HHS Public Access

Author manuscript

J Am Chem Soc. Author manuscript; available in PMC 2020 September 01.

Published in final edited form as:

J Am Chem Soc. 2019 November 06; 141(44): 17588–17600. doi:10.1021/jacs.9b06410.

Sequence-Defined Macrocycles for Understanding and Controlling the Build-up of Hierarchical Order in Self-Assembled 2D Arrays

James R. Dobscha[†,‡], **Henry D. Castillo**[†,‡], **Yan Li**[†], **Rachel E. Fadler**[†], **Rose D. Taylor**[†,§], **Andrew A. Brown**[†,‖], **Colleen Q. Trainor**[†,⊥], **Steven L. Tait**[*], **Amar H. Flood**[*,†]

[†]Molecular Materials Design Laboratory, Department of Chemistry, Indiana University, 800 East Kirkwood Avenue, Bloomington, Indiana 47405, United States

Abstract

Anfinsen's dogma that sequence dictates structure is fundamental to understanding the activity and assembly of proteins. This idea has been applied to all manner of oligomers but not to the behavior of cyclic oligomers, aka macrocycles. We do this here by providing the first proofs that sequence controls the hierarchical assembly of nonbiological macrocycles, in this case, at graphite surfaces. To design macrocycles with one (AAA), two (AAB), or three (ABC) different carbazole units, we needed to subvert the synthetic preferences for one-pot macrocyclizations. We developed a new stepwise synthesis with sequence-defined targets made in 11, 17, and 22 steps with 25, 10, and 5% yields, respectively. The linear build up of primary sequence (1°) also enabled a thermal Huisgen cycloaddition to proceed regioselectively for the first time using geometric control. The resulting macrocycles are planar (2° structure) and form H-bonded dimers (3°) at surfaces. Primary sequences encoded into the suite of tricarb macrocycles were shown by scanning-tunneling microscopy (STM) to impact the next levels of supramolecular ordering (4°) and 2D crystalline polymorphs (5°) at solution-graphite interfaces. STM imaging of an AAB macrocycle revealed the formation of a new gap phase that was inaccessible using only C_3-symmetric macrocycles. STM imaging of two additional sequence-controlled macrocycles (AAD, ABE) allowed us to identify the factors driving the formation of this new polymorph. This demonstration of how sequence controls the hierarchical patterning of macrocycles raises the importance of stepwise syntheses relative to one-pot macrocyclizations to offer new approaches for greater understanding and control of hierarchical assembly.

Graphical Abstract

[*]**Corresponding Authors:** tait@indiana.edu, aflood@indiana.edu.

[§](R.D.T.) Department of Chemistry, College of Wooster, 1189 Beall Avenue, Wooster, OH.

[‖](A.A.B.) Department of Chemistry, Louisiana Tech University, 1 Adams Boulevard, Ruston, LA.

[⊥](C.Q.T.) Department of Chemistry, Hillsdale College, 33 East College Street, Hillsdale, MI.

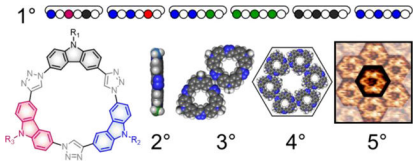
[‡]**Author Contributions:** J.R.D. and H.D.C. contributed equally to this work.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.9b06410.

General experimental procedures, synthetic schemes, and spectra; additional STM images and characterization (PDF)

The authors declare no competing financial interest.



INTRODUCTION

Sequence information is foundational to nature's structural and functional complexity. We know that the primary sequence of amino acids controls protein folding[1,2] and that the residues displayed on a protein's exterior surface direct their hierarchical assembly into biological nanostructures.[3–5] This bottom-up control of molecular architectures has long motivated interest in establishing correlations between sequence information and structure in natural systems, e.g., DNA[6–8] and proteins,[9–11] and it is now being explored in artificial systems like sequence-controlled polymers[12–19] and sequence-defined foldamers.[20–27] Biomimetic sequence control is an exciting new challenge for artificial systems. To help bridge that gap, structure-property relationships are being explored with DNA origami[28–33] and de novo protein design.[34–37] For example, Ulijn has leveraged the diversity of 20 amino acids in tripeptides to identify those that fold up and provide the right 3D display of side chains to enhance their self-assembly.[38–40] Cyclic molecules offer superior control over the 3D display of functional groups as a result of a smaller set of conformations. In this way, for example, Kirshenbaum and Logan designed sequence-specific cyclic peptoids to bind and disrupt protein-protein interfaces.[41] Use of these approaches with nonbiological macrocycles are not known.[42] We explore the use of primary sequences (1°, Figure 1a) encoded into macrocycles (Figure 1c–e) with planar secondary structures (2°) that are shape persistent in order to limit conformations (Figure 1b) and develop new syntheses using elementary chemistries and high building block versatility to make three-component sequences (AAA, AAB, and ABC type). We show how sequence controls the build up of hierarchical structures on graphite surfaces into dimers (3°), local superstructures (4°), and then polymorphs (5°).

