Supramolecular Chemistry

DOI: 10.1002/anie.200903478

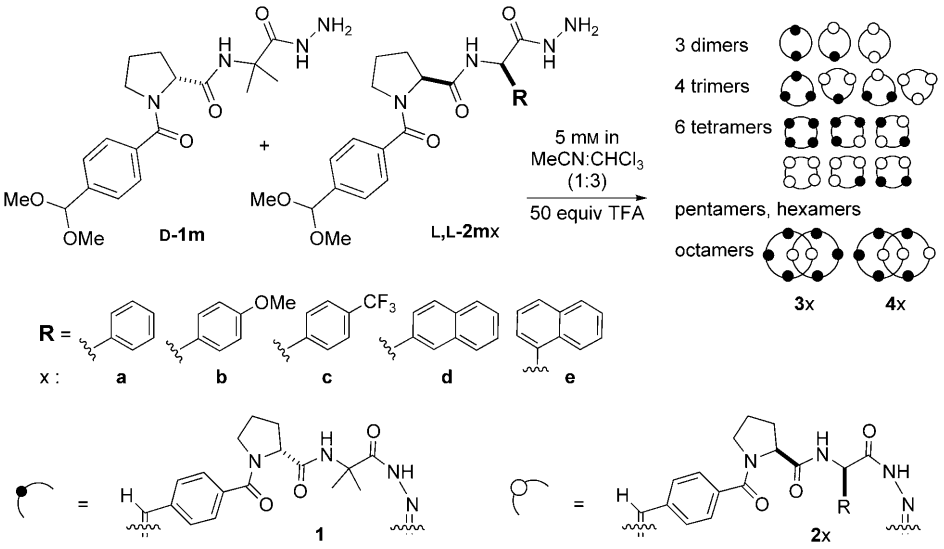
Synthesis of Interlocked 56-Membered Rings by Dynamic Self-Templating**

*Mee-Kyung Chung, Peter S. White, Stephen J. Lee, and Michel R. Gagné**

Reversible templated synthesis is a powerful tool for the synthesis of supramolecular structures that range from solids (e.g., MOFs) to complex molecular assemblies and new host–guest combinations.[1] The error-correction features of reversible synthesis have provided insights into host–guest thermodynamics while also offering unusual structural examples in which free energy is minimized. One such example was the hexameric [2]-catenane nucleated by acetylcholine from dipeptide hydrazone monomers.[2] The strong binding of acetylcholine and the small concentration of the catenane in its absence suggested a binding site at the intersection of the two interlocked macrocyclic rings. This example represents one of numerous cases wherein dynamic templating was used to discover synthetic receptors by dynamic combinatorial chemistry (DCC).[3]

We report herein the spontaneous self-assembly of new octameric [2]-catenanes under dynamic assembly conditions. The compounds are assembled of two different peptide monomers; their formation is highly diastereoselective, and an X-ray crystal structure reveals structural features that are likely repeated in previously discovered receptors composed of these monomers.[4] We additionally report that the solid-state structure is maintained in solution (as determined by NMR spectroscopy), and we postulate a driving force for this assembly.

When building block D-**1m**[5] was combined with L,L-**2ma** and trifluoroacetic acid (TFA), a mixture of cyclic oligomers was obtained (Scheme 1); the dynamic library (DL) included



Scheme 1. Generation of dynamic libraries (DL) of oligomers from pairs of hydrazide/aldehyde monomers.

cyclic dimers, trimers, tetramers, pentamers, hexamers, and two octamers (Figure 1a). LC-MS analysis revealed that the major cyclic octamer **3a** had a mass of 2723, corresponding to six units of **1** and two units of **2a**, while the minor octamer **4a** had a mass of 2771 (five units of **1** and three units of **2a**). The formation of large oligomers is unusual, as the smaller dimers and trimers are entropically favored. MS/MS analysis of **3a** showed the principal daughter ion to be the tetrameric $[\mathbf{1}_3\mathbf{2a}_1+H]^+$; no other tetrameric ions were detected (Figure 1b).[6] As no heptameric, hexameric, or pentameric fragments were detected, these data suggested an octameric [2]-catenane wherein each tetrameric macrocycle contained three units of **1** and a single unit of **2a**. Monomers having *i*Pr, Bn, or CH_2-indole (Trp) R groups in L,L-**2mx** did not generate the corresponding catenanes.

