JACS ARTICLES

Published on Web 11/29/2010

Atomistic Simulations of 2D Bicomponent Self-Assembly: From Molecular Recognition to Self-Healing

Carlos-Andres Palma,[†] Paolo Samorì,* and Marco Cecchini*

ISIS-CNRS 7006, Université de Strasbourg, 8 allée Gaspard Monge, 67000 Strasbourg, France

Received September 1, 2010; E-mail: samori@unistra.fr; mcecchini@unistra.fr

Abstract: Supramolecular two-dimensional engineering epitomizes the design of complex molecular architectures through recognition events in multicomponent self-assembly. Despite being the subject of in-depth experimental studies, such articulated phenomena have not been yet elucidated in time and space with atomic precision. Here we use atomistic molecular dynamics to simulate the recognition of complementary hydrogen-bonding modules forming 2D porous networks on graphite. We describe the transition path from the melt to the crystalline hexagonal phase and show that self-assembly proceeds through a series of intermediate states featuring a plethora of polygonal types. Finally, we design a novel bicomponent system possessing kinetically improved self-healing ability in silico, thus demonstrating that a priori engineering of 2D self-assembly is possible.

Introduction

Molecular self-assembly of artificial architectures with precise 3D arrangements[1] is key to the bottom-up fabrication of multifunctional materials. As simplified models, 2D supramolecular networks are ideal test beds for unraveling the principles of multicomponent self-assembly. Although various experimental techniques exist for the exploration of materials with atomic precision,[2] the level of understanding of the subtle details of the process of 2D supramolecular self-assembly is still limited, and full control over the resulting phase is far from being reached. This is exemplified by a series of recent scanning probe microscopy analyses, which independently reported the coexistence of numerous structurally diverse self-assembled phases at the solid−liquid interface[3] ranging from 2D amorphous[4] and polygonal networks[5] to crystals.[6] In addition, it has been proposed that the self-assembled architectures at surfaces may actually correspond to either equilibrium[7] or nonequilibrium states, which can be reversible[4a] or irreversible.[5b] Consequently, any bottom-up material design would require a thorough and a priori knowledge of both the thermodynamics and the kinetics of molecular self-assembly on the time and length scales relevant to the process, which includes an accurate description of the structure of the multicomponent system as a function of temperature, pressure, and relative concentrations of the molecular modules (i.e., phase diagram) as well as of the transition pathways and barriers to the thermodynamically stable state. Only such a detailed knowledge may ultimately allow one to bridge the gap between the concepts of chemical design and self-assembly into custom architectures, thus paving the way toward supramolecular engineering.

In this manuscript we simulate the 2D self-assembly of a model bicomponent system, i.e., melamine and a bis(N_1-hexyl-uracil) module, exposing complementary recognition sites (*mel*·*ura*, Figure 1A). The melamine's donor−acceptor−donor (DAD) coupling with acceptor−donor−acceptor (ADA) imidic moieties in the *mel*·*ura* bicomponent systems together with the 3-fold symmetry imposed by the melamine cornerstone is known to lead to porous networks on graphite, which were shown by scanning tunneling microscopy (STM) to feature a variety of polygonal types, including pentagons, hexagons, and heptagons.[4a] Such a structural diversity offers the opportunity to explore the nature of 2D polygonal networks on the submolecular scale and cast light onto the structural evolution to the crystalline phase.

Results

The study of self-assembly at the solid−liquid interface by molecular dynamics is accomplished through the description

[†] Present address: Max Planck Institute for Polymer Research, Ackermannweg 10, D-55128 Mainz, Germany.

(1) (a) Whitesides, G. M.; Mathias, J. P.; Seto, C. T. *Science* **1991**, *254*, 1312. (b) Fujita, M.; Oguro, D.; Miyazawa, M.; Oka, H.; Yamaguchi, K.; Ogura, K. *Nature* **1995**, *378*, 469. (c) Lehn, J. M. *Science* **2002**, *295*, 2400. (d) Yaghi, O. M.; O'Keeffe, M.; Ockwig, N. W.; Chae, H. K.; Eddaoudi, M.; Kim, J. *Nature* **2003**, *423*, 705. (e) Rothemund, P. W. K. *Nature* **2006**, *440*, 297.

(2) Barth, J. V.; Costantini, G.; Kern, K. *Nature* **2005**, *437*, 671.

