO}3c, shorter distances between the naphthyl group and the F- or OH-substituted proline moieties agreed with the expected enhancement of the CH−π's electrostatic component, which would be stabilized by a more positively charged CH interacting with the arene's quadrupole.[19] Despite the weaker π−donor characteristics of Ph versus naphthyl,[21] the decrease in the CH−π distances of ^{F}3d^{Ph} compared to ^{H}3a coupled with the stronger thermodynamic stability of F vs non-F catenanes (see our next paper[14]) reinforce the argument that the CH−π interaction can be electronically manipulated, and that the more acidic CH bonds of F-substituent prolines enhance the strength. Zondlo has previously noted this effect in short peptide model systems in water.[19]

On the opposite face of the arene clip, the interactions between the aryl group and Ac_4c, Ac_6c, and 4-Cl-phenyl residues (^{F}3d^{Ph}, ^{F}3b and ^{HO}3c) were variable. The combination of a naphthyl group and the Ac_6c moiety was structurally similar to ^{H}3a, except that the conformationally restricted Ac_6c group (cf. CH_3) orients a single CH into the aromatic ring, which shifts this H resonance to higher field than any other CH−π group in the ^{1}H NMR spectrum, vide infra. The longer distance between the Ph centroid and the Ac_4c group in ^{F}3d^{Ph} (cf. ^{H}3a, Table 2) indicates that the weaker π−donor character of a Ph group causes a lengthening of the CH−π distances, implying that the core might be less tightly packed. In the same way,

dx.doi.org/10.1021/ja302345n | J. Am. Chem. Soc. 2012, 134, 11415−11429

catenane ^{HO}**3c** displays lengthened $\pi-\pi$ interactions in one of its two rings, suggesting that the optimum CH$-\pi-\pi$ arrangement cannot be simultaneously achieved in both halves of the molecule.

2.1.4. Aryl Barrel. The gross features of the catenanes are established by the connection of the peripheral β-turn loops to the dense core. This connection is achieved through 1,4-hydrazone/benzamide linkages that create a barrel-like structure. The aromatic rings interact with one another through a combination of slipped face-to-face and edge-to-face arrangements (Figure 5).[20,22] Of the eight aromatic rings making up this barrel, a number of the centroid-to-centroid distances were illuminating. For catenanes ^{H}**3a**, ^{F}**3b**, and ^{HO}**3c**, each of which has a type II′ β-1 turn, the A and A′ aryl rings (Figure 5) occupy positions at the interface of the interlocked region. The arene rings interact through a slipped face-to-face orientation with a nearly invariant plane-to-plane distance of 3.86 Å. A similar geometry but significantly shorter distance (by ∼0.2 Å) was noted for ^{F}**3d**^{Ph}, which adopts a single type I′ β-1 turn (^{F}**3d**^{Ph}) (Table 3 and SI1 Table 6). The combination of a type II′ and a type I′ in the two interacting β-1 turns establishes an H-bond between the proline amide carbonyl in ring 2 and the amide NH of the Ac$_4$c in ring 1 (N$-$O distance = 3.175 Å).[23] It thus seems reasonable that weaker CH$-\pi$ interactions in the core (i.e., Ph vs naphthyl)[21] create a more flexible core that can compensate by optimizing aromatic interactions in the barrel region.

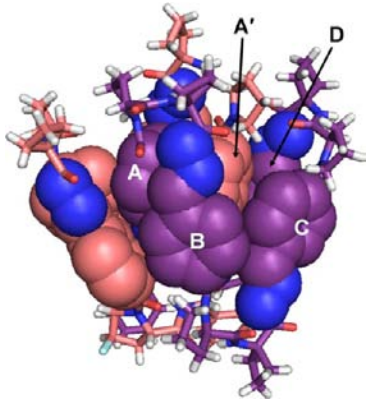


Figure 5. X-ray structural representations of ^{F}**3d**^{Ph}, emphasizing how the 1,4-hydrazone/benzamide linkages establish an aryl barrel through $\pi-\pi$ stacking interactions (face-to-face and edge-to-face). The A, B, C, and D linkages are directly connected to the core L-NaphGly (β-4), D-Pro (β-2), Ac$_4$c (β-2), and L-Pro (β-4) residues, respectively.

Other centroid distances informed on the width of the barrel and hinted at potential for compensating effects in this region. For example, compared to ^{H}**3a**'s barrel, those of ^{F}**3d**^{Ph} and ^{HO}**3c** are smaller; in particular, the A-to-C centroid distances are considerably shorter, implying that these structures benefit from greater aromatic stabilization. The barrel size, as estimated by the product of the A$-$C and B$-$D distances support such a compression in ^{F}**3d**^{Ph} and ^{HO}**3c**.

2.2. NMR Analysis. 2.2.1. CH$-\pi$ Interactions. The C_2-symmetric structure of catenane ^{H}**3a** was found to persist in solution, and despite its size, high-resolution spectra could be obtained in pyridine-d_5, which acted to spread the chemical shifts of the ^{1}H NMR resonances.[7] Close contacts between a single methyl group of Aib and a proline residue to the naphthyl ring in the clip was noted by their upfield-shifted resonances, and appropriate cross-peaks in the ROESY spectrum.[7] Catenane ^{F}**3b** also provided a sharp spectrum with well-dispersed signals that corresponded to a C_2-symmetric structure (6$-$8% MeOH-d_4/CDCl$_3$) (Figure 6a).

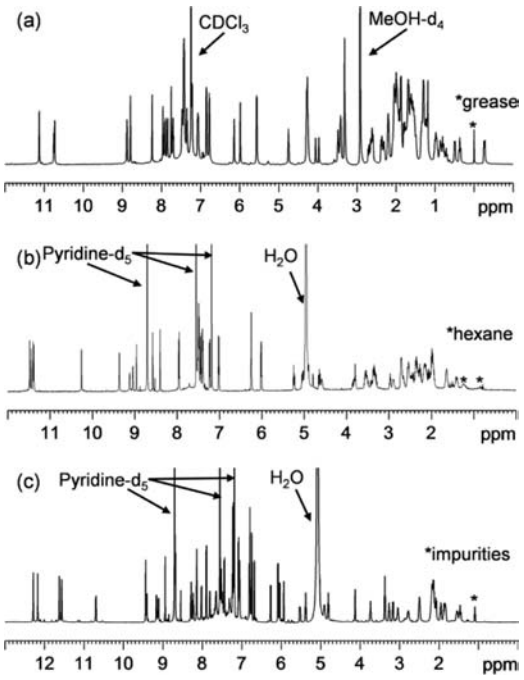


Figure 6. 500.1 MHz ^{1}H NMR spectra at 20 °C of (a) ^{F}**3b** in 7% MeOH-d_4/CDCl$_3$; (b) ^{F}**3d**^{Ph} in pyridine-d_5; (c) ^{HO}**3c** in pyridine-d_5.

Table 3. Direct Comparison of the Centroid-to-Centroid Distances (Å) in the Aryl Barrel of ^{H}3a, ^{F}3b, ^{F}3d^{Ph}, and ^{HO}3c

			Δ^{F}**3d**^{Ph} $-$ ^{H}**3a** (Å)[b]		Δ^{HO}**3c** $-$ ^{H}**3a** (Å)[b]	
	^{H}**3a** (Å)[a]	Δ^{F}**3b** $-$ ^{H}**3a** (Å)[b]	ring 1	ring 2	ring 1	ring 2
A to A′ (C_2-related ring)	3.855	0.010	−0.196	−0.196	−0.044	−0.044
A to C (intra-ring)	8.938	0.028	−0.813	−0.775	−1.162	−0.943
B to D (intra-ring)	9.121	0.249	0.298	0.234	0.683	0.289
barrel size (Å^2)	81.523[c]	2.488[d]	−4.994[d]	−4.971[d]	−5.287[d]	−6.290[d]

[a]Distances (Å) between imaginary centroids computed for each aryl linkage in the barrel as identified in Figure 5. [b]Differential distances (Å) describing both close contacts and barrel widths. Two values are reported for ^{F}**3d**^{Ph} and ^{HO}**3c**, as these compounds are non-C_2-symmetric in the solid state. Negative values described shorter distances. [c]Barrel size is simply obtained from a multiplication of two barrel distances, i.e., (distance of A to C) × (distance of B to D). [d]Differential barrel size (Å^2). Negative values described shrunken barrel areas.

dx.doi.org/10.1021/ja302345n | J. Am. Chem. Soc. 2012, 134, 11415−11429

The F-substituted L-proline residues provided a useful handle for structural analysis as the geminal H* participates in the CH−π interactions of the clip (Figure 7). The H* was assigned