Since we employed two building blocks of different chirality, we investigated the effect of different stereoisomers on [2]-catenane formation. With the exception of the mirror-image combination (L-**1m** and D,D-**2ma**), all combinations failed to generate any octameric [2]-catenanes.[7] The high diastereoselectivity and strong 3:1 monomer preference

[*] Dr. M.-K. Chung, Dr. P. S. White, Prof. Dr. M. R. Gagné
Department of Chemistry
University of North Carolina at Chapel Hill
Chapel Hill, NC 27599-3290 (USA)
Fax: (+1) 919-962-2388
E-mail: mgagne@unc.edu
Dr. S. J. Lee
US Army Research Office
P.O. Box 12211, Research Triangle Park, NC 27709 (USA)

[**] We thank the Defense Threat Reduction Agency (DTRA) for support (HDTRA1-08-1-0045). We thank Dr. Marc ter Horst for helpful discussions of NMR spectroscopy techniques, and Prof. Marcey Waters (UNC Chapel Hill).

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200903478.

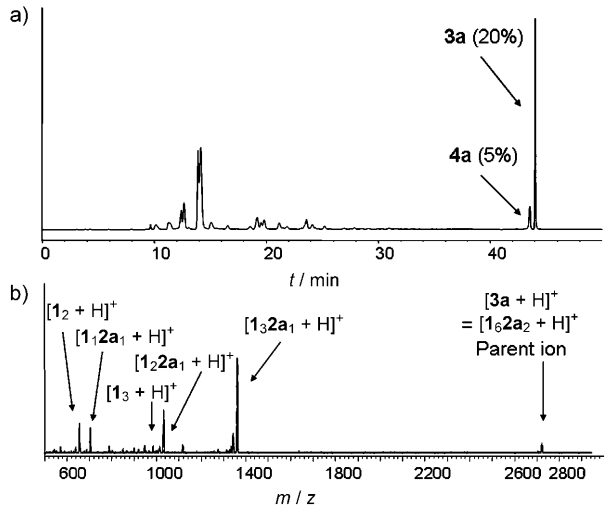


Figure 1. a) HPLC–UV trace (289 nm) at day seven of an untemplated DL from D-**1m** and L,L-**2ma** (1:1, 5 mM in $MeCN/CHCl_3$ = 1:3, 50 equiv TFA); b) Collision-induced dissociation (CID) MS/MS spectrum of **3a** at a collision energy of 140 V.

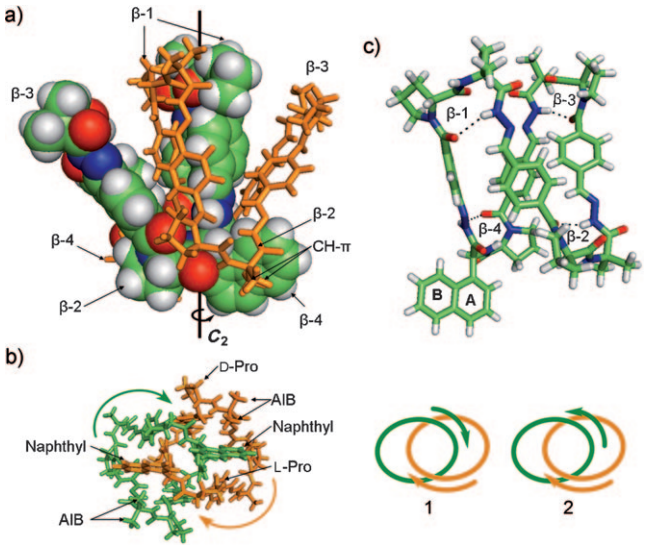


Figure 2. a) The X-ray structure of **3e** displaying two identical interlocked tetramers. One ring is shown as a space-filling representation, while the symmetry-related ring is shown in a stick representation. b) View from the bottom of (a) along the C_2 axis. Two possible enchained diastereomers are illustrated in (1) and (2). Octamer **3e** adopts the diastereomeric form (1). c) A single tetrameric ring of **3e** displaying four β-turns (β-1 to β-4).