(3) (a) Elemans, J. A. A. W.; Lei, S. B.; De Feyter, S. *Angew. Chem., Int. Ed.* **2009**, *48*, 7298. (b) Ciesielski, A.; Palma, C. A.; Bonini, M.; Samorì, P. *Adv. Mater.* **2010**, *22*, 3506.

(4) (a) Palma, C. A.; Bjork, J.; Bonini, M.; Dyer, M. S.; Llanes-Pallas, A.; Bonifazi, D.; Persson, M.; Samorì, P. *J. Am. Chem. Soc.* **2009**, *131*, 13062. (b) Otero, R.; Lukas, M.; Kelly, R. E. A.; Xu, W.; Laegsgaard, E.; Stensgaard, I.; Kantorovich, L. N.; Besenbacher, F. *Science* **2008**, *319*, 312.

(5) (a) Llanes-Pallas, A.; Palma, C. A.; Piot, L.; Belbakra, A.; Listorti, A.; Prato, M.; Samorì, P.; Armaroli, N.; Bonifazi, D. *J. Am. Chem. Soc.* **2009**, *131*, 509. (b) Zwaneveld, N. A. A.; Pawlak, R.; Abel, M.; Catalin, D.; Gigmes, D.; Bertin, D.; Porte, L. *J. Am. Chem. Soc.* **2008**, *130*, 6678.

(6) (a) Theobald, J. A.; Oxtoby, N. S.; Phillips, M. A.; Champness, N. R.; Beton, P. H. *Nature* **2003**, *424*, 1029. (b) Schlickum, U.; Decker, R.; Klappenberger, F.; Zoppellaro, G.; Klyatskaya, S.; Ruben, M.; Silanes, I.; Arnau, A.; Kern, K.; Brune, H.; Barth, J. V. *Nano Lett.* **2007**, *7*, 3813.

(7) Garrahan, J. P.; Stannard, A.; Blunt, M. O.; Beton, P. H. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 15209.

(8) (a) Langevin, P. *C. R. Acad. Sci.* **1908**, *146*, 530. (b) Brooks, B. R.; et al. *J. Comput. Chem.* **2009**, *30*, 1545.