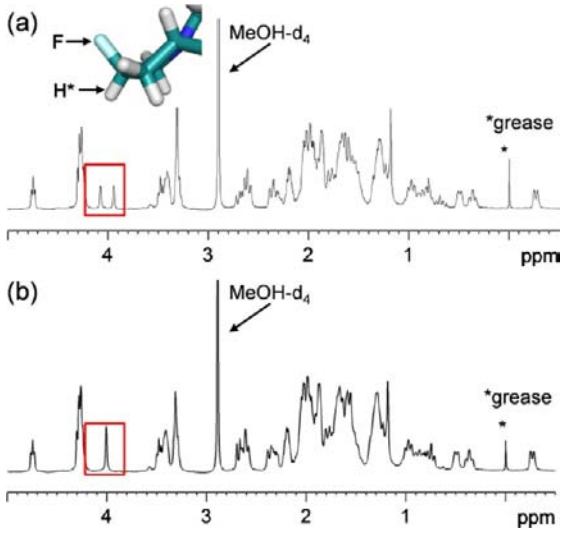


Figure 7. Portion (δ = −0.5 to 5.0 ppm) of 400.0 MHz ^{1}H NMR spectra of ^{F}**3b** at 20 °C in 7% MeOH-d_4/$CDCl_3$: (a) ^{19}F coupled and (b) ^{19}F decoupled.

to a doublet at 4.0 ppm (J_{HF} = 53.1 Hz), upfield shifted from the monomer (δ = 5.1 ppm, J_{HF} = 52.4 Hz). This assignment was confirmed by comparison to ^{1}H NMR spectra obtained with ^{19}F decoupling (Figure 7 and SI1 Figure 5). Cross-peaks indicative of CH−π interactions between the L-^{F}Pro CH's and the naphthyl group were also observed in the ROESY spectrum (Figure 8).[24] Consistent with its solid-state structure, three L-^{F}Pro CH cross-peaks to the naphthyl group were noted (see inset Figure 8). The ROESY spectrum of ^{F}**3b** also showed that one cyclohexyl hydrogen (δ = −0.2 ppm) was significantly shielded by its close interaction with the opposite face of the arene (vide supra). Unlike the Aib case, the CH group participating in the C−H−π interaction does not exchange with a non-interacting position (CH_3 rotation) and is thus more significantly shifted by its aromatic interaction. The obvious near-neighbor contacts from the solid-state structures are clearly maintained in solution.

The remaining catenanes, ^{F}**3d**^{Ph} and ^{HO}**3c**, provided high-resolution ^{1}H NMR spectra in pyridine-d_5 solvent (Figure 6b,c). Contrasting their non-C_2-symmetric solid-state structures, however, these spectra were symmetric, suggesting time-averaged dynamic behavior. In the case of ^{F}**3d**^{Ph}, this would require interconversion of either the β-1 type I′ into a type II′ turn or the isomerization of the other loop's type II′ turn into a type I′ to regenerate C_2 symmetry. A more subtle slipped π/π to edged π/π exchange is required for macrocycle equilibration in ^{HO}**3c**.

Despite representing time-averaged spectra, the proline CH units participating in the CH−π interactions of both catenanes are diagnostic and upfield-shifted in their respective ROESY spectra (SI1 Figures 11a and 13a). The CH−π interacting CH_2 unit of the Ac_4c residue in ^{F}**3d**^{Ph} are also upfield-shifted, and cross-peaks to the Naph group in its ROESY spectrum are observed. Unlike the Ac_6c case, which orients the equatorial CH to the arene, the Ac_4c analogue projects both geminal CH's

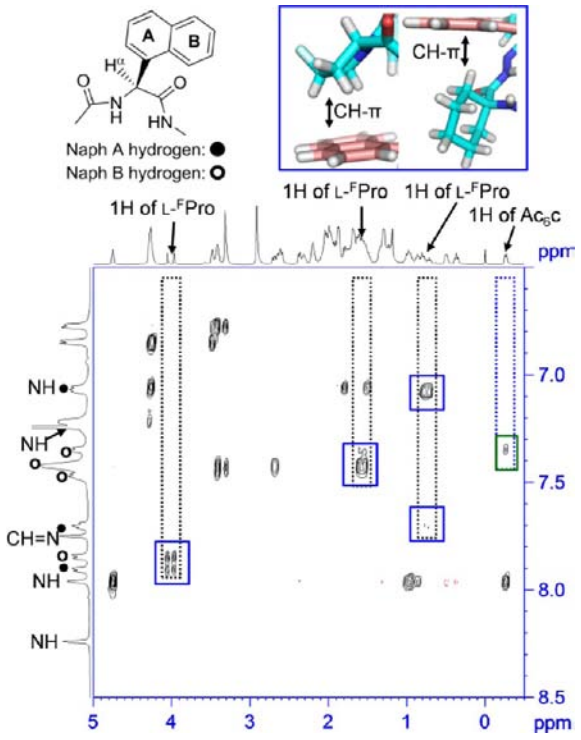


Figure 8. Portion of the 2D 600.1 MHz ROESY spectrum of ^{F}**3b** (F1 δ = −0.5 to 5.0 ppm; F2 δ = 6.5−8.5 ppm; 7% MeOH-d_4/$CDCl_3$, 20 °C). Inset: solid-state structure showing CH−π interactions of a L-proline residue and an Ac_6c group to the naphthyl group.

toward the arene, and both thus exhibit upfield shifting and cross-peaks to the Naph in the ROESY spectrum.

Despite the degree of spectral overlap in the aromatic region of ^{HO}**3c**, the 4-Cl-phenyl residue participating in the $\pi-\pi$ stack could be identified on the basis of a combination of TOCSY and ROESY spectra and its solid-state structure (SI1 Figures 12−14). Indicative were upfield-shifted aryl resonances at δ = 6.74 and 6.80 ppm (the remaining aryl signals resonate from δ = 7.05 to 7.92 ppm), which additionally showed cross peaks to the naphthyl group (SI1 Figure 14).

In addition to providing information relevant to dynamics in the periphery of the catenane, NMR analysis also revealed that some rotational motion is possible in the core of the **3** structure. Since a full assignment of the ^{1}H NMR spectrum of ^{F}**3d**^{Ph} was achieved (ROESY, COSY, TOCSY, and HSQC, SI1 Figures 9−11), these data combined with spectral integration indicated that rotation of the templating Ph−CH moiety was rapid on the NMR time scale as the chemically inequivalent ortho/ortho′ and meta/meta′ positions were time averaged (Figure 9, inset). Thus, despite the apparent high density, the near invariance of the core's structure, and the tightness of the CH−π−CH sandwich, solution dynamics can enable even highly congested groups to rotate on the NMR scale.

2.2.2. $\pi-\pi$ Interactions in the Aromatic Barrel. In the solid-state structure of each catenane, the 1,4-hydrazone/benzamide linkages assemble through a combination of slipped face-to-face and edge-to-face $\pi-\pi$-stacking interactions into a tightly packed aromatic barrel (Figure 5).[20,22] These close contacts were investigated spectroscopically in ^{F}**3d**^{Ph} due to the simplified nature of its aromatic region. Spectra in pyridine-d_5 generated well-dispersed amide NH's, imine CH═N, and aromatic signals. The COSY and TOCSY spectra enabled the aryl barrel

dx.doi.org/10.1021/ja302345n | J. Am. Chem. Soc. 2012, 134, 11415−11429

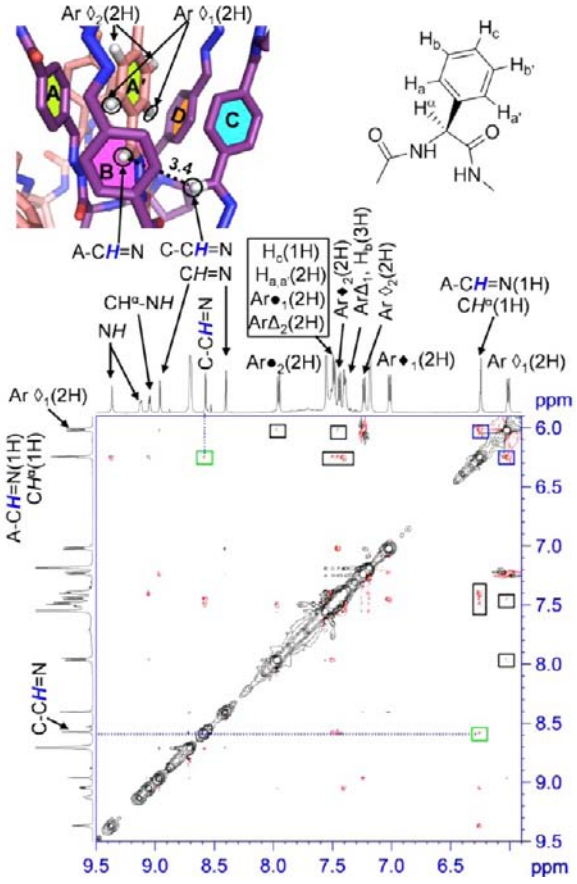


Figure 9. Portion of the 2D 500.1 MHz ROESY spectrum of ^{F}**3d**^{Ph} (F1 δ = 5.9–9.5 ppm; F2 δ = 5.9–9.5 ppm; pyridine-d_5, 20 °C). Ar◆$_{1,2}$, Ar●$_{1,2}$, and Ar$\Delta_{1,2}$ represent the remaining aryl ring CH's (B, C, and D). Inset: solid-state structure indicating Ar◇$_{1,2}$, neighboring imine CH groups in the ring A′ and C (A–C*H*=N, and C–C*H*=N).