suggested a tightly ordered structure that was exceptionally sensitive to component stereochemistry. We also examined L,L-**2m** monomers with variable aryl substituents to study the effect of electronic and steric changes on [2]-catenane preference. Electron-withdrawing or -donating groups in the para position and the 2-naphthyl substituent did not significantly affect the [2]-catenane equilibrium concentration.[8] However, when a 1-naphthyl group was employed (L,L-**2me**), larger quantities of **3e** (29%) and **4e** (19%) were generated. MS/MS analysis of **4e** showed two tetrameric daughter ions ($[\mathbf{1_3 2e}_1 + H]^+$ and $[\mathbf{1_2 2e}_2 + H]^+$), three trimeric ions, and only two dimeric daughter ions. The fact that $[\mathbf{2e}_2 + H]^+$ was not an observed dimeric daughter ion implied that the [2]-catenane contained one [**1·1·1·2e**] ring and one symmetric tetramer with alternating **1** and **2e** units, that is, [**1·2e·1·2e**].

In $CHCl_3$, this same reaction proved to be more selective for **3e**, and an even higher selectivity was observed with a biased DL (D-**1m**/L,L-**2me** = 3:1). Under these conditions, **3e** could be isolated in 64% yield after purification by flash column chromatography.

Fortunately, **3e** could be obtained as single crystals, and high quality X-ray diffraction data were obtained.[9] As expected, the octameric [2]-catenane consisted of two identical interlocked tetramers related by a C_2 axis penetrating a tightly packed central loop, which also relates two floppier loops (Figure 2a). Each tetrameric ring is saddle-shaped with three β turns created by the D-Pro-AIB unit (AIB = 2-aminoisobutyric acid) and a fourth formed by the L-Pro-L-arylGly moiety; each β turn is connected by a rigid, extended 1,4-hydrazone/benzamide linkage (Figure 2c). The unique **2e** unit is located in a dense lower core with the naphthyl ring sandwiched between the β-2 and β-4 turns of the other macrocycle. A Φ,Ψ analysis indicates three type II′ β turns (β-1, β-2, β-3) for the **1** unit and a type VIII turn for the **2e** unit (β-4).[10]

CH–π[8,11] and hydrogen-bonding interactions are well known to stabilize host–guest structures.[12] As an aryl or naphthyl group in **2**x was required to form the catenanes, it can reasonably be inferred that the intercalation of the naphthyl group between a proline and an AIB residue serves a stabilization function. Metrical parameters suggests good geometries to accommodate CH–π interactions between the naphthyl group and the proline residue (3.9–4.1 Å) and the naphthyl group and one AIB methyl group (ca. 3.5 Å).[13] The AIB methyl group preferentially sits over the naphthyl B ring (Figure 1a), while the proline CH units interact with both. Edge-to-face π-stacking interactions are also observed in the intertwined central loop of the structure. The crystal structure additionally revealed that of the eight NH groups, six hydrogen-bond to an amide C=O unit, four of which help to rigidify a β turn and two of which are intermacrocyclic; the remaining two hydrogen atoms point towards solvent (Figure 4a). Finally, like the Sanders hexameric [2]-catenane receptor, **3e** is observed as a single enchained diastereomer (Figure 2b, (1)).

Unlike the Sanders hexameric [2]-catenane receptor, which gives broad line spectra in the absence of acetylcholine, **3e** provides sharp spectra with well-dispersed signals; [D_5]pyridine as solvent gave especially high quality spectra (Supporting Information, SI Figure 13). Particularly diagnostic were upfield-shifted resonances for the single AIB methyl group and proline CH units participating in the CH–π interactions detected in the X-ray analysis. These close contacts were also apparent in the ROESY spectrum (Figure 3). One AIB methyl group had cross peaks to the naphthyl B protons, while the upfield proline hydrogen atoms are close to both A and B hydrogens (Figure 3). Close contact