10.1021/ja107882e

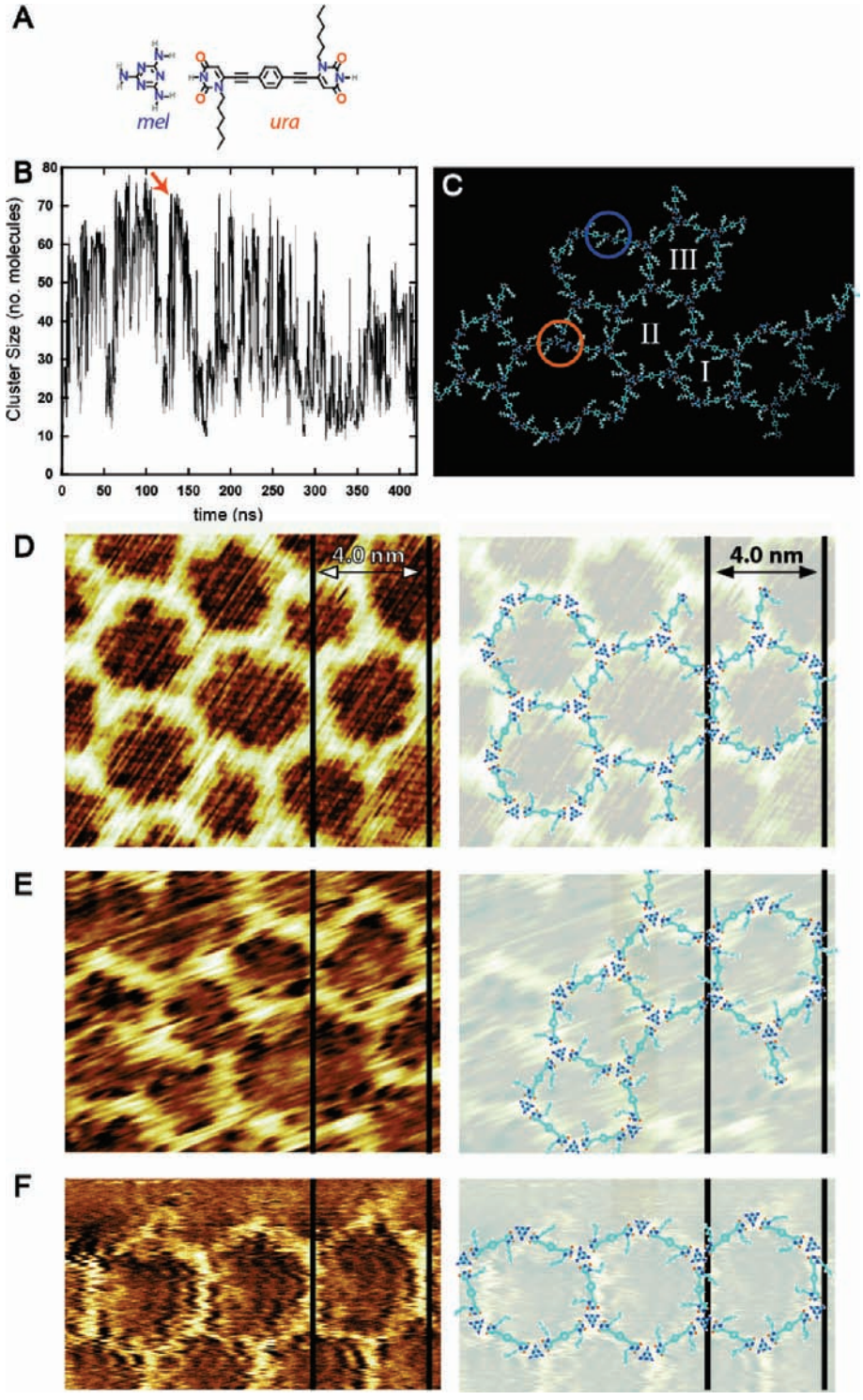

Figure 1. Self-assembly simulations of *mel*·*ura*. (A) Chemical structure of the complementary heterodimer. The *mel*·*ura* recognition is based on donor−acceptor−donor (DAD) coupling with acceptor−donor−acceptor (ADA) imidic moieties. (B) Time series of the size (i.e., number of molecules) of the largest supramolecular cluster in a typical fast-annealing trajectory (black line), i.e., a simulation performed at 400 K with temperature spikes up to 525 K. (C) Molecular structure of a 73-mer cluster (out of 80) extracted at 137 ns (red arrow in Figure 1B). Roman numbers indicate supramolecular polygons and red and blue circles highlight *mel*·*mel* and *ura*·*ura* homorecognition events, respectively. (D−F) Comparison between clusters extracted from the simulations and experimental STM images. On the left-hand side, the supramolecular pattern of experimentally observed polymorphic networks is shown. On the right-hand side, molecular snapshots extracted from the simulations are superposed to the STM images. The correlation between the complex experimental patterns and the simulated assemblies is striking.

of the sequence of events (from molecular recognition to phase properties) at an atomic level of detail. We chose Langevin dynamics[8] with continuum electrostatics in conjunction with a restraining harmonic potential to mimic the presence of graphite, which is referred to as the implicit substrate. The cornerstone (*mel*) and the linker (*ura*) were modeled by the Merck molecular force field (**MMFF**),[9] while atomic charges were modified ad hoc to fit the *mel*·*ura* interaction energy profile estimated by

(9) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490.

density functional theory.[4a] The quality of the implicit substrate was assessed by comparing the lifetime and the apparent diffusion constant of a preformed *mel·ura* dimer with those obtained with an all-atom representation of a graphene slab; the latter was reported to yield adsorption energies for a series of aromatic compounds on graphite in good agreement with experimental data.[10] The analysis confirmed that the implicit model provides a representative description of the substrate while allowing for microseconds sampling in a few weeks of calculation. Significantly, the *mel·ura* dimer physisorbed on the implicit substrate was found to have a lifetime exceeding 100 ns at temperatures below 400 K. For this reason all simulations were performed at temperatures above 400 K, where self-assembly is reversible and the model mimics the dynamic experimental conditions.[11] All simulations were performed with the program CHARMM[8b,12] in a constant dielectric continuum electrostatics model with a cutoff of 18 Å for the nonbonded interactions. The SHAKE[13] algorithm was used to fix the length of all covalent bonds, thus allowing for an integration time step of several femtoseconds. Further details on the modeling of the self-assembly modules and graphite substrate and the simulation conditions are given in the Supporting Information.

