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

Efficient Self-Assembly in Water of Long Noncovalent Polymers by Nucleobase Analogues

Brian J. Cafferty,[†] Isaac Gállego,[†] Michael C. Chen,[†] Katherine I. Farley,[†] Ramon Eritja,[‡] and Nicholas V. Hud*,[†]

[†]Department of Chemistry and Biochemistry, Parker H. Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, United States

[‡]Institute for Research in Biomedicine, Parc Científic de Barcelona, Barcelona 08028, Spain

Supporting Information

ABSTRACT: Molecular self-assembly is widely appreciated to result from a delicate balance between several noncovalent interactions and solvation effects. However, current design approaches for achieving self-assembly in water with small, synthetic molecules do not consider all aspects of the hydrophobic effect, in particular the requirement of surface areas greater than 1 nm^2 for an appreciable free energy of hydration. With the concept of a minimum hydrophobic surface area in mind, we designed a system that achieves highly cooperative self-assembly in water. Two weakly interacting low-molecular-weight monomers (cyanuric acid and a modified triaminopyrimidine) are shown to form extremely long supramolecular polymer assemblies that retain water solubility. The complete absence of intermediate assemblies means that the observed equilibrium is between free monomers and supramolecular assemblies. These observations are in excellent agreement with literature values for the free energy of nucleic acid base interactions as well as the calculated free energy penalty for the exposure of hydrophobic structures in water. The results of our study have implications for the design of new self-assembling structures and hydrogel-forming molecules and may provide insights into the origin of the first RNA-like polymers.

Elucidating the physiochemical principles that govern molecular self-assembly in water is essential for understanding biological systems at the molecular level, for developing new materials, and for nanotechnology.[1,2] Current efforts to improve the self-assembly of synthetic molecules in water include increasing the number and arrangement of H-bonds,[3] adding complementary electrostatic charges for salt-bridge formation,[4] and covalently joining multiple recognition units with a flexible linker.[5,6] Hierarchical assembly methods involving intermediate structures with large planar hydrophobic surfaces that stack to form noncovalent polymers have been demonstrated using both synthetic[7−9] and biologically derived molecules.[10] However, a simple description of what is needed to harness the hydrophobic effect is still lacking, which limits its use as a design parameter for artificial systems. Both theoretical studies and recent experimental results indicate that hydrophobic surfaces must be larger than a minimum size (ca. 1 nm^2) in order for the hydrophobic effect to facilitate molecular self-assembly.[11−14] Inspired by such results, we hypothesized that small, water-soluble molecules that associate weakly in water would form highly ordered supramolecular assemblies if the intermediate structures formed by the association of monomers were to create hydrophobic surfaces with dimensions greater than 1 nm^2. Here we show the application of this principle to the generation of some of the longest supramolecular polymers ever observed in water from well-studied recognition elements that were previously limited to linear polymer formation in aprotic solvents.

For over 20 years, the triazines melamine and cyanuric acid (**CA**) and the related pyrimidines 2,4,6-triaminopyrimidine (**TAP**) and barbituric acid have been used as recognition units to construct many different three-dimensional noncovalent assemblies, including the formation of six-membered rosettes.[15,16] Early studies with these motifs emphasized the importance of H-bonding in supramolecular assemblies and were accordingly carried out in "noncompeting" aprotic solvents. Later attempts to assemble the complementary monomeric molecules in water resulted in the formation of irregular precipitates[17] or the total loss of molecular recognition[7] unless multiple units were joined by a hydrophilic covalent linker.[6] We were intrigued by the potential of these historically significant recognition units because (1) the surface area of a six-membered rosette (e.g., $\mathbf{TAP_3CA_3}$) is 1.7 nm^2 and (2) the similarity of these monomer units to nucleic acid bases presents the possibility of understanding the energetics of assembled structures in terms of well-characterized molecular interactions. Specifically, decades of research into RNA and DNA structures have produced some of the most accurate values available for the standard-state free energies of H-bond formation and $\pi-\pi$ stacking in water.