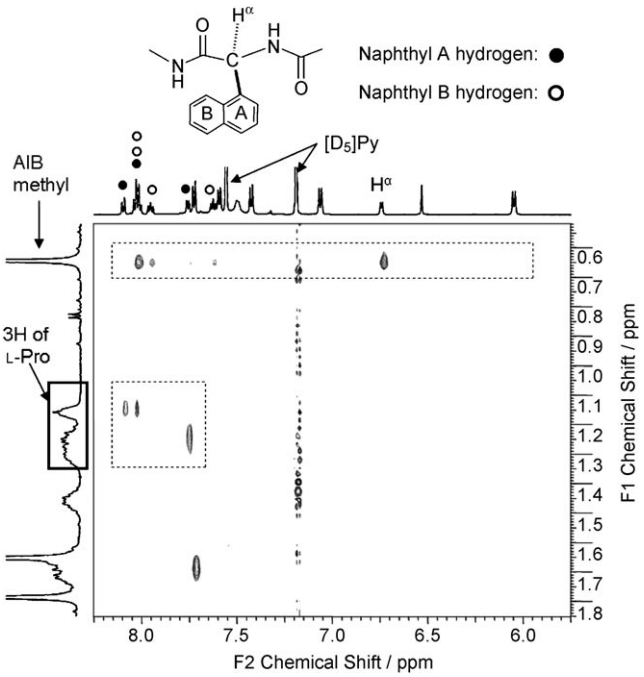


Figure 3. A portion of the 2D 599.8 MHz ROESY spectrum of **3e** (F1 $\delta = 0.5$–1.85 ppm; F2 $\delta = 5.75$–8.25 ppm; [D_5]pyridine, 20 °C).

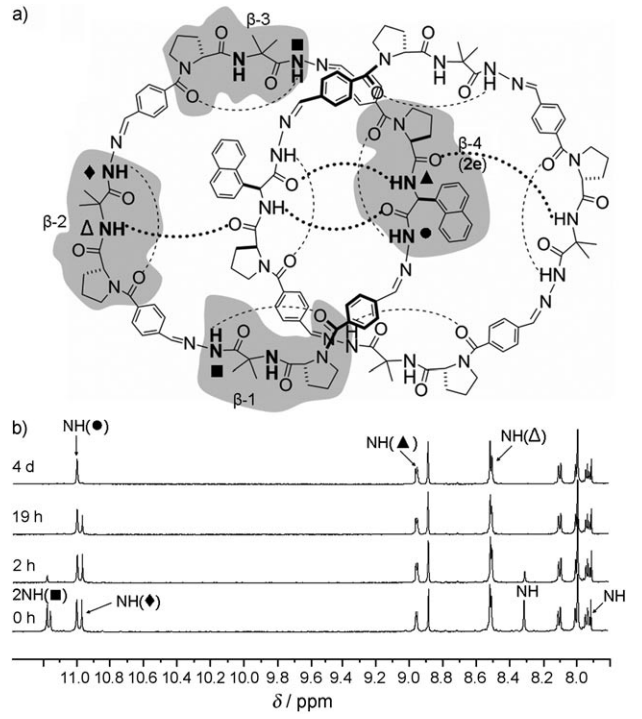


Figure 4. a) Illustration of the **3e** structure; dotted lines refer to hydrogen bonds noted in the X-ray structure. b) A portion ($\delta = 7.8$–11.4 ppm) of the 599.8 MHz ^{1}H NMR spectra of **3e** ([D_4]MeOH, 20 °C), recorded at different times from mixing.

between the upfield AIB methyl group and H^α was consistent with a C–C distance of 3.0 Å in the X-ray structure. In short, the observed through-space interactions are completely consistent with the solid-state structure being maintained in solution.

In the downfield portion of the ^{1}H NMR spectra, four hydrazone NH resonances (♦, ●, 2■) were observed from $\delta =$ 10.7–11.3 ppm (Figure 4 and Supporting Information, SI Figure 13). Their assignments were deduced from cross peaks in the ROESY spectrum (Supporting Information, SI Figure 20). The remaining amide NH groups could be differentiated from imine CH resonances by the lack of correlations to the carbon atoms in the HSQC spectrum (Supporting Information, SI Figures 18 and 19). The amide NH group (▲) in the β-4 turn was uniquely coupled to the adjacent benzylic CH group (COSY; Supporting Information, SI Figures 14 and 15). The amide NH group (△) in the β-2 turn was also deduced by cross-peak analysis of the ROESY spectrum (Supporting Information, SI Figures 21 to 23).