Many macrocycles are composed of repeating monomers, and thus, sequence information is latent to their structure. To lay a foundation for exploring sequence information, however, synthetic control over sequence must be achieved. Examples of synthetically accessible macrocycles intrinsically encoded with sequence information have been known since the total synthesis of chlorophyll A.[43,44] Macrocycles in which we can identify sequences ex post facto have been reported; e.g., the sequence of nonsymmetric calixarenes have been shown to influence guest recognition,[45–47] the performance of supramolecular catalysts,[48–50] as well as intramolecular reactivity[51] but sequence-level information was never addressed explicitly. Sequence has also been indirectly investigated in heteromacrocycles, defined as those macrocycles that do not have a regularly repeating backbone. These include expanded porphyrins,[52–59] amide-based macrocycles that form nanopores and transport ions,[60] and a coumarin-containing triamide macrocycle used as a sensor.[61] Use of a flexible heteromacrocycle that was explicitly encoded with a three-unit sequence to enable unidirectional motion in a catenane[62] showcases the potential benefits of this type of

structural control. However, explicit investigation of the role of sequence on the hierarchical ordering of the many macrocycles capable of self-association[63–68] is unexplored.

Motivation for sequence-controlled macrocycle syntheses must overcome the traditional priorities of an expedient synthesis. Shape-persistent macrocycles are made using either convergent one-pot reactions or multistep pathways. Step-wise routes are tenable but often longer and have traditionally been given less value relative to one-pot methods. This valuation is seen in developments of one-pot preparations of macrocycles after their stepwise syntheses.[69–73] With the one-pot approach, the macrocycles offer greater scalability and, with just one repeating monomer, have more straightforward structure-property relationships. These benefits are achieved at the expense of diversity; only a limited set of interactions can be encoded in the macrocycle. Nonsymmetric macrocycles are sometimes made in one-pot, e.g., porphyrins,[74–76] but statistical product distributions (Figure 2a) limit yields and impede isolation. These one-pot reactions also fail to control the sequence. One approach to circumventing these limitations has been to incorporate orthogonal dynamic covalent chemistry into the one-pot synthesis of heteromacrocycles.[77,78] While this method minimizes side products and allows for the expedient synthesis of multicomponent heteromacrocycles, it still does not offer site-specific control over sequence. Thus, inverting synthetic priorities while elevating the importance of sequence remains the best approach for controlling the build up of hierarchical order. The payoff for the synthetic investment comes when a correlation can be made between the primary sequence and the resulting higher levels of structural order (1° → 2° → 3° → 4° → 5°) to control the final architecture.

The hierarchical ordering of molecules can be readily seen at surfaces using scanning tunneling microscopy (STM). The resulting real-space images produced by STM (e.g., Figure 1e) enable the detailed study of the hierarchical assembly of macrocycles,[79–81] but sequence information is infrequently investigated. Sequence is expected to impact molecular ordering at surfaces primarily by altering intermolecular contacts.[82] Packing relies on molecular shape and edge-directed interactions between macrocycles, which are intimately connected to the sequence of noncovalent interactions encoded into the macrocycle.[83] Despite the >2000 studies of molecular assembly at solid-solution interfaces,[84] only a few incorporate sequence information. In one case, the precise geometric substitution of a porphyrin was used to control assemblies formed on gold surfaces.[85] Porphyrins with two nitrile groups displayed in either a linear (180°) or perpendicular (90°) manner form either a line- or square-shaped assemblies, respectively. In another instance, rectangular, tetrasubstituted dehydro[24]annulene macrocycles were shown to adopt unique packing patterns depending on the arrangement of substituents either along the wider or narrower edge of the macrocycle.[86] Sequence programming has been explored in linear peptides to define a 2D pattern that builds up hierarchically into either linear or hexagonal lattices.[87] Thus, and despite these supramolecular assemblies showing a connection between latent sequence information and their surface patterns, the role of sequence in the hierarchical assembly of macrocycles needs to be investigated.

We describe a high-yielding synthetic pathway to sequence-defined macrocycles called tricarbazole triazolophanes (tricarb)[67,88] and the use of sequence to identify complementary driving forces controlling their self-assembly into 2D superstructures. The synthetic pathway

developed herein was used to create four new tricarb macrocycles with A-Tz-A-Tz-A-Tz-, A-Tz-A-Tz-B-Tz-, and A-Tz-B-Tz-C-Tz-type sequences (Figure 3a,b; hereafter abbreviated AAA, AAB, and ABC type). To establish the viability of the method, we first prepared the AAA-type and C_3-symmetric macrocycle **TC-10** (Figure 3c) for comparison to the product of a previously described one-pot reaction.[67] This stepwise pathway also afforded us the opportunity to then showcase the regioselectivity of a thermal Huisgen cycloaddition reaction by using covalent control instead of by supramolecular control seen previously.[89–91] The new stepwise method allows us to prepare sequence-defined macrocycles of lower symmetry comprising either AAB-type (**TC-6618** and **TC-66HEG**, Figure 3c) or ABC-type (**TC-610MeCy**, Figure 3c) structures with uniquely substituted carbazole building blocks (**6** = hexyl, $-C_6H_{13}$; **10** = decyl, $-C_{10}H_{21}$; **18** = octadecyl, $-C_{18}H_{37}$; **HEG** = hexaethylene glycol monomethyl ether, $-(CH_2CH_2O)_6CH_3$; **MeCy** = methylcyclohexyl, $-CH_2C_6H_{11}$). The power and efficiency with which the stepwise pathway can be utilized to encode sequence information within a tricarb macrocycle allowed us to probe the role of sequence on surface assembly. Macrocycle **TC-6618** was observed to form two new surface polymorphs not seen with the totally symmetric **TC-10** macrocycle;[67] a gap phase and a disordered phase. To identify the driving forces underpinning these new phases we leveraged the stepwise synthesis to make AAB-type macrocycle **TC-66HEG** and ABC-type **TC-610MeCy** (Figure 3c). Assembly studies then showed that the newly characterized gap phase arises as a consequence of the information encoded within the substituents of the **TC-6618** macrocycle, the long octadecyl chain, and not as a result of lower molecular symmetry. These studies verify that sequence can control the hierarchical patterning of shape-persistent macrocycles on planar graphite surfaces and that new polymorphs can be produced only when using nonsymmetric compounds. These findings help elevate the role of high-yielding stepwise syntheses in investigating and controlling how sequence information programs the build up of hierarchical structure.