The self-assembly of *mel·ura* was investigated by Langevin molecular dynamics (MD) simulations of an 80-molecule stoichiometric mixture adsorbed on the substrate. Twenty independent runs were done at 400 K for 500 ns by employing a "fast-annealing" protocol that features temperature spikes at 525 K. Since the *mel·ura* dimer dissociates in the subnanosecond time scale at 500 K, the procedure results in an efficient conformational search. During annealing, a plethora of discrete supramolecular structures based on heteromers as large as 70-mers rapidly assemble and disassemble (see Figure 1B); they are hereafter referred to as clusters. The motif of a transient cluster extracted at 137 ns is portrayed in Figure 1C. It shows a 73-mer combining a pentagon (I), a heptagon (II), and a hexagon (III) as well as the existence of *mel·mel* and *ura·ura* homorecognition events. When compared with STM images, the correlation between the structure of the simulated clusters and the experimental patterns is striking. Figure 1D−F shows three structural motifs extracted from the fast-annealing trajectories superimposed on STM recorded images. Interestingly, out of hundreds of clusters composed of >50-mers, the largest "crystallike" fragment from the simulations was composed of only three consecutive hexagons (Figure 1F), out of nine that can possibly form in the 80-molecule crystalline configuration. Conversely, hexagonal crystalline domains are ubiquitously observed in the experiments. This suggests that the large heterogeneous population of polygonal networks is a kinetic product, which suddenly forms in the early steps of 2D self-assembly. To confirm this observation, the difference in conformational free energy (ΔG) between potential nuclei of archetypical size, i.e., the smallest polygons, was investigated by the confinement approach (see Supporting Information for details).[14] The free energy of the supramolecular hexagon, which is structurally compatible with the crystalline pattern, was compared to that of the branched pentagon, a supramolecular discrete architecture steering formation of polygonal networks. The analysis yields a ΔG of 2.3 ± 1.0 kcal/mol at 400 K in favor of the hexagon, implying a theoretical ratio between the populations of the hexagon and the pentagon at equilibrium amounting to 18. Conversely, the average over the "fast-annealing" simulations revealed essentially equal populations, their ratio being 0.66 (Figure S15, Supporting Information). These results provide strong evidence that polygonal defects, such as pentagons, are kinetically formed species, promoting formation of noncrystalline phases. Moreover, taking into account that crystalline domains are found to extend far beyond polygonal domains in STM images,[4a] the hexagonal pattern is predicted to be the thermodynamically stable state. In light of this, the existence of a structural transition from the glass-like network to the crystalline state, i.e., a self-healing process, is expected in the simulations.

To prove the existence of such a transition, we first approached the problem by constant temperature MD simulation of stoichiometric *mel·ura*. At 375 K the system shows no structural evolution to ordering on the submicrosecond time scale, leading only to 2D amorphous polymers. At 420 K the system is more dynamic and large molecular clusters spontaneously assemble. Interestingly, the clusters exhibit a steady growth on the nanosecond time scale along with large fluctuations in size on the tens of nanoseconds, leading to an overall structural increase on the hundreds of nanoseconds (Figure 2A, black line). The comparison of the time-dependent trend of the number of hetero- and homorecognition events (see Figure 2A) reveals that the former (in blue) inversely mirrors the latter (in red). Such an observation suggests that although hetero-hydrogen bonding has a constant increase with time, continuous replacement of heterointeractions with homointeractions and vice versa takes place. This is a consequence of the similar binding energies amounting to 13 kcal/mol for *ura·ura* and 11 kcal/mol for *mel·mel* if compared to 18 kcal/mol for *mel·ura*. The competitive interplay gives rise to additional barriers, which effectively trap the system in a glass-like state[15] on the simulation time scale. To test this hypothesis and speed up the glass to crystal transition in silico, the interaction energy of all homorecognition events (i.e., *mel·mel* and *ura·ura*) was artificially disfavored by scaling them down by a factor of 100 while preserving all other simulation conditions. This corresponds to an idealized system essentially devoid of homorecognition, referred to as *mel·ura** (see Supporting Information for details). When *mel·ura** is simulated at 420 K, the clusters grow very rapidly (cf. Figure 2A and 2B), and a marked structural evolution to ordering is detected by the appearance of a peak at 41 Å in the radial distribution function $g(r)$. Such a peak corresponds to the unit cell of a fully crystalline pattern, and its intensity gauges the degree of crystallinity (Figure 2C). Analysis of the evolution of the system in time reveals a series of structural intermediates with increased ordering, as depicted in Figure 2D. In the subnanosecond range (blue region) the average cluster is composed of three interacting molecules (state A). Between 1 and 5 ns, the cluster size increases and 20-mers