The downfield hydrazone resonances each nucleate their respective β turns by hydrogen bonding, while the amide NH groups in β-1 and β-3 apparently orient towards solvent. Although no direct NMR spectroscopic evidence for hydrogen bonding was apparent for the amide NH groups in β-2 and β-4 (▲,△), the X-ray structure showed these NHs to participate in intermacrocyclic hydrogen bonding.

H/D exchange reactions were informative. In [D_4]MeOH, two of four hydrazone NH groups exchanged quickly (in less than 2 h), while a third (♦) was gone after four days; a fourth NH group (●) was untouched (Figure 4). The two amide NH groups in the β-1 and β-3 turns were also found to exchange. When this experiment was repeated in [D_5]pyridine/D_2O, all NH groups exchanged except for the two amide NH groups (▲ and △) in β-2 and β-4 ($\delta = 8.98$ and 8.51 ppm, respectively). In the X-ray crystal structure, it is these two turns which sandwich the naphthyl group.

The exchange experiments are thus consistent with a densely packed, relatively inaccessible region that is composed of the β-2 and β-4 turns. Additional floppier elements are also apparent, with β-1 and β-3 being more prone to solvation and H/D exchange. Cooperative hydrogen bonding and CH–π interactions thus serve to rigidify and stabilize a core of the structure that is efficiently nucleated by the arene ring.

The structural results described herein articulate a unique solution to the minimization of free energy in a mixed monomer system. The three-dimensional assembly interlocks two 56-membered macrocycles and is perfectly diastereoselective with respect to the enchainment stereoisomer. It has a significant preference for a 3:1 ratio of macrocyclic ring components, and it is highly selective for a specific combination of diastereomeric monomers.[14] X-ray structural investigations reveal a dense self-complementary structure that efficiently utilizes functionality arranged by the Pro-AIB templated β turns to grasp the complementary structure. Liberal use of intra- and intermacrocyclic hydrogen bonds, π stacking, and CH–π interactions conspire to stabilize an entropically disfavored structure that persists in solution. This analysis serves as an excellent structural model for cases in which these subunits assemble into efficient hosts for the binding of guests.

Experimental Section

Generation of DLs for LC-MS analyses: 5 mM DLs were prepared on a 1 mL scale. D-**1m** (5 μmol) and L,L-**2mx** (5 μmol) were separately dissolved in 1 mL of a mixture of MeCN and $CHCl_3$ (1:3). 0.5 mL of the D-**1m** solution was mixed with 0.5 mL of the L,L-**2mx** solution and TFA (50 equiv, 250 μmol, 19 μL). The resultant solution was allowed to sit for several days until a steady state was reached (as determined by LC-MS analyses).

Generation and isolation of **3e**: Three 5 mM biased DLs (D-**1m**/L,L-**2me** = 3:1) were prepared on a 20 mL scale. Each DL was prepared by dissolving the mixture of D-**1m** (29.4 mg, 75.0 μmol) and L,L-**2me** (12.3 mg, 25.0 μmol) in $CHCl_3$ (20.0 mL). DLs were allowed to sit for four days, and then Et_3N (50 equiv, 5000 μmol, 0.7 mL) was added to each DL. To remove the precipitates, the solutions were filtered and all three were combined. The volatiles were removed in vacuo, and $CHCl_3$ (60 mL) was added. This solution was washed with water (3 × 60 mL) and then dried over anhydrous $MgSO_4$. More $CHCl_3$ (60 mL) was added, and the mixture was heated to 50 °C (to alleviate the poor solubility of **3e** in $CHCl_3$) and then filtered to remove $MgSO_4$. Volatile components were removed in vacuo, and pure **3e** was obtained in 64.0% (67.8 mg) yield after flash column chromatography on silica gel (8% MeOH/$CHCl_3$). Recrystallization of **3e** from either pyridine by slow evaporation at RT or pyridine/water at low temperature yielded crystals suitable for X-ray analysis.