CH's and the phenyl CH's in the core to be distinguished (SI1 Figures 9 and 10), while the HSQC spectrum differentiated the amide NH's from the imine CH's by the latter's association to an imine carbon (N=CH) (SI1 Figure 11b). In addition to signals in the expected spectral region, two resonances were consistently found upfield shifted in the ^{1}H NMR spectra (δ = 5.9–6.5 ppm).

Key to their deconvolution was the strong through-space correlation between the most upfield of these CH's (Ar◇$_1$ (2H), δ = 6.1 ppm) and a neighboring imine CH group (A-C*H*=N (1H), δ = 6.25 ppm) (ROESY, blue boxes in Figure 9), which overlaps with CH$^{\alpha}$ in this compound. In the Figure 9 inset, these correspond to the circled hydrogens, one of which is being viewed through aromatic ring B (hence the upfield shift). Each of these resonances could then be correlated to a network of three symmetry-inequivalent sets of aryl CH's (black boxes in Figure 9), thus establishing their close contact and the spatial proximity of the remaining 1,4-substituted rings (B, C, and D). Additionally noted was a strong correlation (green boxes in Figure 9) between two imine CH's (A–C*H*=N and C–C*H*=N in Figure 9 inset), consistent with one particularly close contact in the solid-state structure (3.4 Å). The high degree of structural order as measured by the preciseness of inter-aromatic contacts, and homology, as assessed by the degree to which key ROESY cross-peaks repeat across the series of catenanes, imply that the aryl barrel structure is conserved across the series and is thus an important contributor to catenane stability.

2.2.3. H/D Exchange. The four hydrazone NH's (δ = 10.7–11.3 ppm) and the four amide NH's of ^{H}**3a** resonate in the downfield portion of the ^{1}H NMR spectra (pyridine-d_5, 20 °C).[7] The amide signals were differentiated from the imine CH signals in the HSQC spectrum by the latter's correlation to their imine carbons. The β-4 amide NH was specifically identified by its COSY cross peak to the adjacent CH$^{\alpha}$ of the aryl glycine. Since these NH groups (except for the β-1 and β-3 amide NH's) participate in intra- and inter-macrocyclic hydrogen-bonding interactions (vide supra), H/D exchange reactions were utilized to investigate the inherent structural flexibility of the catenane structure.

In pyridine-d_5/D$_2$O, all NH groups of ^{H}**3a** exchanged within 15 h with the exception of the amide NH's of the β-2 and β-4 turns (Figure 10); these latter amide NH's were unchanged

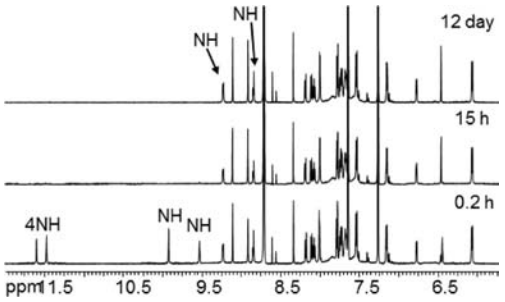


Figure 10. Portion (δ = 5.8–11.8 ppm) of the 599.8 MHz ^{1}H NMR spectra of ^{H}**3a** (pyridine-d_5 (0.6 mL) + D$_2$O (0.03 mL), 20 °C), recorded at different time from mixing.

even after 12 days. Since they are positioned near the Aib and L-Pro units participating in the CH–π interactions and additionally form inter-ring HBs (solid state), these data imply that the core created by the β-2 and β-4 turns is densely packed and relatively inaccessible. In contrast, the loops comprised of the β-1 and β-3 turns are relatively flexible and solvent accessible according to H/D exchange.

Catenane ^{HO}**3c**, which adopts a non-symmetric solid-state structure but a C_2-averaged ^{1}H NMR spectrum, was subjected to similar H/D exchange experiments to determine if this phenomenon was indicative of a more dynamic structure; if the non-symmetric solid-state structure was simply a crystal packing effect, then an enhanced H/D exchange rate was not expected. Additionally hinting at a "looser" arrangement were the longer CH–π and π–π distances between the NaphGly and the L-4-Cl-PhGly units. In this case, all NH groups exchanged within 2.6 days, and the adjacent C*H*$^{\alpha}$ transformed from a doublet into a broad singlet lacking the normally observed *H*C–N*H* vicinal coupling (Figure 11a). The contrasting exchange of the core amide NH's (β-2 and β-4), in particular, indicates that the core is less tightly packed in solution and more accessible than the Aib variant. Since the intermacrocyclic HB lengths in the core are similar to ^{H}**3a**, this higher reactivity appears to be driven by a weaker CH–π–π clip as compared to the CH–π–CH clip. Alternatively, the dynamics that accompany a structural rocking between the parallel and edge-to-face aromatic orientations might enhance the catenane's accessibility to H/D exchange (Figure 4e,f).

For catenane ^{F}**3d**^{Ph}, the unique β-1 turn of one ring and the longer CH–π distances coupled with the symmetric solution

dx.doi.org/10.1021/ja302345n | J. Am. Chem. Soc. 2012, 134, 11415–11429

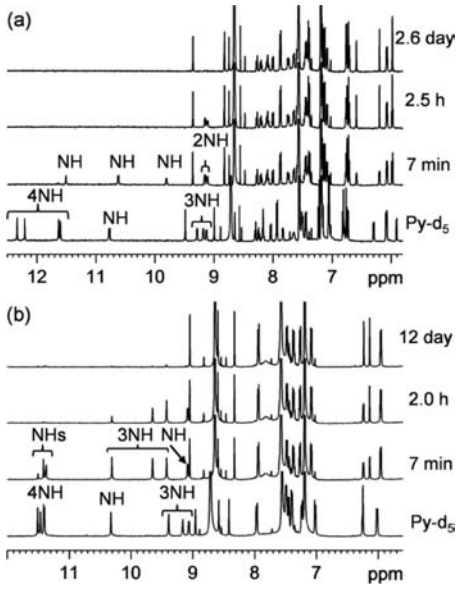


Figure 11. (a) Portion (δ = 5.8−12.5 ppm) of the 500.1 MHz ^{1}H NMR spectra of ^{HO}**3c** (pyridine-d_5 (0.6 mL) + D_2O (0.03 mL), 20 °C), recorded at different time from mixing. The bottom spectrum is the same portion of ^{HO}**3c** in pyridine-d_5 before adding D_2O. (b) Portion (δ = 5.6−12.0 ppm) of the 500.1 MHz ^{1}H NMR spectra of ^{F}**3d**^{Ph} (pyridine-d_5 (0.6 mL) + D_2O (0.03 mL), 20 °C), recorded at different time from mixing. The bottom spectrum is the same portion of ^{F}**3d**^{Ph} in pyridine-d_5 before adding D_2O.

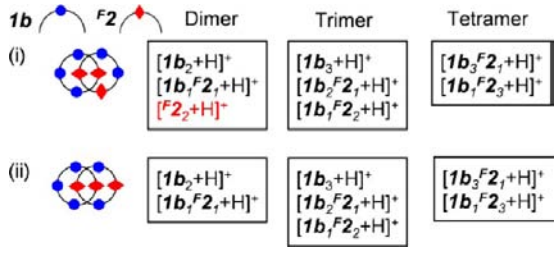
structure suggested a compound that was more flexible and dynamic than the parent ^{H}**3a**. In this case, all hydrazone NH groups exchanged within 2 h in pyridine-d_5/D_2O and the H/D exchange of all amide NH groups was complete within 12 days (Figure 11b). The rates of H/D exchange rates therefore follow the trend of ^{HO}**3c** > ^{F}**3d**^{Ph} > ^{H}**3a**. For the catenane ^{F}**3b**, its poor solubility in water-miscible polar solvents precluded its study in an analogous fashion.

