Mesoscale Folding: A Physical Realization of an Abstract, 2D Lattice Model for Molecular Folding

Insung S. Choi, Marcus Weck, Noo Li Jeon, and George M. Whitesides*

Department of Chemistry and Chemical Biology
Harvard University, 12 Oxford Street
Cambridge, Massachusetts 02138

Received May 30, 2000

Understanding protein folding—the translation of one-dimensional information contained in an amino acid sequence into a functional, three-dimensional structure—remains a central challenge to biophysics.[1−9] The many approaches to the folding problem have included detailed experimental studies,[2,3] modeling using molecular dynamics,[4−7] and abstract theoretical studies of folding on regular lattices.[8,9] Theory and experiment are most successful when they are connected, and the more abstract, statistical−mechanical models have never enjoyed the opportunities for testing and improvement that direct comparison of theory with experiment would afford.

The objective of this work is to provide a simple, physical realization of folding on a regular lattice. Our model is based on the two-dimensional (2D) folding of chains of hexagonal plates, floating at a fluid−fluid interface (Figure 1a). These chains are tethered together by flexible links and interact by directional capillary forces. We have described the characteristics of unconstrained (that is, unlinked) plates in similar environments.[10] Here we link these plates within a distance comparable to that over which the capillary forces between hydrophobic faces operate. When these systems are suspended at the water/perfluorodecalin (H_2O/PFD) interface and agitated by swirling on an orbital shaker, the plates with each chain self-assemble into regular structures determined both by the pattern of the hydrophobic and hydrophilic faces of each plate and by the sequence of the plates in a chain.

We fabricated the linked chains in poly(dimethylsiloxane) (PDMS) using a combination of photo- and soft lithography (Figure 2). The entire chain—plates and linkers—was fabricated as a unit, using a mold prepared by rapid prototyping[11] and two-level photolithography (see Supporting Information).[12] The linkers were thin PDMS threads (100 μm × 100 μm), designed as zigzags to increase flexibility. Faces and edges were differentiated into hydrophobic and hydrophilic sets using procedures described previously:[10] the linkers were hydrophilic.

Figure 1b describes the nomenclature we use to describe these plates. The nomenclature describes the shape of the basic unit (H for hexagon), the relative location of the linkers, and the pattern of the hydrophobic faces. In cases where a unit has two linkers, we put a prefix—ortho (*o*), meta (*m*), and para (*p*)—in front of the symbol H to describe the relative position of the two linkers to each other, in analogy to the nomenclature used to describe arenes. The hydrophobic faces are labeled as numbers in square parenthesis following the H. The numbering begins with the face

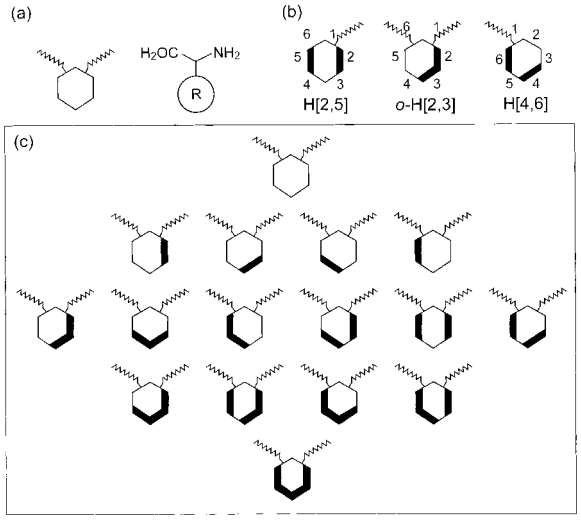

Figure 1. (a) Analogy between our models and molecular amino acids. (b) Nomenclature of the system used. The hydrophobic and hydrophilic faces are indicated by thick and thin lines, respectively. The number of the edges is clockwise, with "1" being a face with a linker. (c) Sixteen possible patterns of hydrophilic and hydrophobic faces for hexagons of the type *o*-H.

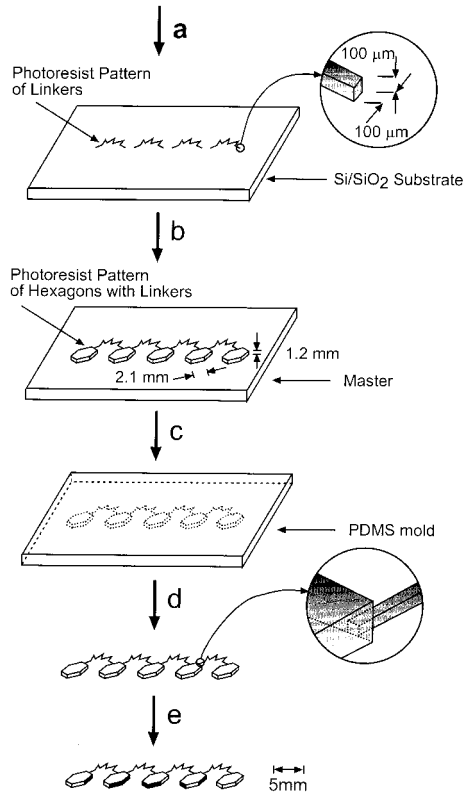

Figure 2. Schematic outline of the process used to fabricate the chains. The photoresist patterns (the "master") were generated using two-level photolithography. A mold was formed by casting and curing a prepolymer of PDMS against the master (c); the final chain was prepared in PDMS using the mold (d). Faces and edges were differentiated into hydrophobic and hydrophilic sets (e).