RESULTS AND DISCUSSION

Convergent One-Pot Synthesis of Self-Assembling Macrocycles for Benchmarking.

The C_3-symmetric AAA-type tricarb macrocycles studied in this work, **TC-6** (AAA),[88] **TC-10** (BBB),[67] and **TC-18** (CCC), were synthesized from difunctional building block **7** (Scheme 1) substituted with the corresponding alkyl chains (see **7a-c**). Novel tricarb macrocycle **TC-18** was synthesized following a previously developed seven-step pathway.[88] An octadecyl ($-C_{18}H_{37}$) chain was substituted on the bridgehead nitrogen of carbazole in the first step (**2c**) followed by introduction of the azide and alkyne functional groups at the 3 and 6 positions of carbazole over five steps to give the difunctional octadecyl-building block **7c**. Carbazole **7c** was then subjected to a one-pot copper-catalyzed azide-alkyne cycloaddition (CuAAC) to give tricarb macrocycle **TC-18** in 68% yield. This sequence and the yields are typical of the one-pot pathway.[67,88]

Stepwise Synthesis of Sequence-Defined Tricarb Macrocycles.

The design of a stepwise synthesis of the tricarb macrocycle scaffold relies on the progressive growth of an oligomer in a controlled and high-yielding manner to establish the primary sequence. Examining the possible synthetic pathways as based on the various

preparations of building blocks **1–7** led us to a stepwise sequence that takes advantage of a protection-deprotection scheme to control the growth of the macrocyclic precursors (Scheme 2). In this approach, amine groups and trimethylsilyl (TMS) protected alkyne moieties were employed as precursors to reactive azide and alkyne groups, respectively. The amine-substituted carbazoles (Reaction 4) can be converted easily to azides in excellent yields,[92] while TMS groups (Reactions 2 and 5) can be quantitatively deprotected with mild base. Furthermore, neither the amine nor TMS-protected alkynyl groups were anticipated to undergo side-reactions during the proposed triazole-forming steps (Reactions 1 and 3), making them ideal protecting groups for the efficient coupling of building blocks. Satisfyingly, a carbazole building block functionalized with an amine and TMS-protected alkynyl group displayed no reactivity under the CuAAC conditions employed herein.

The viability of the proposed stepwise synthesis was confirmed using the previously established **TC-10** macrocycle[67] as the target. TMS-protected carbazole **7b-TMS** was made in six steps while amine-substituted carbazole **8b** was made by subjecting **5b** to Sonogashira coupling with TMS-acetylene followed by base deprotection. These two carbazoles (**7b-TMS** and **8b**) were then used to form crescent **dimer-1010** (step I, Scheme 2) using CuAAC conditions followed by base deprotection. This dimer was then subjected to another CuAAC (step II, Scheme 2) with carbazole **7b-TMS** to grow the oligomer and produce the amine-substituted precursor of crescent **trimer-101010**. The amino group of the trimer intermediate was converted to an azide, and the alkyne moiety was activated by base deprotection to give crescent **trimer-101010**. This trimer was closed by a third CuAAC reaction to afford **TC-10** in 80% yield (step III, Scheme 2).

To verify the build up of the oligomer's primary sequence and the higher symmetry upon macrocyclization, we recorded the ^{1}H NMR spectra of all of the key intermediates along the synthetic pathway (Figure 4). We observe the aromatic region becoming more complex at each stage up until the final ring-closing reaction. The addition of each building block is expected, and found, to introduce seven new aromatic resonances: one from the newly formed triazole ring and six peaks from the additional carbazole. Consistently, crescent **dimer-1010** displays 13 aromatic proton resonances (Figure 4b), while crescent **trimer-101010** displays 20 (Figure 4c). This signature then simplifies to seven signals following macrocyclization (Figure 4d). This simplified ^{1}H NMR spectrum is identical to the one for the **TC-10** macrocycle formed by the original one-pot reaction.[67]

We can compare the efficiencies of the two synthetic approaches. The one-pot method is expedient and high-yielding. Starting from carbazole, macrocycle **TC-10** is made from a single building block in just seven total steps with an overall yield of 40%. In comparison, the stepwise pathway requires six steps to synthesize azido building block **7b-TMS** from carbazole (55% yield over six steps), one subsequent step to synthesize alkynyl building block **8b** from compound **5b** (95% yield), two coupling reactions and one deprotection step to reach crescent **trimer-101010** (63% yield over three steps), and a final coupling reaction to give macrocycle **TC-10** in an overall yield of 25% for the 11 total steps. While the overall yield of the stepwise route (25%, 11 steps) is lower than the convergent one-pot synthesis (40%, eight steps), this is the only way to expediently make tricarb macrocycles of any sequence (Figure 2). Thus, the benefit of this approach is sequence control.