Together these results indicate that H/D exchange reactions can be used to probe for dynamic processes that perturb the structure of what is otherwise a highly compact assembly. While perhaps related to the dynamic processes that symmetrize the ^{1}H NMR spectra, these latter "breathing" dynamics occur on the sub-millisecond time scale and are thus considerably faster than H/D exchange. Perturbations in the metrical parameters thus correlate well to the longer time scale processes that H/D exchange reveals.

3. Self-Assembled [2]-Catenanes with Rings of Different Mass (^{F}4b). *3.1. LC-MS/MS Analyses of ^{F}4b and ^{H}4a by LC-QTOF.* Previous MS/MS analysis of the minor [2]-catenane ^{H}**4a** implied that it was composed of one [**1a**·**1a**·**1a**·^{H}**2**] ring and a symmetric tetramer with alternating **1a** and ^{H}**2** units, [**1a**·^{H}**2**·**1a**·^{H}**2**].[7] This suggestion was based on the lack of a [^{H}**2**$_2$+H]$^+$ dimeric daughter ion (Scheme 2). To confirm the structure, we investigated [2]-catenane ^{F}**4b** since it was the major compound in the corresponding DSA mixture. LC-MS/MS analysis on purified ^{F}**4b** and a DSA solution containing the catenane ^{F}**4b** (plus others) by LC-QTOF significantly improved our previous experiments' resolution, sensitivity, and mass accuracy.

For both samples, the doubly charged ion ([^{F}***4b*** + 2H]$^{2+}$, m/z = 1588.7219) was selected for fragmention on the basis of its high abundance under the electrospray conditions. Both analyses showed two tetrameric daughter ions ([***1b***$_3$^{F}**2**$_1$ + H]$^+$ and [***1b***$_2$^{F}**2**$_2$ + H]$^+$), three trimeric ions, and two dimeric daughter ions ([***1b***$_2$ + H]$^+$ and [***1b***$_1$^{F}**2**$_1$ + H]$^+$) with good amplitude (Figure 12 and SI1 Figure 15). The key dimeric daughter ion [^{F}**2**$_2$ + H]$^+$ (Scheme 2), which only occurs in the non-alternating [**1b**$_2$^{F}**2**$_2$] structure, was observed in both cases, but with only trace intensity.

Scheme 2. Possible Daughter Ions for the Isomers of ^{F}4b



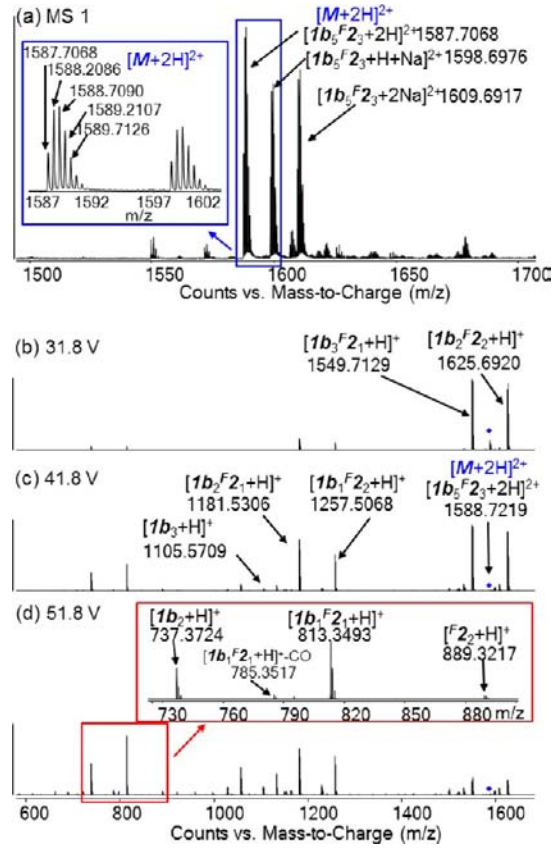


Figure 12. (a) Targeted doubly charged ion in MS1 stage for LC-MS/MS analyses. LC-QTOF (CID MS/MS) spectra of the isolated ^{F}**4b** at a collision energy of (b) 31.8, (c) 41.8, and (d) 51.8 V. The y-axis is normalized.

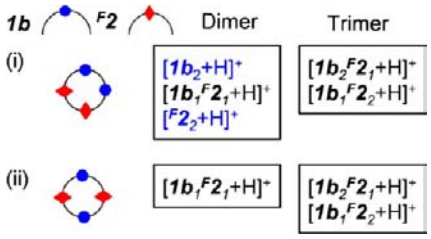
Since we have previously demonstrated that hydrazones are susceptible to ion−ion crossover in the mass spectrometer, it was possible that this fragment formed during the electrospray ionization process.[25] An LC-MS/MS analysis of the DSA sample containing ^{H}**4a** provided the same net result: only traces of [^{H}**2**$_2$ + H]$^+$ ion (SI1 Figure 16).

To investigate the possibility that the traces of [^{F}**2**$_2$ + H]$^+$ were being generated through ion-crossover pathways, an alternative MS approach was developed. By manipulating the MS parameters, the catenanes's two tetrameric ions could be

dx.doi.org/10.1021/ja302345n | J. Am. Chem. Soc. 2012, 134, 11415−11429

generated in their singly charged state in MS stage 1 (Scheme 3). The structures of these ions were then ascertained by a

Scheme 3. Possible Daughter Ions for the Isomers of Tetrameric $[1b_2{}^F2_2]$ Macrocycle



secondary fragmention in MS2. If the selected tetrameric ring was alternating ((ii) in Scheme 3), only the hetero-dimeric daughter ion would be observed, while structure (i) would provide all three possible dimeric ions.

Following this protocol, a pure sample of ^{F}**4b** was decomposed into its two tetrameric singly charged ions ($[\boldsymbol{1b_3{}^F2_1} + H]^+$ and $[\boldsymbol{1b_2{}^F2_2} + H]^+$) in the MS1 stage. The key tetrameric ion $[\boldsymbol{1b_2{}^F2_2} + H]^+$ was then selected and additionally fragmented. Since all possible dimeric daughter ions were detected (Figure 13 and SI1 Figure 17), these data suggested that the structure of ^{F}**4b** was in fact (i) in Scheme 2.

A similar experiment on ^{H}**4a**, that is, the decomposition of ^{H}**4a** into two tetrameric singly charged ions ($[\boldsymbol{1a_3{}^H2_1} + H]^+$ and $[\boldsymbol{1a_2{}^H2_2} + H]^+$) in the MS1 stage followed by the selection and fragmentation of the key tetrameric ion $[\boldsymbol{1a_2{}^H2_2} + H]^+$, showed all three possible dimeric daughter ions ($[\boldsymbol{1a_2} + H]^+$, $[\boldsymbol{1a_1{}^H2_1} + H]^+$, and $[^H\boldsymbol{2_2} + H]^+$), although the key homodimeric ion $[^H\boldsymbol{2_2} + H]^+$ was of low intensity (SI1 Figure 18). We therefore conclude that the more favorable constitutional isomers of ^{H}**4a** and ^{F}**4b** are in fact (i) in Scheme 2, and that tetramer fragmentation to $[^H\boldsymbol{2_2} + H]^+$ is unfavorable. These new data support a correction to our previous assignment as (ii) in Scheme 2.[7]

3.2. Solid-State Analysis. The catenane ^{F}**4b** was obtained as single crystals, and high-quality X-ray diffraction data were obtained. The octameric [2]-catenane consists of two different interlocked tetramers: ring 1 is composed of three **1b** and one ^{F}**2** units, while ring 2 contains two units of **1b** and ^{F}**2** arranged in a non-alternating manner [**1a**·**1a**·^{F}**2**·^{F}**2**] (Scheme 2 (i) and Figure 14). This structure confirmed the MS data and additionally added that the D-Pro-Ac$_6$c occupied the β-1 and β-2 turn positions, while L-^{F}Pro-L-NaphGly occupied the β-3 and β-4 positions.