* To whom correspondence should be addressed.

(1) Karplus, M.; Shakhnovich, E. *Protein Folding: Theoretical Studies of Thermodynamics and Dynamics.* In *Protein Folding*; Creighton, T. E., Ed.; W. H. Freeman: New York, 1992.
(2) Fersht, A. R. *Curr. Opin. Struct. Biol.* **1997**, *7*, 3.
(3) Eaton, W. A.; Munoz, V.; Thompson, P. A.; Chan, C.-K.; Hofrichter, J. *Curr. Opin. Struct. Biol.* **1997**, *7*, 10.
(4) Shakhnovich, E. I. *Curr. Opin. Struct. Biol.* **1997**, *7*, 29.
(5) Dill, K. A.; Chan, H. S. *Nature Struct. Biol.* **1997**, *4*, 10.
(6) Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G. *Annu. Res. Phys. Chem.* **1997**, *48*, 545.
(7) Dobson, C. L.; Karplus, M. *Curr. Opin. Struct. Biol.* **1999**, *9*, 92.
(8) Crippen, G. M. *J. Mol. Biol.* **1996**, *260*, 467.
(9) Istrail, S.; Schwartz, R.; King, J. *J. Comput. Biol.* **1999**, *6*, 143.
(10) Bowden, N.; Choi, I. S.; Grzybowski, B. A.; Whitesides, G. M. *J. Am. Chem. Soc.* **1999**, *121*, 5373.
(11) Qin, D.; Xia, Y.; Whitesides, G. M. *Adv. Mater.* **1996**, *8*, 917.
(12) Madou, M. *Fundamentals of Microfabrication*; CRC Press: Boca Raton, FL, 1997; pp 1−50.

10.1021/ja001874o CCC: $19.00
Published on Web 11/14/2000

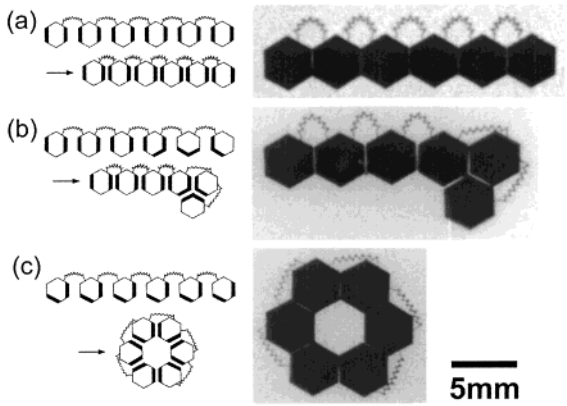

Figure 3. 2D folding of linked six hexagonal plates. (a) H[2,5]−(*o*-H[2,5])$_4$−H[3,6], (b) H[2,5]−(*o*-H[2,5])$_2$−*o*-H[2,3,5]−*o*-H[3,4]−H[5,6], and (c) H[2,4]−(*o*-H[2,4])$_4$−H[3,5].

to which the linker is attached, and continues in a clockwise direction. For example, "*o*-H[2,3]" indicates a hexagonal plate having two adjacent linkers and two adjacent hydrophobic faces (Figure 1b).

Figure 3 shows the outcome of 2D folding of structures comprising six linked hexagonal plates; these structures illustrate the influence of the pattern of hydrophobic faces of the hexagonal plates on the structures that form. There are 16 possible patterns of hydrophilic and hydrophobic faces for hexagons of the type *o*-H—a number close to the 20 common amino acids—assuming that the chain is directional (Figure 1c). Each chain has its own sequence of hexagonal plates with patterns of hydrophobic and hydrophilic faces; this sequence and these patterns direct its folding into a final structure.

The chains were placed at the interface in an extended conformation. The folding took place in seconds, and required no or only gentle agitation. We performed each experiment at least 10 times with each chain. The structures shown in Figure 3 formed in more than 90% of the trials. In these systems, the interactions between hydrophobic and hydrophobic faces are strongly attractive, the interactions between hydrophilic and hydrophilic faces are weakly attractive, and the interactions between hydrophobic and hydrophilic faces are strongly repulsive.[10] Figure 3a shows a chain of the sequence H[2,5]−(*o*-H[2,5])$_4$−H[3,6], and the structure it forms on self-assembly. The structure is the one expected from aggregates of [1,4] hexagons not constrained by linkers.[10] The structure shown in Figure 3b was derived from the sequence H[2,5]−(*o*-H[2,5])$_2$−*o*-H[2,3,5]−*o*-H[3,4]−H[5,6]. This sequence was designed to have a trimeric domain. As a control, we used three [1,4], one [1,2,4], and two [1,2] hexagons unconstrained by linkers; these experiments generated a number of structures, with no single one dominant. The sequence, H[2,4]-(*o*-H[2,4])$_4$−H[3,5], formed a cyclic hexamer. Unconstrained [1,3] hexagons self-assemble into zigzag structures with only a few cyclic hexamers.[10]

Figure 3 establishes two characteristics of this model of folding: (1) that linking can substantially increase the probability of formation of ordered structures, (2) that information coded in the pattern of hydrophobic edges in the chain determines the final structure.