Next, we highlight the versatility of the stepwise synthesis by preparing broken symmetry macrocycles. The first target was an AAB-styled macrocycle with two hexyl side chains ($-C_6H_{13}$) and one octadecyl chain ($-C_{18}H_{37}$). This specific combination was selected for the purposes of investigating surface assembly (vide infra). Satisfyingly, tricarb macrocycle **TC-6618** (AAC) was synthesized using our generalized three-step pathway (Scheme 2) from the appropriate building blocks. Hexyl ($-C_6H_{13}$) substituted monomers **7a-TMS** and **8a** were used to form crescent **dimer-66** over two steps with an overall yield of 84%. Crescent **dimer-66** was then combined with octadecyl ($-C_{18}H_{37}$) substituted building block **7c-TMS** to give crescent **trimer-6618** over three steps with an overall yield of 74%. The crescent **trimer-6618** was then closed to give macrocycle **TC-6618** with a 75% yield for this final step (12% overall yield over 20 steps starting from carbazole).

When we evaluated the ^{1}H NMR spectrum of macrocycle **TC-6618** (Figure 5a, black), and despite the fact that this is a lower symmetry macrocycle, it displayed only seven aromatic protons typical of the higher symmetry macrocycles (Figure 3a). This higher apparent symmetry was also present in the aromatic region of the ^{13}C NMR spectrum.[93] Only the alkyl region (10–40 ppm, Figure 5b, black) showed unique resonances arising from both hexyl and octadecyl side chains.

We extended the success of the stepwise synthesis to a second two-component, AAB-type macrocycle. **TC-66HEG** (AAD) was made in four steps totaling 45% yield (6% yield over 20 steps starting from carbazole) using the hexaethylene glycol substituted monomer **7d-TMS** together with crescent **dimer-66**, which was a common intermediate in the preparation of macrocycle **TC-6618**. Much like the AAC macrocycle **TC-6618**, the ^{1}H NMR spectrum of **TC-66HEG** only displayed seven aromatic protons.[93] This finding suggests that the use of simple functional groups on the carbazole building blocks in these AAB-type tricarb macrocycles have little influence on the symmetry of the macrocycle's core. This outcome is consistent with the fact that the carbazole substituents are directed outward from the macrocyclic core.

To accentuate further the versatility of the stepwise synthesis, we prepared the ABC-type macrocycle **TC-610MeCy** (ABE). The methylcyclohexyl ($-CH_2C_6H_{11}$, MeCy)-substituted monomer **7e-TMS** was reacted with hexyl-substituted monomer **8a** to give crescent **dimer-6MeCy** in 84% yield over two steps. Crescent **dimer-6MeCy** and decyl monomer **7b-TMS** were then combined to give crescent **trimer-610MeCy** in a 70% yield over three steps. Crescent **trimer-610MeCy** was then closed to give macrocycle **TC-610MeCy** in 75% yield (5% yield over 24 steps starting from carbazole). Only macrocycle **TC-610MeCy** showed the expected lower symmetry in the ^{1}H NMR spectrum with splitting in the aromatic resonances (Figure 6).

The overall yields of the sequence-controlled tricarb macrocycles (5–25%) are reminiscent of the overall yields obtained in the statistical synthesis of multiply substituted porphyrin macrocycles (<10%).[94–97] Isolation and purification of porphyrins benefits from the variety of visible colors associated with this class of molecule. Consequently, the ability to separate various products from within a complex mixture of similar species (Figure 2a) often offsets the lack of control over the distribution of products. However, the ability to isolate a specific

target does not necessarily guarantee access to any specific target. For tricarb, and many other macrocycles, they are not highly colored. Thus, the controlled and stepwise build up of sequence is the only way to expediently synthesize and isolate AAB and ABC-type tricarb macrocycles with any arbitrary sequence.

Macrocycle Formation by Regioselective Thermal Huisgen.

One of the structural features of the premacrocyclic crescent trimers is their nearby placement (Figure 7b) of the terminal azide and alkyne groups for reaction. The spatial proximity and orientation of these two reactive functionalities suggested to us that macrocyclization could proceed in a regioselective way by a thermal Huisgen cycloaddition. Under thermal conditions, Huisgen cycloaddition reactions typically give a mixture of 1,4- and 1,5-triazole regioisomers (Figure 7a).[98] We hypothesized that the rigidity of the crescent backbone would geometrically inhibit formation of the 1,5-adduct. Satisfyingly, heating the tridecyl-substituted crescent **trimer-101010** at 100 °C for 4 days produced tricarb macrocycle **TC-10** in 65% yield (Figure 7b). Similar regioselective thermal Huisgen cycloadditions reactions have been carried out using different strategies to effect spatial control in the solid state,[99–102] as well as in solution, by using components encapsulated in a resorcinarene capsule[90] or a cucurbit[6]uril.[89,91] However, to the best of our knowledge, this is the first example of a thermal Huisgen reaction to selectively produce the 1,4-triazole adduct as a result of geometric control.

Hierarchical Assembly of Macrocycles on Graphite Surfaces.