We constructed a two-component self-assembling system consisting of **CA** and **TAPAS**, a derivative of **TAP** in which one exocyclic amine is modified with a succinate group (Figure 1A). This modification was performed to inhibit sterically the formation of nonrosette assemblies (e.g., ribbons and sheets), which otherwise would result in coprecipitation of the unmodified **TAP** and **CA** motifs, and to increase the solubility of the desired **TAPAS−CA** rosette assembly. Equimolar

Received: December 12, 2012
Published: February 8, 2013

dx.doi.org/10.1021/ja312155v | J. Am. Chem. Soc. 2013, 135, 2447−2450

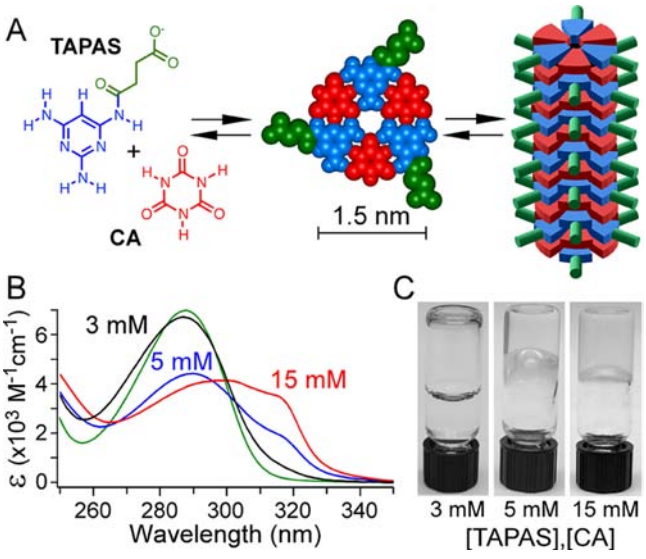

Figure 1. (A) Chemical structures of succinate-conjugated 2,4,6-triaminopyrimidine (**TAPAS**) and cyanuric acid (**CA**), a **TAPAS**−**CA** rosette, and the proposed polymer formed by stacking of rosettes. (B) UV spectra of solutions with only **TAPAS** (green curve) and 1:1 **TAPAS**/**CA** mixtures at 3, 5, and 15 mM. All of the spectra were taken at 20 °C. **CA** does not absorb light within the UV range displayed. (C) Inverted-bottle test of gel formation by solutions with the same monomer concentrations as in (B). The increasing intensity of the band at 320 nm in (B) is indicative of ring stacking and correlates with the gel rigidity in (C). See the SI for more experimental details.