We extended the strategy of folding hexagonal plates to the folding of more complex structures (Figure 4). Figure 4a shows a folded structure of seven linked hexagonal plates, H[2,4]−(*o*-H[2,4])$_3$−*o*-H[2,3,5]−*o*-H[3,4]−H[3,5,6]. This sequence contains two of the folded motifs shown in Figure 3 (a cyclic hexamer and a trimer). This compact structure formed in greater than 95%

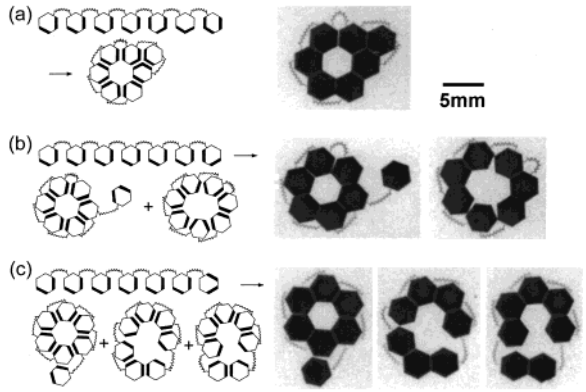

Figure 4. 2D folding of linked seven hexagonal plates. (a) H[2,4]−(*o*-H[2,4])$_3$−*o*-H[2,3,5]−*o*-H[3,4]−H[3,5,6], (b) H[2,4]−(*o*-H[2,4])$_5$−H[4,6], and (c) H[2,4]−(*o*-H[2,4])$_5$−H[2,4].

of trials. Other sequences—H[2,4]−(*o*-H[2,4])$_5$−H[4,6] (Figure 4b) and H[2,4]−(*o*-H[2,4])$_5$−H[2,4] (Figure 4c)—gave mixtures of structures rather than unique, compact structures.

This system has both strengths and weaknesses as a model with which to explore aspects of folding of both of molecules and of non-molecular structures. The strengths are that: (i) it is very simple and flexible; (ii) capillarity is a conceptually well-understood interaction,[13] albeit complicated to model; (iii) capillarity is related to the hydrophobic effect,[14,15] and thus directly relevant to a biologically significant interaction;[16] (iv) the chains can be designed conveniently using a computer, and fabricated easily using rapid prototyping;[11] (v) the number of hexagons, the dimensions and patterns of hydrophobic faces on them, and the length, location, and stiffness of the linkers between them—all can be easily controlled; (vi) arbitrary 2D shapes can be substituted for the hexagons;[17] (vii) the relative densities of the various phases involved, and thus the value of the capillary interactions, can be adjusted over substantial ranges.[18,19] The weaknesses of this system are that it involves 2D rather than 3D folding (and that, as a consequence, the number of structures that can form is much smaller than with proteins or polypeptides with the same number of units) and that it uses only two types of interactions—steric and capillary—rather than the larger range important for molecules. For all of its simplicity, the system provides a physical realization of the abstract analytical models that are used to investigate molecular folding. This approach also suggests—when applied to objects in micrometer ranges—routes of assembling aggregated structures that would otherwise be difficult to obtain.

Acknowledgment. This work was supported by the National Science Foundation (CHE-9901358, ECS-9729405). M.W. thanks the German Academic Exchange Service (DAAD) for a postdoctoral fellowship.

Supporting Information Available: Experimental procedure (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA001874O

(13) Rowlinson, J. S.; Widom, B. *Molecular Theory of Capillary*; Clarendon Press: Oxford, 1982.
(14) Blokzijl W.; Engberts, J. B. F. N. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 1545.
(15) Hummer, G.; Garde, S.; Garcia, A. E.; Paulaitis, M. E.; Pratt, L. R. *J. Phys. Chem. B* **1998**, *102*, 10469.
(16) *Water and Biological Macromolecules*; Westhof, E., Ed.; CRC Press: Boca Raton, FL, 1993.
(17) Weck, M.; Choi, I. S.; Jeon, N. L.; Whitesides, G. M. *J. Am. Chem. Soc.* **2000**, *122*, 3546.
(18) Bowden, N. Oliver, S. R. J.; Whitesides, G. M. J. *Phys. Chem. B* **2000**, *104*, 2714.
(19) Velev, O. D.; Lenhoff, A. M.; Kaler, E. W. *Science* **2000**, *287*, 2240.