Our sequence-defined synthesis provides the ability to control the substituents on the exterior of the tricarb macrocycles. A critical next step in valorizing this stepwise synthesis is to investigate and verify that the sequence of the substituted carbazoles are important for directing the build-up of hierarchical order (Figure 1). Specifically, we investigate how these substituents influence the macrocycle's patterns of 2D assembly into superstructures (4°) and polymorphs (5°) on flat surfaces. We ultimately found that being able to predetermine the substituents gave us insights that were not possible when using only C_3-symmetric macrocycles.

Previous work[67] has shown that the symmetric **TC-10** macrocycle forms two related but different 2D crystalline polymorphs on highly oriented pyrolytic graphite (HOPG): honeycomb (Figure 1e, top) and flower (Figure 1e, bottom) surface morphologies.[67] These morphologies are built up by the side-on packing of macrocycles into rosettes, which constitute a preferred quaternary (4°) structure. The rosettes define a central pore in which are disposed some of the decyl chains from the six constituent macrocycles. Rosette fusion results in the high-density honeycomb phase. An alternative arrangement of this quaternary structure produces the flower phase. Instead of being fused, rosettes are spaced regularly apart as mediated by coadsorbed solvent molecules.[67] The flower phase is present at low concentrations (75 μM) and metastable at higher concentrations where the honeycomb dominates. At room temperature, the initial flower phase is replaced by the honeycomb phase in minutes at 100 μM and days at 5 μM. With the packing relying on the lateral H-bonding between macrocycles (Figure 1c), the three decyl chains of **TC-10** were assumed[67]

to play a subservient role in controlling the hierarchical structure of these two polymorphs. We sought to investigate this assumption by altering the relative lengths of the alkyl chains.

The easiest way to test this idea is to lengthen all alkyl chains in a symmetric manner and examine the distribution of flower and honeycomb phases (Figure 8a–c). To this end, we synthesized (Scheme 1) the AAA-type tricarb macrocycle **TC-18** bearing three long octadecyl ($-C_{18}H_{37}$, Figure 8b). We anticipated that the octadecyl chains would display an increased affinity for the graphite surface leading to their enhanced adsorption relative to the decyl chains of **TC-10**. Consequently, we would expect preferential stabilization of a flower-like phase with an expanded unit cell relative to **TC-10**. Contrary to expectation, when a solution of **TC-18** dissolved in the nonvolatile solvent 1,2,4-trichlorobenzene (TCB) was deposited onto an HOPG surface and subsequently imaged using STM, we saw both the honeycomb and flower phases as a function of concentration (Figure 8d–f). We were also able to observe bright features within the STM images of **TC-18** attributed to the kind of multilayer stacking seen previously with **TC-10**.[93] However, the observation of these stacking events was rare, and so we have elected to focus on 2D self-assembly in this work. At low concentrations, below 75 μM, the flower phase (Figure8d) is observed to be the dominant packing pattern on the surface. At moderate concentrations, between 75 and 100 μM, a mixture of the flower and honeycomb phases is seen (Figure8e). Finally, at high concentrations, above 100 μM, the **TC-18** macrocycle predominantly forms the honeycomb phase (Figure8f). The flower phase formed by **TC-18** was also observed to be metastable with the honeycomb phase at room temperature. At low concentrations, the flower phase is supplanted by the honeycomb phase over the course of hours. At high concentrations, this transformation occurs faster and is completed within minutes. This behavior is identical to the 2D self-assembly of the **TC-10** macrocycle. The inability to influence the relative stability of the flower phase simply by symmetrically lengthening all the alkyl chains motivated designs of sequence-defined tricarb macrocycles.

Sequence-Controlled Macrocycles Direct Emergence of New Surface Superstructures (4°) and Polymorphs (5°).

Our observations that macrocycle **TC-6** only formed the honeycomb phase at concentrations at which macrocycles **TC-10** and **TC-18** were observed to form the flower phase led us to hypothesize that a balance between three factors led to the formation and metastability of the flower phase: the lateral interactions between adsorbed macrocycles; the adsorption affinity of the side chains for the surface; and the steric interactions between side chains sequestered within the rosette pore. As a means to precisely investigate the interactions that direct surface patterning, we designed a two-component macrocycle that displays properties complementary to the packing demands of the flower phase, macrocycle **TC-6618** (AAC, Figure 9a). Macrocycle **TC-6618** employs an octadecyl ($-C_{18}H_{37}$) chain to drive surface interactions that are stronger than decyl chains ($-C_{10}H_{21}$), while hexyl ($-C_6H_{13}$) chains were chosen to maximize the disparity in surface adsorption strength relative to the octadecyl chain, to minimize, while producing a macrocycle where the total number of carbons contained within the sum of alkyl chains was equivalent to **TC-10**. Given the predispositions of these constituent alkyl chains, we expected the flower polymorph to be preferred: we anticipated the long surface-adsorbing octadecyl chains would be directed

outward from the rosettes to disfavor the honeycomb. The smaller hexyl chains were expected to play secondary structure-directing roles and would either be adsorbed inside the rosette pores or backfolded into solution to ultimately favor the flower phase.

We were surprised to observe two new phases and a total of three phases from tricarb macrocycle **TC-6618** (Figure 9). At low concentrations (2.5 μM), a new phase emerges that displays a darker line in the STM imaging that has the appearance of a gap (Figure 9b) in the more typical 2D surface packing of the macrocycles. At 25 μM, domains of honeycomb appear to coexist with the gap phase. This honeycomb has an identical unit cell ($a = b = 2.9 \pm 0.1$ nm) to the honeycomb phase of **TC-10**.[67] As concentration is further increased, the honeycomb's coverage grows, and above 200 μM, a gap is not observed on the surface (Figure9c). At high concentrations (≥200 μM), a new state (Figure9d) with high density and reduced order is observed to coexist with the honeycomb. The gap phase formed by the **TC-6618** macrocycle was also found to be metastable with respect to the honeycomb phase, with the transition from gap to honeycomb slowly occurring over the course of 24 h from the initial scan at room temperature.