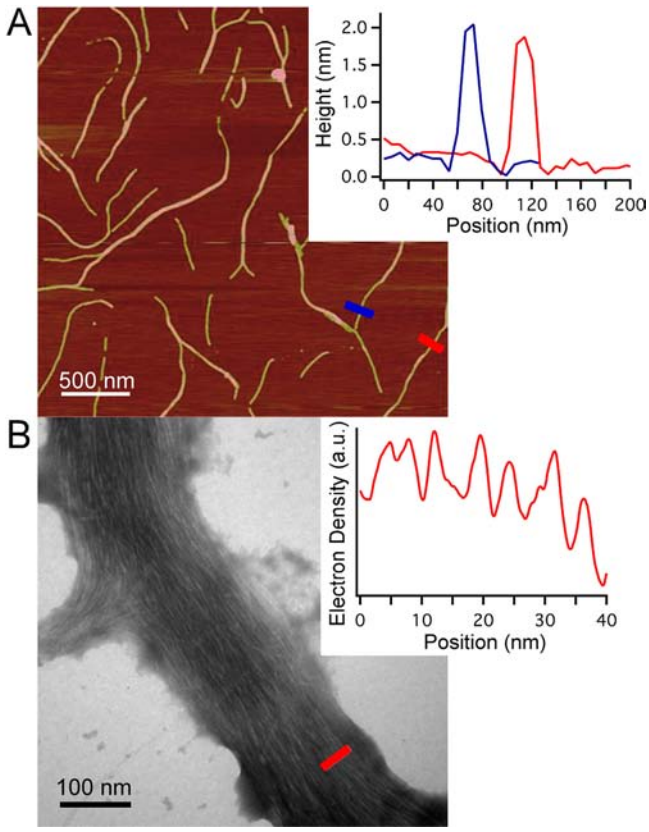

Figure 2. Images and measurements of **TAPAS**−**CA** supramolecular assemblies. (A) AFM topographic image and (B) TEM micrograph of self-assembled **TAPAS**−**CA** fibers. The insets show the profiles delimited by the red and blue lines in the main panels. The concentration of each monomer used for the assemblies was (A) 5 and (B) 10 mM.

mixtures of **TAPAS** and **CA** (pH 7, 200 mM sodium phosphate) remain in solution at concentrations above 50 mM each, whereas mixtures of unmodified **TAP** and **CA** immediately precipitate at concentrations above 10 mM. The UV absorption spectrum of **TAPAS** changes substantially in the presence of **CA**, with the rise of a longer-wavelength absorption band at concentrations above 3.5 mM [Figure 1B; also see the Supporting Information (SI)], a spectral feature associated with the ordered stacking of chromophores into J-type aggregates.[18] Mixtures containing **TAPAS** and **CA** at 5 mM or higher form shear-thinning hydrogels (Figure 1C) for which the rates of gel formation and gel rigidification increase with monomer concentration (e.g., solutions containing each monomer at 5 mM gel within 10 min, while 30 mM solutions gel in seconds).

The architecture of the **TAPAS**−**CA** assemblies was visualized by atomic force microscopy (AFM) and transmission electron microscopy (TEM). Both techniques revealed large (>1 μm) linear and branched fibrillar structures with extremely high aspect ratios (Figure 2). AFM showed that the diameter of a single fiber perpendicular to the image plane (i.e., the height) was between 1.5 and 2.2 nm, corresponding to the predicted width of a rosette. Large aggregated, fibrillar structures created higher-order network assemblies that were also observed in some of the AFM images (Figure S2 in the SI), consistent with the formation of a gel matrix. TEM micrographs also showed the formation of long individual fibers that were laterally associated into bifurcating bundles. An electron density profile of one such bundle confirmed the width of an individual fiber to be ca. 2 nm (Figure 2B).

In a previous study, Fenniri and co-workers demonstrated the formation of micrometer-length polymer bundles from monomers with similar H-bonding edges.[8] However, the stacking surface of their monomers was larger (i.e., bicyclic), and the formation of polymers longer than 100 nm required covalent tethering of two monomers. Thus, to the best of our knowledge, the **TAPAS**−**CA** assemblies are the longest supramolecular polymers generated to date by untethered, monocyclic monomers in water.

NMR spectroscopy was used to follow the association of **TAPAS** and **CA** as a function of concentration to determine their association constant. The ^{1}H spectrum of equimolar mixtures of **TAPAS** and **CA** from 1 to 3 mM exhibited **TAPAS** chemical shifts identical to those in the spectrum of **TAPAS** in the absence of **CA**, with resonance intensities directly proportional to the **TAPAS** concentration (Figure 3A). At concentrations above 3.5 mM, the integrated intensity of the **TAPAS** resonances remained identical to that at 3.5 mM. The lack of an observed change in the **TAPAS** chemical shifts or signal intensity or the appearance of new **TAPAS** resonances indicates that **TAPAS** molecules were in slow exchange between their unbound and assembled states (i.e., exchange time constant >100 ms). In these experiments, the **TAPAS**−**CA** assemblies were too large to be observed by solution NMR spectroscopy because of extreme resonance line broadening, consistent with the structures having very large molecular weight, such as those observed by AFM and TEM. The formation of a **TAPAS**−**CA** assembly is therefore a highly cooperative transition in which only free monomers and large assemblies coexist (i.e., there are no intermediate structures).[19] Additional evidence of a phase transition was also provided by