(10) Bjork, J.; Hanke, F.; Palma, C. A.; Samorì, P.; Cecchini, M.; Persson, M. Submitted for publication.
(11) We previously reported that the presence of a supernatant solvent in STM studies confers reversibility to self-assembly at the solid−liquid interface.[4a] In the computations, such a reversibility is achieved by tuning the temperature of the system rather than the effective dielectric constant (Supporting Information section 2f).
(12) Brooks, B.; Bruccoleri, R.; Olafson, B.; States, D.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(13) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(14) (a) Cecchini, M.; Krivov, S. V.; Spichty, M.; Karplus, M. *J. Phys. Chem. B* **2009**, *113*, 9728. (b) Tyka, M. D.; Clarke, A. R.; Sessions, R. B. *J. Phys. Chem. B* **2006**, *110*, 17212.
(15) Zallen, R. *The Physics of Amorphous Solids*; Wiley-VCH: New York, 1998.

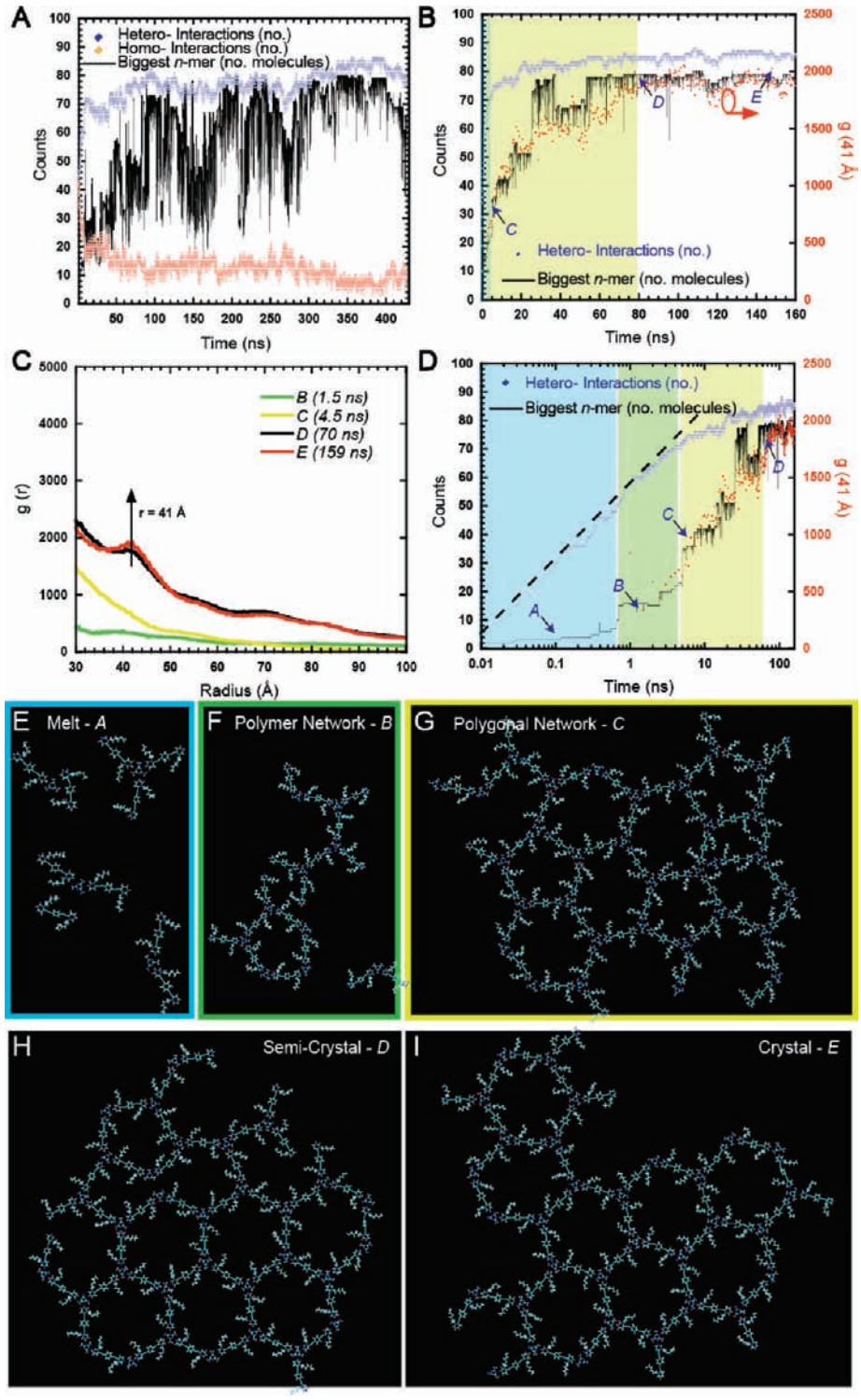

Figure 2. Self-assembly simulations of *mel·ura**. (A) Time evolution of *mel·ura* at 420 K. The time series of the size of the largest cluster and those of the number of hetero- and homorecognition events are shown in black, red, and blue, respectively. (B) Time evolution of *mel·ura** at 420 K. The size of the largest cluster (black) and the number of heterorecognition events (blue) are shown. By construction, homorecognition events are absent (see main text). The figure also shows the time series of $g(r)$ at $r = 41$ Å (red dots). (C) Radial distribution function, $g(r)$, determined at different time frames of the *mel·ura** simulation. (D) Logarithmic plot of B. Four distinct phase behaviors are apparent. (E−I) Structural representation of the intermediate states identified in D.

assemblies appear (B, green region). At this stage, a big step in the average cluster size is observed (C) followed by a drop in the slope of the heteroassociation rate (blue line). Both the $g(r)$ and the cluster size increase until they reach a plateau at about 100 ns (state D, yellow region). Overall, four main structural transitions are identified. They correspond to the oligomerization of the melt (state A, Figure 2E), the appearance of polymeric networks (state B, Figure 2F), clustering into 2D polygonal networks (state C, Figure 2G), and relaxation to the crystalline state (state D, Figure 2H and 2I). The transition from state C to D in Figure 2 provides the first direct proof of self-healing occurring in silico. The same structural evolution to the crystalline state occurring through a series of polygonal intermediates was observed in multiple simulation trajectories.