The gap polymorph present at low concentrations of the **TC-6618** macrocycle is characterized by STM imaging (Figure 10) to have two types of macrocycle-bearing rows. The imaging (Figure 10a) and the corresponding model (Figure 10c) show honeycomb-like rows of tricarb macrocycles that propagate in one dimension and which are separated by low contrast features, the "gap". The honeycomb-like (Figure 10c, purple) rows retain the privileged quaternary structure of six-macrocycle rosettes, but unlike the honeycomb phase, these rosettes are fused with two neighbors instead of three. This number of intermolecular contacts in the honeycomb rows is between that of the honeycomb (all macrocycles have three contacts) and flower phase (two contacts). This phase also displays high-contrast zigzag rows (Figure 10c, green). Along these zigzag rows, the macrocycles do not form rosettes. Instead, each macrocycle only makes lateral contacts with two neighbors indicative of a new quaternary structure. Transitions between zigzag and honeycomb rows are also seen as line defects (Figure 10c, black dot) within the 2D lattice.

Molecular modeling of the possible packing structures that could account for the darker low-contrast gap region suggests that one or two of the long octadecyl chains are adsorbed within the interstitial space. In addition, the shorter hexyl chains of the macrocycle are believed to be backfolded into solution or adsorbed to the surface either inside the rosette pores or into the gap spaces. Based on these ideas, the emergence of a new self-assembled structure from the **TC-6618** macrocycle suggests that the length of the alkyl chains in sequence-controlled macrocycles can be tuned to play a more distinct role in surface self-assembly on graphite. Alternatively, the new gap phase may instead arise simply as a consequence of the lowered symmetry.

To investigate the role of molecular symmetry on the 2D packing of tricarb macrocycles and to highlight the utility of the sequence-defined synthesis, a three-component ABC macrocycle was designed. Three relatively low surface-affinity substituents were selected as a means to differentiate the role that symmetry plays in directing the 2D surface-assembly of tricarb macrocycles from the effect of long surface-adsorbing alkyl chains. A

methylcyclohexyl group was chosen for its weak adsorption to graphite,[103] while hexyl and decyl chains were chosen because their respective C_3-symmetric macrocycles (**TC-6** and **TC-10**) form packing phases in which the alkyl chains do not play a significant role in self-assembly: As noted above, macrocycle **TC-10** only forms flower and honeycomb phases,[67] while **TC-6** forms honeycomb and a nonordered phase.[93] The observation that both **TC-6** and **TC-6618** form the nonordered packing phases suggests that this phase originates from the low affinity hexyl chains and not as a result of decreased macrocycle symmetry.

STM imaging of the sequence-defined **TC-610MeCy** macrocycle at the solution-graphite interface revealed the formation of only two packing polymorphs: honeycomb and, not surprisingly given the poor-surface interacting hexyl and methyl cyclohexyl substituents, the nonordered state (Figure 11). The gap phase was not observed. At the lower concentrations (≤2.5 μM) needed to form the gap phase, we instead see the **TC-610MeCy** macrocycle forming the honeycomb phase. At 100 μM, domains of the nonordered phase coexist with honeycomb. As concentration is increased further to 1 mM, the surface coverage of the nonordered phase grows such that it is the only one observed. The absence of a gap phase suggests that its emergence in **TC-6618** arises from driving forces beyond the nonsymmetric display of functional groups. Rather, the surface affinity and local interactions involving the long octadecyl chain are also key for accessing this packing morphology. Additionally, the nonordered phase is a kinetically trapped state and can be converted to the honeycomb phase by solvent annealing with toluene.[93] We acknowledge that a more detailed investigation of the factors that give rise to the nonordered packing phase is warranted and will be further discussed in a later work.

To provide an independent test of the hypothesis that surface interactions encoded by the long octadecyl chain direct the gap phase, we designed macrocycle **TC-66HEG** (Figure 12a) with an isosteric hexaethylene glycol (HEG) chain in place of the octadecyl chain. The chains are similar in length (19 and 18 heavy atoms, respectively); however, glycol chains encode for minimal graphite surface affinity.[104] We anticipated that the low surface-affinity chain would stop formation of the gap phase at low concentrations. Gratifyingly, the gap phase is not observed when solutions of **TC-66HEG** are deposited onto HOPG. Instead, macrocycle **TC-66HEG** only forms the honeycomb phase (Figure 12b) at those concentrations where **TC-6618** forms the gap phase (2.5 to 50 μM). This observation is consistent with the idea that the surface interactions of the octadecyl chains lead to their adsorption into the gap region (Figure 12c). By contrast the hexaethylene glycol of **TC-66HEG** is likely backfolded (Figure 12d) for honeycomb formation.