dx.doi.org/10.1021/ja312155v | J. Am. Chem. Soc. 2013, 135, 2447−2450

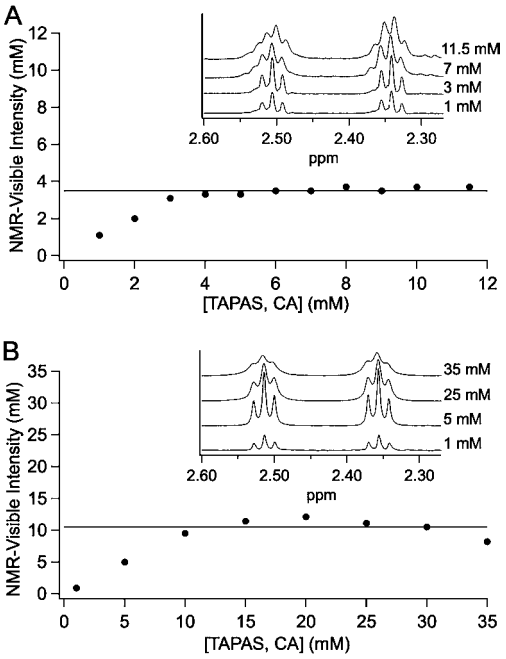

Figure 3. Plots of the apparent solution-phase concentration of **TAPAS** (as determined by methylene ^{1}H NMR resonance integration) vs the actual **TAPAS** concentration in 1:1 solutions with **CA** at (A) pH 7 and (B) pH 8. Horizontal lines indicate the concentration observed for supramolecular assembly (ca. 3.5 mM at pH 7 and ca. 10 mM at pH 8). The insets show representative spectra of the **TAPAS** methylene protons. Line broadening in the more concentrated samples likely occurred because the solvent became viscous with gel formation. The peculiar methylene resonance line structure observed for the pH 7 spectra at higher concentrations appears to be the result of residual dipolar coupling, which can be considerable in samples containing linear supramolecular assemblies that are partially aligned by the magnetic field.[20]

dynamic light scattering (DLS), which revealed a dramatic increase in scattering intensity when the concentration of each monomer was above 3.5 mM (Figure S3A). Cooling to 5 °C decreased the minimal assembly concentration 5-fold, to below 1 mM (Figure S4).

Our NMR results show similarities and differences in comparison with those for previously reported self-assembling systems. For example, previous studies of small-molecule systems that self-assemble in water through the stacking of rosettes[7] or G-tetrads (see below)[21] were observed directly by solution NMR spectroscopy, indicating the formation of supramolecular polymers much smaller than those reported here. In a different study, Whitesides and co-workers also observed a plateau in the ^{1}H NMR signal intensity associated with the self-assembly of bis-modified melamine and bis-modified **CA** monomers in an organic solvent, but in their system, signal loss occurred over several hours, indicating much slower growth of large structures compared with that reported here for monomeric recognition units.[22]

The free energy associated with **TAPAS** and **CA** monomer incorporation into supramolecular assemblies can be estimated from the free energy of dilution corresponding to the free **TAPAS** and **CA** concentrations that coexist with these assemblies. At 20 °C and pH 7, the 3.5 mM free monomer concentration corresponds to a standard-state free energy of dilution of −3.3 kcal·mol^{-1}. This value is in excellent agreement with free energy values previously determined for interactions between nucleobases within folded RNA structures. Specifically, each H-bond contributes ca. −0.5 kcal·mol^{-1} to the stability of an RNA secondary structure,[23] and the stacking of two six-membered rings (derived from the overlap of purine bases on opposite strands of a duplex) contributes −1.8 kcal·mol^{-1} to duplex stability.[24,25] On the basis of these values, the free energy for the three H-bonds and $\pi-\pi$ stacking of each monomer in stacked **TAPAS−CA** rosettes is predicted to be −3.3 kcal·mol^{-1}.