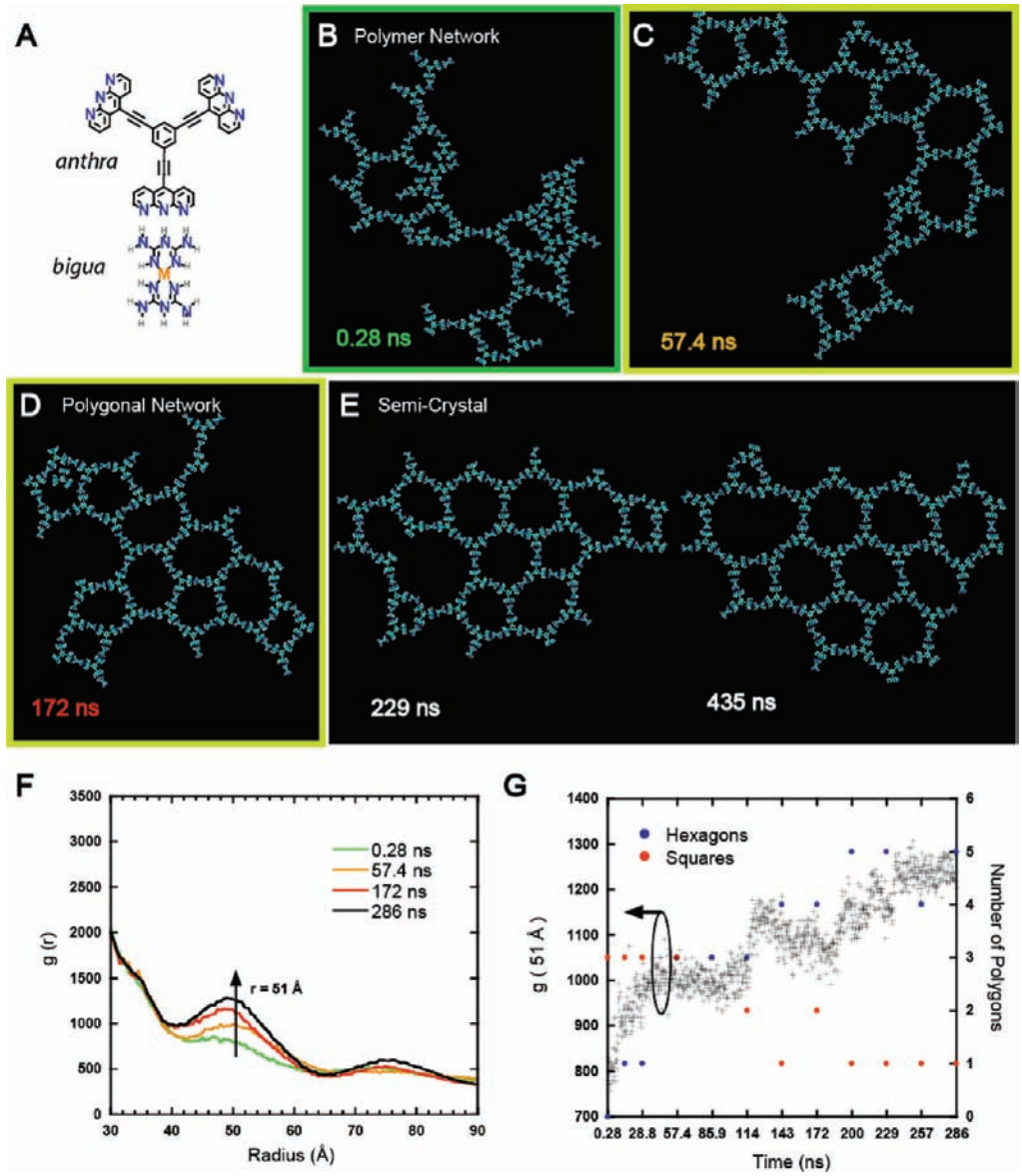

Figure 3. Self-assembly simulations of *anthra·bigua*. (A) Chemical structure of the complementary heterodimer. The *anthra·bigua* recognition is based on donor−donor−donor coupling with acceptor−acceptor−acceptor triaza-anthracene moieties. The metal (M) of the *bigua* module was replaced by an artificial bond in the simulations. (B−E) Molecular snapshots extracted from a typical self-assembly simulation at 520 K depicting the structural intermediates leading to the semicrystalline phase. (F) Radial distribution function, $g(r)$, at chosen times. The band centered at $r = 51$ Å, which corresponds to the unit cell vector length for a fully crystalline pattern, shows the evolution of the system to ordering. (G) Time series of $g(r)$ at a distance of 51 Å. Blue and red dots, respectively, correspond to the number of hexagons and squares sampled in the simulation as a function of time.

A visual representation of the *mel·ura** transition pathway from the melt to the hexagonal phase is given in the Supporting Information (see video S2). The simulations show that two-dimensional crystallization mediated by hydrogen-bonding recognition proceeds via a discrete series of metastable (kinetic) intermediate states until the most stable (thermodynamic) supramolecular architecture is reached. In a way, these results provide a microscopic picture of what is known as Ostwald's rule of stages.[16]

By capitalizing on the knowledge acquired on the *mel·ura** system, we formulate a novel bicomponent featuring simultaneously improved self-healing kinetics with enhanced thermodynamics control. For this purpose, a system naturally devoid of homorecognition was designed by exploiting two new complementary modules, the former exposing three parallel hydrogen bond donors that recognize three parallel hydrogen-bond acceptors belonging to the second molecule (*anthra·bigua*, Figure 3A). In such a unique triple-recognition process, primary interactions are further stabilized by secondary intermolecular interactions which all feature an attractive nature. For the sake of computational simplicity the metal in the *bigua* module was replaced with an artificial bond bridging the two biguanide moieties. In such a system, the model's heterorecognition association energy amounts to 52 kcal/mol. By following the strategy used for *mel·ura*, the self-assembly of an 80-molecule *anthra·bigua* bicomponent was simulated at 520 K and dielectric constant of 2.5.[17] Figure 3B−E shows the structural evolution of the system in a typical simulation trajectory. The

(16) (a) Ostwald, W. *Z. Phys. Chem.* **1897**, *22*, 289. (b) Chung, S. Y.; Kim, Y. M.; Kim, J. G.; Kim, Y. J. *Nat. Phys.* **2009**, *5*, 68.