CONCLUSIONS

We have shown that the primary sequence (1°) of sequence-controlled macrocycles encodes for the hierarchical build-up of ordering on graphite surfaces. This control is expressed at the 4° level of organization in the many-molecule superstructures (rosettes and zigzag rows) and at 5° levels of the surface polymorphs (flower-honeycomb, gap-honeycomb-disorder, honeycomb-disorder). This work was investigated with tricarb macrocycles bearing either one (AAA **TC-6**, BBB **TC-10**, CCC **TC-18**), two (AAC **TC-6618**, AAD **TC-66HEG**), or three (ABE **TC-610 MeCy**) different substituents made in 11–25 total steps. The design of

these macrocycles required the development of a stepwise synthesis, thus inverting the typical synthetic predilection for one-pot macrocyclizations. Along the way, the final macrocycle-forming reaction of the stepwise route was carried out regioselectively for the first time in a geometrically controlled Huisgen cycloaddition under thermal conditions without intermolecular ordering. We found that the AAC tricarb macrocycle with one octadecyl and two hexyl chains assembled into the new gap phase. Only by making an AAD and ABE macrocycle were we able to show that the gap phase arises from the local interactions generated upon the adsorption of the octadecyl chain and not from the chain's bulk or the macrocycle's lowered symmetry. We note that, in some cases, significant changes to the peripheral functional groups do not have a noticeable impact on the packing structure. For example, **TC-10** and **TC-18** produce similar packing structures, in spite of nearly doubling the peripheral alkyl length. It seems that, in this and other cases, the adjustments to the peripheral units do not shift the balance away from other intermolecular interactions. This study provides the framework necessary to shift the perceived importance of synthesis from one-pot macrocyclizations to stepwise methods to enable sequence to control the build up of hierarchical structure of macrocycles. Ultimately, this shift in perspective offers the potential for greater control and a deeper understanding of hierarchical assembly.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

J.R.D., Y.L., S.L.T., and A.H.F. thank the National Science Foundation (DMR 1533988) for support; H.D.C. thanks Indiana University and the Arts and Sciences President's Diversity Initiative Fellowship for support; R.E.F., R.D.T., A.A.B., and C.Q.T. thank the National Science Foundation (CHE 1460720) for support through Indiana University's Research Experiences for Undergraduates (REU) in Materials Chemistry, Nanoscience, and Assembly.

REFERENCES

(1). Anfinsen CB The formation and stabilization of protein structure. Biochem. J 1972, 128, 737–749. [PubMed: 4565129]

(2). Anfinsen CB Principles that Govern the Folding of Protein Chains. Science 1973, 181, 223–230. [PubMed: 4124164]

(3). Cusick ME; Klitgord N; Vidal M; Hill DE Interactome: gateway into systems biology. Hum. Mol. Genet 2005, 14, R171–R181. [PubMed: 16162640]

(4). De Las Rivas J; Fontanillo C Protein-Protein Interactions Essentials: Key Concepts to Building and Analyzing Interactome Networks. PLoS Comput. Biol 2010, 6, e1000807. [PubMed: 20589078]

(5). Wang B; Xie Z-R; Chen J; Wu Y Integrating Structural Information to Study the Dynamics of Protein-Protein Interactions in Cells. Structure 2018, 26, 1414–1424. [PubMed: 30174150]

(6). Douglas SM; Marblestone AH; Teerapittayanon S; Vazquez A; Church GM; Shih WM Rapid prototyping of 3D DNA-origami shapes with caDNAno. Nucleic Acids Res. 2009, 37, 5001–5006. [PubMed: 19531737]

(7). Castro CE; Kilchherr F; Kim D-N; Shiao EL; Wauer T; Wortmann P; Bathe M; Dietz H A primer to scaffolded DNA origami. Nat. Methods 2011, 8, 221. [PubMed: 21358626]

(8). Han D; Pal S; Nangreave J; Deng Z; Liu Y; Yan H DNA Origami with Complex Curvatures in Three-Dimensional Space. Science 2011, 332, 342–346. [PubMed: 21493857]

(9). Lupas A; Van Dyke M; Stock J Predicting Coiled Coils from Protein Sequences. Science 1991, 252, 1162–1164. [PubMed: 2031185]

(10). Jones DT Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol 1999, 292, 195–202. [PubMed: 10493868]

(11). Krissinel E; Henrick K Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr., Sect. D: Biol. Crystallogr 2004, 60, 2256–2268. [PubMed: 15572779]

(12). Ogura Y; Artar M; Palmans ARA; Sawamoto M; Meijer EW; Terashima T Self-Assembly of Hydrogen-Bonding Gradient Copolymers: Sequence Control via Tandem Living Radical Polymerization with Transesterification. Macromolecules 2017, 50, 3215–3223.

(13). Pfeifer S; Zarafshani Z; Badi N; Lutz J-F Liquid-Phase Synthesis of Block Copolymers Containing Sequence-Ordered Segments. J. Am. Chem. Soc 2009, 131, 9195–9197. [PubMed: 19522508]

(14). Barnes JC; Ehrlich DJC; Gao AX; Leibfarth FA; Jiang Y; Zhou E; Jamison TF; Johnson JA Iterative exponential growth of stereo- and sequence-controlled polymers. Nat. Chem 2015, 7, 810. [PubMed: 26391080]

(15). Golder MR; Jiang Y; Teichen PE; Nguyen HVT; Wang W; Milos N; Freedman SA; Willard AP; Johnson JA Stereochemical Sequence Dictates Unimolecular Diblock Copolymer Assembly. J. Am. Chem. Soc 2018, 140, 1596–1599. [PubMed: 29356516]