At pH 7, free **CA** (pK_a = 6.9) and the pyrimidine ring of free **TAPAS** (pK_a = 6.5; see the SI) both exist in their neutral and ionized states in approximately in equal proportions. The assembly of a supramolecular structure by **TAPAS** and **CA** would be expected to maintain a net zero charge within the stacked rosettes, as otherwise the electrostatic repulsion in the low-dielectric interior of this structure would prevent its formation. In a solution with pH above the pK_a of **CA** (i.e., >6.9), three protons per rosette must be abstracted from the solvent upon formation of the supramolecular assembly. As a test of this prediction, the concentration dependence of the **TAPAS−CA** assembly was investigated by ^{1}H NMR spectroscopy at pH 8 (Figure 3B). In this case, the plateau in the integrated resonance intensity was observed at ca. 10 mM **TAPAS** and **CA**. This 3-fold increase in the concentration required for **TAPAS−CA** assembly is in excellent agreement with the free energy difference for abstraction of a proton at pH 7 versus pH 8. We estimate the free energy of monomer incorporation into the polymeric assemblies at pH 8 to be −2.7 kcal·mol^{-1} on the basis of the free **TAPAS** and **CA** concentrations of 10 mM. As three protons per rosette must be captured from solution at a free energy that is 1.34 kcal·mol^{-1} greater at pH 8 than at pH 7, this average change in free energy of 0.67 kcal·mol^{-1} per monomer is in excellent agreement with the apparent change in the **TAPAS**/**CA** association free energy from −3.3 kcal·mol^{-1} at pH 7 to ca. −2.7 kcal·mol^{-1} at pH 8.

While the assembly of the monomers is an enthalpically driven process, as indicated by thermal denaturation experiments (Figure S4), the very long assemblies observed by AFM and TEM as well as the absence of evidence for intermediate structures (e.g., individual rosettes, rosette dimers, trimers, etc.) in our ^{1}H NMR spectra can be explained by the hydrophobic effect. Theoretical studies estimate the free energy for cavity formation in water around a hydrophobic molecule or assembly to be ca. 8 kcal·nm^{-2}·mol^{-1} (calculated for a sphere of radius 0.8 nm).[13] On the basis of this value, we estimate the free energy cost of exposing the two hydrophobic faces of a single rosette (each face with surface area of 1.7 nm^2) to be on the order of 27 kcal·mol^{-1}. The same positive free energy would apply to the two solvent-exposed ends of a rosette stack and thus may be the origin of the highly cooperative nature of **TAPAS−CA** self-assembly. That is, the free energy associated with doubling the average stack length (i.e., combining pairs of existing stacks to eliminate half of the total number of exposed rosette faces) is on the order of −27 kcal·mol^{-1}. In contrast, the free energy of mixing associated with decreasing the concentration of the stacks by a factor of 2 is only +0.3 kcal·mol^{-1}. In view of the large net free energy for stack consolidation, the hydrophobic effect is expected to drive the assembly of rosettes into polymers with lengths that would be limited only by cyclization and the steady-state breakage rates that result from kinetic fluctuations, a prediction that is fully

dx.doi.org/10.1021/ja312155v | J. Am. Chem. Soc. 2013, 135, 2447−2450

consistent with the observed phase transition between free monomers and large supramolecular assemblies.