The structure of ^{F}**4b** is similarly flower-like, but compared to ^{F}**3b**, it has one of its petal-like loops kinked away from the pillar-like central region. Each tetrameric macrocycle is again saddle-shaped, with four β-turns connected by an extended 1,4-hydrazone/benzamide linkage. The metrical parameters of each β-turn indicate that, while all five D-Pro-Ac$_6$c units adopt type II′ turns and the two L-^{F}Pro-L-NaphGly units in the core adopt pseudo type VIII or type VIII turns (β-4 turns in ring 1 and ring 2, respectively), the unique L-^{F}Pro-L-NaphGly unit adopted a type I turn (β-3 turn in ring 2) (SI1 Table 7). This latter turn is that most commonly encountered in L-Pro-L-X structures.[26] Conversion of a type I turn to the type VIII required for β-4 requires an ∼130° rotation[27] of the L-NaphGly i+2 carbonyl (Ψ_{i+2}). This carbonyl is then positioned to participate in a key inter-ring HB (SI1 Tables 7 and 9). In hindsight, the greater

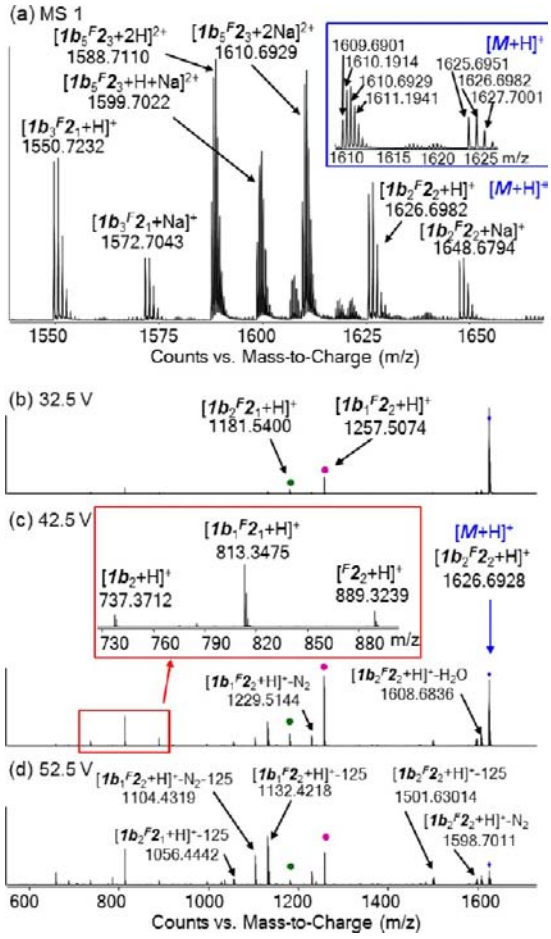


Figure 13. (a) Tetrameric singly charged ion ($[\boldsymbol{1b_2{}^F2_2} + H]^+$) generated in the MS1 stage for LC-MS/MS analyses. LC-QTOF (CID MS/MS) spectra of the targeted tetrameric ion at a collision energy of (b) 32.5, (c) 42.5, and (d) 52.5 V. The y-axis is normalized. Two filled circles (green and pink) indicated two trimeric daughter ions ($[\boldsymbol{1b_2{}^F2_1} + H]^+$ and $[\boldsymbol{1b_1{}^F2_2} + H]^+$), respectively.

adaptability of the β-1 and β-3 turns suggests that if any subunit were to be replaced by a less optimal version, it would be β-1 or β-3 (especially); the β-2 and β-4 sites would be the least accommodating to a non-optimal replacement.

Compared to the identical macrocycles in ^{F}**3b**, ring 1 exhibits larger torsional angle deviations (SI1 Table 7). While the core was invariant across the **3** series, the β-4 metrical parameters of ^{F}**4b** deviate from ^{F}**3b**, suggesting significant relay of remote structural information back to the core, and an alternative structural solution to minimizing the overall energy. Although the turns in the core deviate from those in ^{F}**3b**, the structure still effectively sandwiches the naphthyl ring between the D-^{F}Pro and Ac$_6$c units of the other macrocycle with CH−π metrical parameters not unlike those observed in the **3** series (SI1 Table 8). Based on CH---π distances, the interactions between proline and NaphGly in each ring are weaker than in the C_2-symmetry analogue ^{F}**3b** but still stronger than those observed in ^{H}**3a**. The interactions between NaphGly and Ac$_6$c in each ring are, however, weaker than those observed in both ^{F}**3b** and ^{H}**3a**, which is likely the result of the adjacent accommodating β-3 turn of an unfavorable turn type (type I).

Like its symmetric analog ^{F}**3b**, ^{F}**4b** has six turn-stabilizing hydrogen bonds in β-1, β-2, and β-3 and six inter-ring HBs (between β-2 and β-4 of both rings and between the β-4 turns

dx.doi.org/10.1021/ja302345n | J. Am. Chem. Soc. 2012, 134, 11415−11429

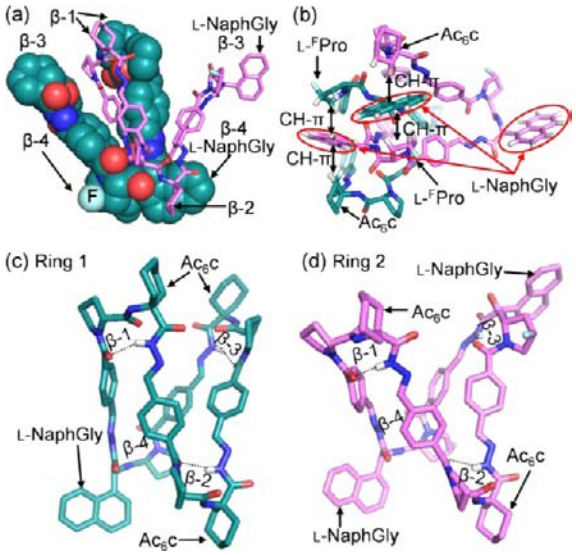


Figure 14. (a) X-ray structure representations of ^{F}**4b**, with one ring displayed in space-filling and the second inter-locked ring in stick. (b) Bottom view looking down onto the CH−π−CH sandwiches. Among three L-^{F}Pro-L-NaphGly residues, the two core residues form CH−π interactions, while the β-3 loop unit does not. (c) A single tetrameric ring of ^{F}**4b** composed of three **1b** and one ^{F}**2** units (ring 1, [**1b**$_3$^{F}**2**$_1$]). (d) A single tetrameric ring of ^{F}**4b** composed of two **1b** and two ^{F}**2** units (ring 2, [**1b**$_2$^{F}**2**$_2$]). Only selected hydrogen atoms are displayed.

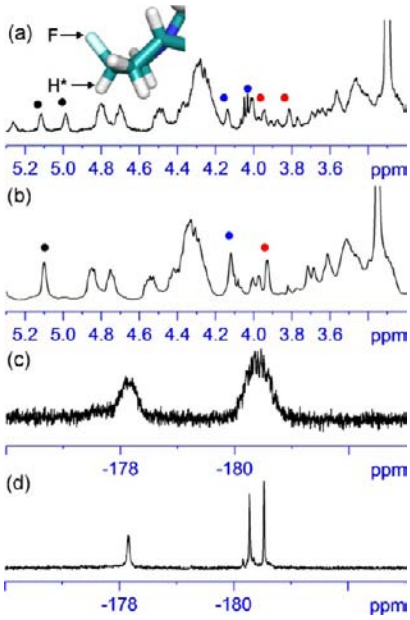


Figure 15. Portion (δ = 3.2−5.3 ppm) of the 400.0 MHz ^{1}H NMR spectra of ^{F}**4b** at 20 °C in 7% MeOH-d_4/CDCl$_3$ (a) coupled with ^{19}F and (b) decoupled with ^{19}F. The filled circles (red, blue, and black) indicate the H* signals of three L-^{F}Pro residues described in the insert picture. A portion (δ = −176 to −183 ppm) of the 400.0 MHz ^{19}F NMR spectra of ^{F}**4b** at 20 °C in 7% MeOH-d_4/CDCl$_3$, (c) ^{1}H coupled and (d) ^{1}H decoupled.

of each ring) (SI1 Table 9). The 1,4-hydrazone/benzamide linkages connecting the β-turns also similarly arrange to form an aryl barrel (SI1 Table 10). Like catenane ^{F}**3d**^{Ph}, which contains weak CH−π interaction units (PhGly-Ac$_4$c), ^{F}**4b** also displays a shortened inter-ring HBs next to the weak CH−π moiety (NaphGly-Ac$_6$c). Moreover, this effect also propagates to the adjacent aryl barrel region where the interfacial (A−A′) and cross barrel (A−C) distances also shorten. It is clear that these non-covalent interactions are linked phenomena.

Of course, since **4** is a minor product in most DSA reactions, the combination of F-proline, naphthylglycine, and Ac$_6$c appears to be uniquely able to compensate for an unfavorable unit at β-3.

3.3. NMR Analysis. Like ^{F}**3b**, the catenane ^{F}**4b** also provided a sharp spectrum with well-dispersed signals in pyridine-d_5 (SI1 Figure 19). The eight hydrazone NH's and three (doublet) amide NH's (adjacent to CH$^{\alpha}$ in the L-NaphGly residue) were well resolved and are consistent with an asymmetric structure having two different tetrameric macrocycles and three L-Pro-L-NaphGly residues. The lack of symmetry, however, made for significant complexity, which could be partially overcome by direct comparison to the symmetrical and similarly well resolved ^{F}**3b** (7% MeOH-d_4/CDCl$_3$). Although this solvent choice caused partial H/D exchange of some NH's, the spectra were comparable to ^{F}**3b**.