(17) For the *anthra-bigua* system, the dielectric constant was changed to reduce the dissociation energy (E_{diss}) of the dimer to 20 kcal/mol. Also, the temperature was increased to 520 K to have kT/E_{diss} comparable with that of *mel-ura*.

melt phase rapidly aggregates into a 2D amorphous glass in tenths of picoseconds until a stable cluster is formed (Figure 3B). The cluster rearranges into a 2D glassy network in about 100 ns (Figure 3C), evolves into a polygonal network (Figure 3D), and finally undergoes a spontaneous transition to a semicrystal featuring five hexagonal pores (Figure 3E). The latter further evolves into a six-hexagons- structure in about 0.5 μs. Similar results were obtained in multiple simulations runs (see Supporting Information). Strikingly, the de novo designed *anthra·bigua* shows the spontaneous transition to the semicrystalline phase observed for the model system *mel·ura** (see Figure 3F and 3G).

Conclusions

In summary, computer-assisted self-assembly is a most valuable tool for making 2D supramolecular engineering a fully predictive discipline. The results presented in this paper demonstrate that atomistic computer simulations of small organic compounds physisorbed on a substrate are key to identification of the design rules of 2D self-assembly, thus providing a novel framework for the rational optimization of the chemical structure of the molecular modules. We demonstrated that knowledge of the free-energy landscape could be exploited to elucidate bicomponent self-assembly into crystalline networks in silico and showed that crystallization proceeds through a series of glass-like intermediates. The reversible nature of the self-assembly process was evidenced by monitoring the self-healing of a 2D polygonal network into a 2D crystal. By investigating the 2D self-assembly of a model system, we captured the molecular origin of the frustration of the underlying free energy surface and used it to design a novel bicomponent system with improved self-healing kinetics. The latter is expected to show extremely rapid structural reorganization to the crystalline pattern (i.e., submicrosecond) in response to external stimuli (i.e., mechanical or thermal) and is proposed as a prototypical example of a novel class of intelligent materials. In addition, the design of complementary modules exposing either hydrogen-bonding donor or hydrogen-bonding acceptor moieties at the recognition sites, which was introduced to suppress the occurrence of homorecognition, provides a simple solution to the common propensity of preprogrammed multicomponents to undergo phase segregation. The latter is expected by design to confer enhanced material stability to degradation and aging. These findings are key for de novo material design, where understanding the link between the chemical nature of the building blocks and the properties of the resulting architectures is of paramount importance to translate information into function. This problem is highly reminiscent of the relationship between the primary sequence and the three-dimensional structure of proteins.

Acknowledgment. We are grateful to Prof. Martin Karplus for helpful discussions and support and Dr. Anna Llanes-Pallas and Prof. Davide Bonifazi for the conception and design of the flexible uracyl-like linkers. We are also grateful to Dr. Martin Spichty for development of the ^{5}H fast integration routine and MMFF force-field support and to Dr. Francesco Rao for helpful suggestions. Efficient trajectory handling and part of the analysis were done with the help of the program Wordom.[18] This work was granted access to the HPC resources of CCRT/CINES/IDRIS under the allocation 2009 (075114) made by GENCI (Grand Equipement National de Calcul Intensif). This work was financially supported by the EC Marie-Curie RTNs PRAIRIES (MRTN-CT-2006-035810), ITN-SUPERIOR (PITN-GA-2009-238177), the EC FP7 ONE-P large-scale project no. 212311, the NanoSci-E+ project SENSORS, and the International Center for Frontier Research in Chemistry (FRC, Strasbourg). M.C. was supported by a grant of the Human Frontier Science Program.

Supporting Information Available: Scanning tunneling microscopy; molecular modeling and simulation setup; free energy calculations; self-assembly movies for *mel-ura*, *mel-ura** and *anthra-bigua*; complete ref 8b. This material is available free of charge via the Internet at http://pubs.acs.org. Higher resolution movies can be found at http://www-isis.u-strasbg.fr/nanochem/index.php?option=com_content&view=article&id=36&Itemid=41.

JA107882E

(18) Seeber, M.; Cecchini, M.; Rao, F.; Settanni, G.; Caflisch, A. *Bioinformatics* **2007**, *23*, 2625.