Received: June 26, 2009
Published online: October 8, 2009

Keywords: molecular recognition · self-assembly · supramolecular chemistry · template synthesis

[1] For several representative reviews of this broad topic, see: a) P. C. Haussmann, J. F. Stoddart, *Chem. Rec.* **2009**, *9*, 219 – 226; b) K. Ariga, J. P. Hill, Y. Wakayama, *Phys. Status Solidi A* **2008**, *205*, 1249 – 1257; c) J.-M. Lehn, *Chem. Soc. Rev.* **2007**, *36*, 151 – 160; d) J. R. Nitschke, *Acc. Chem. Res.* **2007**, *40*, 103 – 112; e) S. J. Rowan, S. J. Cantrill, G. R. L. Cousins, J. K. M. Sanders, J. F. Stoddart, *Angew. Chem.* **2002**, *114*, 938 – 993; *Angew. Chem. Int. Ed.* **2002**, *41*, 898 – 952.

[2] R. T. S. Lam, A. Belenguer, S. L. Roberts, C. Naumann, T. Jarrosson, S. Otto, J. K. M. Sanders, *Science* **2005**, *308*, 667 – 669. Other examples include: a) K. R. West, R. F. Ludlow, P. T. Corbett, P. Besenius, F. M. Mansfeld, P. A. G. Cormack, D. C. Sherrington, J. M. Goodman, M. C. A. Stuart, S. Otto, *J. Am. Chem. Soc.* **2008**, *130*, 10834 – 10835; b) H. Y. Au-Yeung, G. D. Pantoş, J. K. M. Sanders, *Proc. Natl. Acad. Sci. USA* **2009**, 106, 10466 – 10470.

[3] For reviews, see: a) S. Ladame, *Org. Biomol. Chem.* **2008**, *6*, 219 – 226; b) R. F. Ludlow, S. Otto, *Chem. Soc. Rev.* **2008**, *37*, 101 – 108; c) J.-M. Lehn, *Chem. Soc. Rev.* **2007**, *36*, 151 – 160; d) P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders, S. Otto, *Chem. Rev.* **2006**, *106*, 3652 – 3711.

[4] a) H. Schiltz, M.-K. Chung, S. J. Lee, M. R. Gagné, *Org. Biomol. Chem.* **2008**, *6*, 3597 – 3600; b) M.-K. Chung, C. M. Hebling, J. W. Jorgenson, K. Severin, S. J. Lee, M. R. Gagné, *J. Am. Chem. Soc.* **2008**, *130*, 11819 – 11827; c) J. Liu, K. R. West, C. R. Bondy, J. K. M. Sanders, *Org. Biomol. Chem.* **2007**, *5*, 778 – 786; d) S. M. Voshell, S. J. Lee, M. R. Gagné, *J. Am. Chem. Soc.* **2006**, *128*, 12422 – 12423; e) S. L. Roberts, R. L. E. Furlan, S. Otto, J. K. M. Sanders, *Org. Biomol. Chem.* **2003**, *1*, 1625 – 1633; f) R. L. E. Furlan, Y.-F. Ng, S. Otto, J. K. M. Sanders, *J. Am. Chem. Soc.* **2001**, *123*, 8876 – 8877; g) G. R. L. Cousins, S.-A. Poulsen, J. K. M. Sanders, *Chem. Commun.* **1999**, 1575 – 1576.

[5] The **m** in the compound numbering refers to monomer, which is slightly different from the repeating unit in the cyclic compounds.

[6] LC-MS analyses also showed that there were no larger fragments than tetramers. See the Supporting Information (SI Figure 2).

[7] The combination of D-**1m**/L,D-**2ma**, D-**1m**/D,L-**2ma**, D-**1m**/D,D-**2ma**, L-**1m**/L,D-**2ma**, L-**1m**/L,L-**2ma**, or L-**1m**/L,D-**2ma** failed to generate the octamer. See details in the Supporting Information, SI Figure 3.

[8] K. M. Thomas, D. Naduthambi, N. J. Zondlo, *J. Am. Chem. Soc.* **2006**, *128*, 2216 – 2217.