Since catenane ^{F}**4b** has three F-substituted L-proline moieties and only two are engaged in CH−π interactions, the two upfield-shifted doublets at 3.9 and 4.1 ppm (4.0 ppm in the ^{1}H NMR spectrum of ^{F}**3b**) were assigned to the geminal H* in each of the β-4 turns (Figure 15a). The downfield doublet at 5.0 ppm (5.1 ppm in the ^{1}H NMR spectrum of monomer ^{F}**2**) was assigned to H* in the β-3 turn fluoro-proline of ring 2; ^{19}F-decoupled ^{1}H NMR spectra were supportive (Figure 15b). Using the upfield-shifted H* doublets and the upfield-shifted CH's in the Ac$_6$c as an anchor point, two sets of the F-substituted L-Pro units and the Ac$_6$c protons involved in the CH−π interactions were identified by TOCSY (SI1 Figure 20a). Although spectral complexity made it difficult to identify the Naph groups involved in CH−π interactions, the TOCSY spectrum identified six of the eight 1,4-benzamide groups, which by the process of elimination narrowed the Naph region (SI1 Figure 20b). The HSQC spectrum identified the imine C*H*═N (total 8) and three CH$^{\alpha}$ resonances (SI1 Figure 21).

Based on these data, the proline-naphthyl and Ac$_6$c-naphthyl CH−π interactions could be deduced from the ROESY spectrum (SI1 Figure 23). While the downfield H* assigned to β-3 lacked cross-peaks to the naphthyl group, several upfield-shifted proline CH's did show cross-peaks. Similarly, only two sets of upfield cyclohexyl CH resonances corresponding to the two inequivalent Ac$_6$c CH's displayed such cross-peaks to the Naph groups (SI1 Figure 23).

Like ^{F}**3b** and ^{F}**3d**^{Ph}, uniquely upfield-shifted resonances for aryl and imine CH's in the aryl barrel region of ring 1 and ring 2 were also noted in the ROESY spectrum, this time doubled due to the lack of symmetry (SI1 Figure 24). The conservation of these and numerous other close contacts across the **3** series and in ^{F}**4b** thus implies that the solid-state structures are preserved in solution.

The flexibility and dynamics of ^{F}**4b** were also investigated through H/D exchange reactions in pyridine-d_5/D$_2$O. Within 3.4 days, all NH groups exchanged and the three C*H*$^{\alpha}$s next to the Naph amide NH's collapsed into the broad singlets that are indicative of vicinal H/D exchange (SI1 Figure 25). Overall, ^{F}**4b** is much less dynamic and/or accessible to H/D exchange than a non-folded monomer, but still looser and more dynamic than the symmetric catenane ^{H}**3a**.

dx.doi.org/10.1021/ja302345n | J. Am. Chem. Soc. 2012, 134, 11415−11429

DISCUSSION

We have investigated the features of several octameric [2]-catenane structures derived from dipeptide hydrazide A−B monomers in the solid state, along with solution characterization through high-resolution NMR spectroscopy and H/D exchange. For catenane class **3**, the structures are composed of two C_2-related interlocked rings, each of which is assembled from three units of **1** and one unit of **2**. By contrast, catenane class **4** comes from the asymmetric assembly of a [**1**·**1**·**1**·**2**] and a [**1**·**1**·**2**·**2**] macrocyclic unit. In all cases (class **3** and **4**), monomer **2** is of the L-Pro-L-ArylGly stereochemistry, and yields for the catenane were maximized when the aryl group is naphthyl, though phenyl is sufficient to template self-assembly. The family of catenanes can be described as flower-like with a compact core where the aryl of the arylglycine is pinched between two loops of the saddle-shaped tetrameric ring (Figure 1). The key core formation elements are CH−π interactions provided by a proline on one side and the methyl of an Aib on the other. Six inter-ring hydrogen bonds additionally stabilize the interior core structures. From this highly conserved core emerge petal-like loops that terminate in β-turns. Each Pro-X unit adopts a standard suite of β-turns (types I, I′, II′, and VIII) which have the effect of reversing the chain by ∼180°; deviations from 180° affect the spatial projection of the flat extended hydrazone/benzamide aromatic units. The core turns are invariant and adopt a type II′ turn for β-2 and a type VIII turn for β-4, while the peripheral turns (β-1 and β-3) are more variable (types II, I, and I′). These β-turns additionally organize a tightly packed aryl barrel with little free volume in the static structure, though transient deformations can accommodate aryl rotations that are fast on the NMR time scale. It is in this region of the structure that the two macrocycles thread (Figure 16). The barrel is characterized by an array of slipped face-to-face and edge-to-face π-stacking interactions[20] and resembles the aromatic clusters observed in the interior of proteins.[28] Analysis of these structural features reveals that the catenane is an assembly of interwoven non-covalent interactions and that these interactions combine to maximally stabilize the catenane. For example, when the core CH−π interactions was weakened by substituting a naphthyl for an aryl π-donor (PhGly) or an unfavorable turn (type I β-3 in catenane **4**), reinforcing interactions elsewhere in the structures compensated for the loss of stabilization (contracted inter-ring hydrogen bonds or compressed aromatic interactions in the aryl barrel). The following article examines the interrelated nature of these interactions.[14]

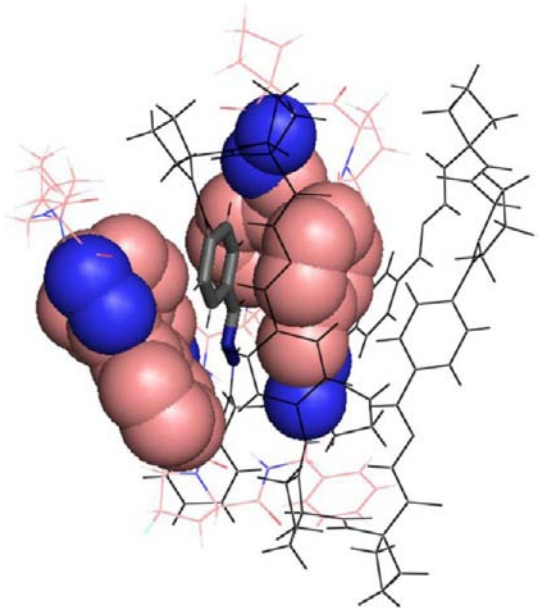


Figure 16. X-ray structural representations of ^{F}**3d**^{Ph} emphasizing how the benzamide linkages can create an aromatic pocket for binding and threading a second benzamide chain.

High-resolution NMR spectroscopy suggests that the solid-state structures are largely conserved in solution, and full assignments were achieved in several cases. H/D exchange in protic D-solvents were also informative, showing that the loop regions undergo H/D exchange at relatively fast rates, while the interior core structures are much less dynamic. In one case, two key amide NH's did not exchange over the course of weeks, even in an aggressive solvent (pyridine/D_2O). We take these data to reflect the degree of motion that is possible in the structure. The loops which are more flexible and exposed are readily deuterated, while the interior core NH's are constrained and inflexible, and are thus less prone to H/D exchange. Comparisons across the series of catenanes reveal that where structural deviations occur from the baseline structure, these distortions correlate with higher H/D exchange rates.

The β-turns are also important, as their structures act to project the 1,4-benzamide−hydrazone linkers that create the aromatic binding pocket. This binding pocket, which can be viewed as being optimized to bind an identical ring, has an opening to accommodate the entrance of a suitable guest (Figure 16). Since a number of turn types are available in a mixed Pro-X library, one can envision how combinations of turn geometries can conspire to create an adaptable class of receptors that through dynamic unit interchange can become optimized for guest-binding.[5]

When one of the tetramers is computationally excised from the interlinked structure previously described, the reversals of peptide direction achieved by the β-turns become clear. The view showing these linkers in Figure 16 is additionally informative. Note how the benzamide groups of the aryl barrel vertically align to create a deep aromatic groove that is open on one side and "binds" the threaded benzamide of the other ring. Such a picture suggests that non-catenated tetramers should be capable of adopting similar conformers that present deep aromatically lined pockets for the recognition of flat structures. Previous work by us and others has shown that non-catenated tetramers comprised of related monomers bind guests with aromatic groups like protonated nucleotides and cinchonidine salts.[9c,d] The current results suggest a mechanism and structure for this molecular recognition.