The results presented here illustrate how relatively simple small molecules can efficiently form supramolecular polymers in water if they assemble into intermediate structures with hydrophobic surfaces having areas greater than ca. 1 nm^2. The H-bonding and $\pi-\pi$ stacking of **TAPAS** and **CA** monomers within our system is similar to the association of nucleobases within RNA and DNA duplexes. However, these associations clearly differ in that the canonical nucleobases (i.e., G, A, C, U, and T) and their corresponding free mononucleosides *do not* form H-bonded Watson−Crick base pairs in water.[26] A notable exception is the G-tetrad formed by guanosine and its derivatives with the aid of cation coordination, which does form polymeric structures in water and can result in gelation.[10,27] Our explanation for **TAPAS−CA** assembly also applies to G-tetrad formation, as the area of a G-tetrad is on the order of 1 nm^2. However, in the case of G-tetrads, two stacked tetrads are observed as an intermediate assembly,[10] and polymer formation is less cooperative than in the **TAPAS−CA** system. This tolerance of intermediate assemblies could be due to the smaller stacking area of the G-tetrad and electrostatic repulsions due to coordinated cations in the center of the G-tetrads and phosphate groups on each monomer, as incremental changes in monomer charge can greatly affect the degree of self-assembly in water.[10,28]

The **TAPAS−CA** assembly system presented here demonstrates that the hexameric rosette can be used as a functional architecture to generate hydrogels, which may be favorable for soft-material design and applications because of their chemical simplicity. For example, rosette nanotubes have recently been used in the formulation of hydrogel materials that enhance tissue growth.[29]

Finally, the inability of the RNA and DNA nucleobases to base-pair and assemble further in water has led to the proposal that the canonical nucleobases are products of evolution.[30] The **TAPAS−CA** system presented here forms polymer assemblies greater than 1 μm in length even at concentrations below 10 mM. Such assemblies contain over 18 000 highly organized monomers. This observation, along with the structural similarity of **TAPAS−CA** pairing with Watson−Crick base-pairing, suggests a possible prebiotic mechanism for *proto*-nucleobase selection from a complex mixture and organization into gene-length polymers even before linkage by a common backbone.

■ ASSOCIATED CONTENT

Supporting Information

Detailed experimental procedures and additional figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

hud@chemistry.gatech.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank A. E. Engelhart, F. A. L. Anet, G. B. Schuster, L. A. Lyon, L. D. Williams, and R. Krishnamurthy for discussions; S. Grijalvo and M. Terrazas for synthesis advice, L. Bottemely for AFM, M. Zhou for mass spectrometry, and the CNC EM center at Georgia Tech. This research was supported by NASA Exobiology [NNX08A014G] (B.J.C., N.V.H.) and jointly supported by the NSF and the NASA Astrobiology Program under the NSF Center for Chemical Evolution [CHE-1004570] (I.G., M.C.C., K.I.F., N.V.H.), the Parker H. Petit Endowment (B.J.C.), and Consejo Superior de Investigaciones Científicas (CSIC) [MEC, SAB2010-0163] (R.E., N.V.H.).