[9] Crystal data for **3e**: $C_{88.5}H_{121.5}N_{18.5}O_{25.5} = C_{76}H_{82}N_{16}O_{12}\cdot 2.5(C_5H_5N)\cdot 13.5(H_2O)$, $M_r = 1852.54$, crystal dimensions 0.28 × 0.37 × 0.44 mm, orthorhombic, space group $P2_12_12$, cell parameters $a = 26.5088(8)$, $b = 18.1541(5)$, $c = 20.0246(5)$ Å, $V = 9636.7(5)$ Å, $T = 100(2)$ K, $Z = 4$, $\rho_{calcd} = 1.277\ Mg\,cm^{-3}$, $\mu = 0.790\ mm^{-1}$, radiation $Cu_{K\alpha}$, $\lambda = 1.54178$ Å. 129815 reflections were collected to a max θ angle of 70.48° (0.82 Å resolution), of which 18203 were independent (average redundancy 7.132, completeness = 99.3%, $R_{int} = 0.0346$, $R_{sig} = 0.0208$) and 16792 (92.25%) were greater than $2\sigma(F^2)$. Data were corrected for absorption effects using the multiscan method (SADABS). Min/max apparent transmission = 0.886, min transmission coefficient = 0.7419, max transmission coefficient = 0.8247. The structure was solved by direct methods (SHELXL-97) and refined by full-matrix least-squares methods on F^2 with 1214 parameters. $R1 = 0.0946$ ($I > 2\sigma(I)$), $wR2 = 0.2736$ for 16792 data and $R1 = 0.0994$, $wR2 = 0.2786$ for all data, GOF = 1.082; max/min residual density 1.043/−0.509 e Å^3 with an RMS deviation 0.096 e Å^3. CCDC 736965 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif.

[10] Proline at the $i+1$ position has long 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 prefers type I and II as the common turns and type III as the most common nonclassic β turn: a) J. R. Lai, B. R. Huck, B. Weisblum, S. H. Gellman, *Biochemistry* **2001**, *41*, 12835 – 12842; b) C. Das, G. A. Naganagowda, I. L. Karle, P. Balaram, *Biopolymers* **2001**, *58*, 335 – 346; c) T. S. Haque, S. H. Gellman, *J. Am. Chem. Soc.* **1997**, *119*, 2303 – 2304; d) H. Santa, M. Ylisirniö, T. Hassinen, R. Laatikainen, M. Peräkylä, *Protein Eng.* **2002**, *15*, 651 – 657. For the dihedral angles, see the Supporting Information, SI Figure 9.

[11] E. A. Meyer, R. K. Castellano, F. Diederich, *Angew. Chem.* **2003**, *115*, 1244 – 1287; *Angew. Chem. Int. Ed.* **2003**, *42*, 1210 – 1250.

[12] a) H.-J. Schneider, *Angew. Chem.* **2009**, *121*, 3982 – 4036; *Angew. Chem. Int. Ed.* **2009**, *48*, 3924 – 3977; b) D. Pal, P. Chakrabarti, *J. Mol. Biol.* **1999**, *294*, 271 – 288; c) L. Biedermannova, K. E. Riley, K. Berka, P. Hobza, J. Vondrasek, *Phys. Chem. Chem. Phys.* **2008**, *10*, 6350 – 6359.

[13] The distances were measured between the carbon atom of an AIB methyl group and three imaginary centroids of the naphthyl unit or between a centroid of the proline unit and the naphthyl centroids (see the Supporting Information SI Figure 10). They all were regarded as CH–π interactions on the basis of the distance cutoff ($d(X\cdots M) < 4.3$ Å). See a) T. Steiner, G. Koellner, *J. Mol. Biol.* **2001**, *305*, 535 – 557; b) M. Brandl, M. S. Weiss, A. Jabs, J. Sühnel, R. Hilgenfeld, *J. Mol. Biol.* **2001**, *307*, 357 – 377.

[14] The exact reason for favoring the D-Pro-AIB unit at the other three positions is not precisely known, although we note that the L-Pro-L-naphthylGly unit is partially tolerated in β-2 position of **4e**.