In addition to displaying a benzamide lined binding pocket, the peptide turns provide points for hydrogen bond donation and acceptance at the barrel's periphery. This display of polar functionality above and below a non-polar interior has rough parallels with cyclodextrins, calixarenes, and cucubiturils, all well-established molecular receptors.[29]

This analysis therefore suggests that an important component of the binding capability of these structures comes from the ability to present an aromatically lined cleft that is predisposed for binding. One presumes that such a binding site could accommodate flat arene rings via π−π interactions as displayed in the [2]-catenane structures described above, but could also utilize strong cation−π and

dx.doi.org/10.1021/ja302345n | J. Am. Chem. Soc. 2012, 134, 11415−11429

perhaps even CH−π interactions to bind alternative structural motifs. The display of polar groups above and below this site could then provide secondary or tertiary points of attachment as the [2]-catenanes take advantage of herein. Although we have no direct evidence for this binding mode in any DCC-selected host−guest complexes, the studies described here provide our first structural glimpses into the mechanisms of molecular recognition by this adaptable class of compounds.

CONCLUSION

The [2]-catenane structures described herein contain many of the hallmarks of a protein. (1) They are relatively large in size (>3 kDa), with a hydrodynamic volume that is large enough to benefit from NMR pulse sequences designed for macromolecules. (2) They have clear components of secondary, tertiary, and quaternary structure, along with a very strong sequence-dependent structure/stability relationship. (3) Like proteins, they additionally contain distinct regions of variable dynamic behavior (as revealed by H/D exchange). Despite these many similarities, however, the [2]-catenanes simultaneously take on the properties of small-molecule receptors, and are thus amenable to solid-state, solution-state, and gas-phase characterization tools. Their structures can be readily manipulated with non-biogenic amino acids to test the effect of electronic and/or steric perturbations, they are amenable to relative stability determinations via their self-assembly properties, as described in the next paper,[14] and they self-assemble in organic solvents. This balance of physical properties positions them at a bridge point between proteins and small molecules, and enables one to begin investigating how diverse non-covalent interactions acting in a multi-valent format affect structure, function, and stability in organic solvents, where the balance of non-covalent forces differ from those in water. These complex three-dimensional structures provide a promising scaffold for pre-organization of functional groups needed to create enzyme-like catalysts.

ASSOCIATED CONTENT

Supporting Information

Synthesis of monomers and their physical data, solid-state metrical parameters, physical data, HPLC traces, MS spectra and various 2D NMR spectra (COSY, TOCSY, HSQC, ROESY) of [2]-catenanes (^{F}**3d**^{Ph}, ^{F}**3b**, ^{HO}**3C**, ^{F}**4b**), LC-MS/MS analyses of ^{F}**4b** and ^{H}**4a**, crystal data of catenane ^{F}**3d**^{Ph}, ^{F}**3b**, ^{HO}**3C** and ^{F}**4b**. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

mgagne@unc.edu; mlwaters@unc.edu

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank the Defense Threat Reduction Agency (DTRA) for support (HDTRA1-10-1-0030). S.J.L. thanks the Army Research Office for support. We thank Dr. Marc ter Horst for helpful discussions of NMR spectroscopy techniques.

REFERENCES

(1) For several recent reviews of this broad topic, see: (a) Chakrabarty, R.; Mukherjee, P. S.; Stang, P. *J. Chem. Rev.* **2011**, *111*, 6810−6918. (b) Cragg, P. J. *Supramolecular Chemistry: From Biological Inspiration to Biomedical Applications*; Springer; Dordecht, London, 2010. (c) Stupp, S. I. *Nano Lett.* **2010**, *10*, 4783−4786. (d) Zayed, J. M.; Nourel, N.; Rauwald, U.; Scherman, O. A. *Chem. Soc. Rev.* **2010**, *39*, 2806−2816. (e) Steed, J. W., Atwood, J. L. *Supramolecular Chemistry*, 2nd ed.; Wiley: Chichester, UK, 2009. (f) Rosen, B. M.; Wilson, C. J.; Wilson, D. A.; Peterca, M.; Mohammad, R. I.; Percec, V. *Chem. Rev.* **2009**, *109*, 6275−6540.

(2) (a) Tewari, A. K. *Bioorg. Med. Chem.* **2008**, *16*, 126−143. (b) Prosinecki, V.; Fraisca, P. F. N.; Gomes, C. M. *Curr. Proteomics* **2007**, *4*, 44−52. (c) Searle, M. S. *J. Chem. Soc., Perkin Trans. 2* **2001**, 1011−1020.

(3) Recent reviews and model systems: (a) Vasudev, P. G.; Chatterjee, S.; Shamala, N.; Balaram, P. *Chem. Rev.* **2011**, *111*, 657−689. (b) Graboroski, S. J. *Chem. Rev.* **2011**, *111*, 2597−2625. (c) Nishio, M. *Phys. Chem. Chem. Phys.* **2011**, *13*, 13878−13900. (d) Takahashi, O.; Kohno, Y.; Nishio, M. *Chem. Rev.* **2010**, *110*, 6049−6076. (e) Chen, Z.; Urban, N. D.; Gao, Y.; Zhang, W.; Deng, J.; Zhu, J.; Zeng, X. C.; Gong, B. *Org. Lett.* **2011**, *13*, 4008−4011. (f) Zhang, P.; Chu, H.; Li, X.; Feng, W.; Deng, P.; Yuan, L.; Gong, B. *Org. Lett.* **2011**, *13*, 54−57. (g) Carroll, W. R.; Zhao, C.; Smith, M. D.; Pellechia, P. J.; Shimizu, K. D. *Org. Lett.* **2011**, *13*, 4320−4323. (h) Zondlo, N. J. *Nat. Chem. Biol.* **2010**, *6*, 567−568. (i) Gan, H.; Gibb, B. C. *Chem. Commun.* **2012**, *48*, 1656−1658. (j) Gibb, C. L.; Gibb, B. C. *J. Am. Chem. Soc.* **2011**, *136*, 7344−7347. (k) Laughrey, Z.; Gibb, B. C. *Chem. Soc. Rev.* **2011**, 363−386.

(4) A few recent model systems relating the protein folding: (a) Zheng, H.; Comeforo, K.; Gao, J. *J. Am. Chem. Soc.* **2009**, *131*, 18−19. (b) Bhayana, B.; Wilcox, C. S. *Angew. Chem., Int. Ed.* **2007**, *46*, 6833−6836. (c) Chen, J.; Im, W.; Brooks, C. L., III *J. Am. Chem. Soc.* **2006**, *128*, 3728−3736. (d) Gardner, R. R.; McKay, S. L.; Gellman, S. H. *Org. Lett.* **2000**, *2*, 2335−2338. (e) Gardner, R. R.; Christianson, L. A.; Gellman, S. H. *J. Am. Chem. Soc.* **1997**, *119*, 5041−5042.

(5) For reviews, see: (a) Otto, S. *Acc. Chem. Res.* **2012**, DOI: 10.1021/ar200246j. (b) Cougnon, F. B. L.; Sanders. *Acc. Chem. Res.* **2012**, DOI: 10.1021/ar200240m. (c) Miller, B. C., Ed. Dynamic Combinatorial Chemistry. In *Drug Discovery, Bioorganic Chemistry and Material Science*; Wiley: Hoboken, NJ, 2010. (d) Reek, J. N. H., Otto, S., Eds. *Dynamic Combinatorial Chemistry*; Wiley-VCH: Weinheim, 2010. (e) Ladame, S. *Org. Biomol. Chem.* **2008**, *6*, 219−226. (f) Ludlow, R. F.; Otto, S. *Chem. Soc. Rev.* **2008**, *37*, 101−108. (g) Lehn, J.-M. *Chem. Soc. Rev.* **2007**, *36*, 151−160. (h) 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.

(6) Cousins, G. R. L.; Poulsen, S. A.; Sanders, J. K. M. *Chem. Commun.* **1999**, 1575−1576.

(7) Chung, M.-K.; White, P. S.; Lee, S. J.; Gagné, M. R. *Angew. Chem., Int. Ed.* **2009**, *48*, 8683−8686.

(8) Lam, R. T. S.; Belenguer, A.; Roberts, S. L.; Naumann, C.; Jarrosson, T.; Otto, S.; Sanders, J. K. M. *Science* **2005**, *308*, 667−669.

(9) (a) Chung, M.-K.; Severin, K.; Lee, S. J.; Waters, M. L.; Gagné, M. R. *Chem. Sci.* **2011**, *2*, 744−747. (b) Ingerman, L. A.; Waters, M. L. *J. Org. Chem.* **2009**, *74*, 111−117. (c) Chung, M.-K.; Hebling, C. R.; Jorgenson, J. W.; Severin, K.; Lee, S. J.; Gagné, M. R. *J. Am. Chem. Soc.* **2008**, *130*, 11819−11827. (d) Bulos, F.; Roberts, S. L.; Furlan, R. L. E.; Sanders, J. K. M. *Chem. Commun.* **2007**, 3092−3093. (e) Nicholas, K. M.; Masaomi, M. *J. Org. Chem.* **2007**, *72*, 9308−9313. (f) Voshell, S. M.; Lee, S. J.; Gagné, M. R. *J. Am. Chem. Soc.* **2006**, *128*, 12422−12423. (g) Roberts, S. L.; Furlan, R. L. E.; Otto, S.; Sanders, J. K. M. *Org. Biomol. Chem.* **2003**, *1*, 1625−1633.