■ REFERENCES

(1) Aida, T.; Meijer, E. W.; Stupp, S. I. *Science* **2012**, *335*, 813.
(2) De Greef, T. F. A.; Smulders, M. M. J.; Wolffs, M.; Schenning, A. P. H. J.; Sijbesma, R. P.; Meijer, E. W. *Chem. Rev.* **2009**, *109*, 5687.
(3) Prins, L. J.; Reinhoudt, D. N.; Timmerman, P. *Angew. Chem., Int. Ed.* **2001**, *40*, 2382.
(4) Rehm, T. H.; Schmuck, C. *Chem. Soc. Rev.* **2010**, *39*, 3597.
(5) Hirschberg, J. H. K. K.; Brunsveld, L.; Ramzi, A.; Vekemans, J. A. J. M.; Sijbesma, R. P.; Meijer, E. W. *Nature* **2000**, *407*, 167.
(6) Ma, M. M.; Bong, D. *Langmuir* **2011**, *27*, 8841.
(7) Fenniri, H.; Packiarajan, M.; Vidale, K. L.; Sherman, D. M.; Hallenga, K.; Wood, K. V.; Stowell, J. G. *J. Am. Chem. Soc.* **2001**, *123*, 3854.
(8) Moralez, J. G.; Raez, J.; Yamazaki, T.; Motkur, K.; Kovalenko, A.; Fenniri, H. *J. Am. Chem. Soc.* **2005**, *127*, 8307.
(9) Bozsonyi, G.; Beingessner, R. L.; Yamazaki, T.; Cho, J.; Myles, A. J.; Malac, M.; Egerton, R.; Kawasaki, M.; Ishizuka, K.; Kovalenko, A.; Fenniri, H. *J. Am. Chem. Soc.* **2010**, *132*, 15135.
(10) Davis, J. T. *Angew. Chem., Int. Ed.* **2004**, *43*, 668.
(11) Davis, J. G.; Gierszal, K. P.; Wang, P.; Ben-Amotz, D. *Nature* **2012**, *491*, 582.
(12) Li, I. T. S.; Walker, G. C. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 16527.
(13) Chandler, D. *Nature* **2005**, *437*, 640.
(14) Raschke, T. M.; Tsai, J.; Levitt, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5965.
(15) Whitesides, G. M.; Simanek, E. E.; Mathias, J. P.; Seto, C. T.; Chin, D. N.; Mammen, M.; Gordon, D. M. *Acc. Chem. Res.* **1995**, *28*, 37.
(16) Lehn, J.-M.; Mascal, M.; Decian, A.; Fischer, J. *J. Chem. Soc., Perkin Trans. 2* **1992**, 461.
(17) Bohanon, T. M.; Denzinger, S.; Fink, R.; Paulus, W.; Ringsdorf, H.; Weck, M. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 58.
(18) Jelley, E. E. *Nature* **1936**, *138*, 1009.
(19) Hirst, A. R.; Coates, I. A.; Boucheteau, T. R.; Miravet, J. F.; Escuder, B.; Castelletto, V.; Hamley, I. W.; Smith, D. K. *J. Am. Chem. Soc.* **2008**, *130*, 9113.
(20) Prestegard, J. H.; Al-Hashimi, H. M.; Tolman, J. R. *Q. Rev. Biophys.* **2000**, *33*, 371.
(21) Wong, A.; Ida, R.; Spindler, L.; Wu, G. *J. Am. Chem. Soc.* **2005**, *127*, 6990.
(22) Choi, I. S.; Li, X. H.; Simanek, E. E.; Akaba, R.; Whitesides, G. M. *Chem. Mater.* **1999**, *11*, 684.
(23) Klostermeier, D.; Millar, D. P. *Biochemistry* **2002**, *41*, 14095.
(24) Mohan, S.; Hsiao, C.; Bowman, J. C.; Wartell, R.; Williams, L. D. *J. Am. Chem. Soc.* **2010**, *132*, 12679.
(25) Frieir, S. M.; Alkema, D.; Sinclair, A.; Neilson, T.; Turner, D. H. *Biochemistry* **1985**, *24*, 4533.
(26) Ts'o, P.; Melvin, I.; Olson, A. *J. Am. Chem. Soc.* **1963**, *85*, 1289.
(27) Yu, Y.; Nakamura, D.; DeBoyace, K.; Neisius, A. W.; McGown, L. B. *J. Phys. Chem. B* **2008**, *112*, 1130.
(28) Besenius, P.; Portale, G.; Bomans, P. H. H.; Janssen, H. M.; Palmans, A. R. A.; Meijer, E. W. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 17888.
(29) Zhang, L.; Rakotondradany, F.; Myles, A. J.; Fenniri, H.; Webster, T. J. *Biomaterials* **2009**, *30*, 1309.
(30) Engelhart, A. E.; Hud, N. V. *Cold Spring Harbor Perspect. Biol.* **2010**, DOI: 10.1101/cshperspect.a002196.