(10) (a) Robinson, J. L. *Acc. Chem. Res.* **2008**, *41*, 1278−1288. (b) Francart, C.; Wieruszeski, J. M.; Tartar, A.; Lippens, G. *J. Am. Chem. Soc.* **1996**, *118*, 7019−7027. (c) MacArthur, M. W. *J. Mol. Biol.* **1991**, *218*, 397−412. (d) Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv. Protein Chem.* **1985**, *37*, 1−109. (e) Kessier, H. *Angew. Chem., Int. Ed.* **1982**, *21*, 512−523.

(11) (a) Nishio, M. *Phys. Chem. Chem. Phys.* **2011**, *13*, 13873−13900. (b) Nishio, M.; Takahashi, O.; Kohno, Y. *Chem. Rev.* **2011**, *110*, 6049−6076. (c) Plevin, M. J.; Bryce, D. L.; Boisbouvier, J. *Nature*

Chem. **2010**, *2*, 466−471. (d) Tewari, A. K.; Dubey, R. *Bioorg. Med. Chem.* **2008**, *16*, 126−143. (e) Gil, A.; Branchadell, V.; Bertran, J.; Oliva, A. *J. Phys. Chem. B* **2007**, *111*, 9372−9379.

(12) (a) Vasilev, L., Fry, D., Eds. *Small-Molecule Inhibitors of Protein-Protein Interactions*; Current Topics in Microbiology and Immunology *348*; Springer-Verlag: Berlin, Heidelberg, 2011. (b) Scheiner, S. *Phys. Chem. Chem. Phys.* **2011**, *13*, 13860−13872. (c) Ibrahim, B. S.; Pattabhi, V. *Biochem. Biophys. Res. Commun.* **2004**, *325*, 1082−1089.

(13) Marsili, S.; Chelli, R.; Schettino, V.; Procacci, P. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2673−2685.

(14) Chung, M.-K.; Lee, S. J.; Waters, M. L.; Gagné, M. R. *J. Am. Chem. Soc.* **2012**, DOI: 10.1021/ja302347q.

(15) With a high-energy fragmentor parameter (375.0 V), LC-TOF analysis of [2]-catenanes (^{H}**3a**, ^{F}**3d**^{Ph}, ^{HO}**3c**, ^{F}**3b**, and ^{F}**4b**) showed no fragments larger than tetramer. See the SI. The characteristic fragmentation pattern of ^{H}**3a** indicating an interlocked structure was further confirmed by its MS/MS analyses: see ref 7.

(16) See details in the SI.

(17) Proline at the *i*+1 position has been known to promote *β*-turn formation. It has also been reported that D-proline in this position favors type I′ and II′ turns, while L-proline at the same position most commonly prefers type I and II, with type VIII occasionally being observed: (a) Lai, J. R.; Huck, B. R.; Weisblum, B.; Gellman, S. H. *Biochemistry* **2001**, *41*, 12835−12842. (b) Das, C.; Naganagowda, G. A.; Karle, I. L.; Balaram, P. *Biopolymers* **2001**, *58*, 335−346. (c) Haque, T. S.; Gellman, S. H. *J. Am. Chem. Soc.* **1997**, *119*, 2303−2304. (d) Santa, H.; Ylisirniö, M.; Hassinen, T.; Laatikainen, R.; Peräkylä, M. *Protein Eng.* **2002**, *15*, 651−657. For the turn types and dihedral angles in detail, see SI1 Table 1.

(18) Hydrogen-bonding interactions are well known to stabilize host−guest structure: (a) Schneider, H.-J. *Angew. Chem., Int. Ed.* **2009**, *48*, 3924−3977. (b) Pal, D.; Chakrabarti, P. *J. Mol. Biol.* **1999**, *294*, 271−288. (c) Biedermannova, L.; Riley, K. E.; Berka, K.; Hobza, P.; Vondrasek, J. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6350−6359.

(19) Thomas, K. M.; Naduthambi, D.; Zondlo, N. J. *J. Am. Chem. Soc.* **2006**, *128*, 2216−2217.

(20) (a) Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem., Int. Ed.* **2003**, *42*, 1210−1250. (b) Hunter, C. A.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1990**, *112*, 5525−5534.

(21) (a) Salonen, L. M.; Ellermann, M.; Diederich, F. *Angew. Chem., Int. Ed.* **2011**, *50*, 4808−4842. (b) Schneider, H.-J. *Angew. Chem., Int. Ed.* **2009**, *48*, 3924−3977.

(22) (a) Tsuzuki, S.; Fujii, A. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2584−2594. (b) Tsuzuki, S.; Honda, K.; Fujji, A.; Uchimaru, T.; Mikami, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2860−2865.

(23) For minimal geometric criteria for hydrogen bond identification in proteins, see: (a) McDonald, I. K.; Thornton, J. M. *J. Mol. Biol.* **1994**, *238*, 777−793. (b) Torshin, I. Y.; Weber, I. T.; Harrison, R. W. *Protein Eng.* **2002**, *15*, 359−363.

(24) L-^{F}Pro CH's and the naphthyl group participating in the CH−*π* interactions were assigned through COSY and TOCSY spectra (SI1 Figures 7 and 8).

(25) Ion−ion reactions of hydrazine mixtures can lead to ions that are an artifact of crossover: Schiltz, H.; Chung, M.-K.; Lee, S. J.; Gagné, M. R. *Org. Biomol. Chem.* **2008**, *6*, 3597−3600.

(26) (a) Park, H. S.; Kim, C.; Kang, Y. K. *Biophys. Chem.* **2003**, *105*, 89−104. (b) Antohi, O.; Sapse, A.-M. *J. Mol. Struct. (Theochem)* **1998**, *430*, 247−258. (c) Aubry, A.; Cung, M. T.; Marraud, M. *J. Am. Chem. Soc.* **1985**, *107*, 7640−7647.

(27) The standard dihedral angle comparison of the type I and VIII turns: type I, Φ_{i+1}, −64°, Ψ_{i+1}, −27°, Φ_{i+2}, −90°, and Ψ_{i+2}, −7°; type VIII, Φ_{i+1}, −72°, Ψ_{i+1}, −33°, Φ_{i+2}, −123°, and Ψ_{i+2}, 121°. See: Hutchinson, E. G.; Thornton, J. M. *Protein Sci.* **1994**, *3*, 2207−2216.

(28) (a) Lanzarotti, E.; Biekofsky, R. R.; Estrin, D. A.; Marti, M. A.; Turjanski, A. G. *J. Chem. Inf. Model.* **2011**, *51*, 1623−1633. (b) Zheng, H.; Cameforo, K.; Gao, J. *J. Am. Chem. Soc.* **2009**, *131*, 18−19. (c) Woll, M. G.; Hadley, E. B.; Mecozzi, S. J.; Gellman, S. H. *J. Am. Chem. Soc.* **2006**, *128*, 15932−15933. (d) Frank, B. S.; Vardar, D.; Buckley, D. A.; McKnight, C. J. *Protein Sci.* **2002**, *11*, 680−687. (e) Maithal, K.; Ravindra, G.; Nagaraj, G.; Singh, S. K.; Balaram, H.; Balaram, P. *Protein Eng.* **2002**, *15*, 575−584.

(29) Recent reviews: (a) Roymon, J.; Rao, C. P. *Chem. Rev.* **2011**, *11*, 4658−4702. (b) Mutihae, L.; Lee, J. H.; Kim, J. S.; Vicens, J. *Chem. Soc. Rev.* **2011**, *40*, 2777−2796. (c) Lee, J. W.; Samal, S.; Selvapalam, N.; Kim, H.-J.; Kim, K. *Acc. Chem. Res.* **2003**, *36*, 621−630. (d) Chen, G.; Jiang, M. *Chem. Soc. Rev.* **2011**, *40*, 2254−2266. (e) Wenz, G; Han, B.-H.; Müller, A. *Chem. Rev.* **2006**, *196*, 782−817. (f) Hardie, M. J. *Chem. Soc. Rev.* **2010**, *39*, 516−527. (g) Hardouin-Lerouge, M.; Hudhoume, P.; Sallé, M. *Chem. Soc. Rev.* **2011**, *40*, 30−43.