(16). Hosono N; Stals PJM; Palmans ARA; Meijer EW Consequences of Block Sequence on the Orthogonal Folding of Triblock Copolymers. Chem. - Asian J 2014, 9, 1099–1107. [PubMed: 24678056]

(17). Lewandowski B; De Bo G; Ward JW; Papmeyer M; Kuschel S; Aldegunde MJ; Gramlich PME; Heckmann D; Goldup SM; D'Souza DM; Fernandes AE; Leigh DA Sequence-Specific Peptide Synthesis by an Artificial Small-Molecule Machine. Science 2013, 339, 189–193. [PubMed: 23307739]

(18). Oh D; Ouchi M; Nakanishi T; Ono H; Sawamoto M Iterative Radical Addition with a Special Monomer Carrying Bulky and Convertible Pendant: A New Concept toward Controlling the Sequence for Vinyl Polymers. ACS Macro Lett. 2016, 5, 745–749.

(19). Hibi Y; Ouchi M; Sawamoto M A strategy for sequence control in vinyl polymers via iterative controlled radical cyclization. Nat. Commun 2016, 7, 11064. [PubMed: 26996881]

(20). Lautrette G; Wicher B; Kauffmann B; Ferrand Y; Huc I Iterative Evolution of an Abiotic Foldamer Sequence for the Recognition of Guest Molecules with Atomic Precision. J. Am. Chem. Soc 2016, 138, 10314–10322. [PubMed: 27428616]

(21). Vallade M; Sai Reddy P; Fischer L; Huc I Enhancing Aromatic Foldamer Helix Dynamics to Probe Interactions with Protein Surfaces. Eur. J. Org. Chem 2018, 2018, 5489–5498.

(22). Liu Y; Parks FC; Zhao W; Flood AH Sequence-Controlled Stimuli-Responsive Single-Double Helix Conversion between 1:1 and 2:2 Chloride-Foldamer Complexes. J. Am. Chem. Soc 2018, 140, 15477–15486. [PubMed: 30346756]

(23). Saha S; Kauffmann B; Ferrand Y; Huc I Selective Encapsulation of Disaccharide Xylobiose by an Aromatic Foldamer Helical Capsule. Angew. Chem., Int. Ed 2018, 57, 13542–13546.

(24). Tsiamantas C; Kwon S; Douat C; Huc I; Suga H Optimizing Aromatic Oligoamide Foldamer Side-Chains for Ribosomal Translation Initiation. Chem. Commun 2019, 55, 7366–7369.

(25). Bao C; Gan Q; Kauffmann B; Jiang H; Huc I A Self-Assembled Foldamer Capsule: Combining Single and Double Helical Segments in One Aromatic Amide Sequence. Chem. - Eur. J 2009, 15, 11530–11536. [PubMed: 19768708]

(26). Goto K; Moore JS Sequence-Specific Binding of m-Phenylene Ethynylene Foldamers to a Piperazinium Dihydrochloride Salt. Org. Lett 2005, 7, 1683–1686. [PubMed: 15844880]

(27). Lehnherr D; Chen C; Pedramrazi Z; DeBlase CR; Alzola JM; Keresztes I; Lobkovsky EB; Crommie MF; Dichtel WR Sequence-defined oligo(ortho-arylene) foldamers derived from the benzannulation of ortho(arylene ethynylene)s. Chem. Sci 2016, 7, 6357–6364. [PubMed: 28567248]

(28). Endo M; Sugita T; Katsuda Y; Hidaka K; Sugiyama H Programmed-Assembly System Using DNA Jigsaw Pieces. Chem. - Eur. J 2010, 16, 5362–5368. [PubMed: 20391568]

(29). Hamblin GD; Rahbani JF; Sleiman HF Sequential growth of long DNA strands with user-defined patterns for nanostructures and scaffolds. Nat. Commun 2015, 6, 7065. [PubMed: 25940750]

(30). Lau KL; Sleiman HF Minimalist Approach to Complexity: Templating the Assembly of DNA Tile Structures with Sequentially Grown Input Strands. ACS Nano 2016, 10, 6542–6551. [PubMed: 27303951]

(31). Chidchob P; Edwardson TGW; Serpell CJ; Sleiman HF Synergy of Two Assembly Languages in DNA Nanostructures: Self-Assembly of Sequence-Defined Polymers on DNA Cages. J. Am. Chem. Soc 2016, 138, 4416–4425. [PubMed: 26998893]

(32). Praetorius F; Dietz H Self-assembly of genetically encoded DNA-protein hybrid nanoscale shapes. Science 2017, 355, eaam5488. [PubMed: 28336611]

(33). Gerling T; Kube M; Kick B; Dietz H Sequence-programmable covalent bonding of designed DNA assemblies. Sci. Adv 2018, 4, eaau1157–eaau1157. [PubMed: 30128357]

(34). Koga N; Tatsumi-Koga R; Liu G; Xiao R; Acton TB; Montelione GT; Baker D Principles for designing ideal protein structures. Nature 2012, 491, 222. [PubMed: 23135467]

(35). Currin A; Swainston N; Day PJ; Kell DB Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently. Chem. Soc. Rev 2015, 44, 1172–1239. [PubMed: 25503938]

(36). Huang P-S; Boyken SE; Baker D The coming of age of de novo protein design. Nature 2016, 537, 320. [PubMed: 27629638]

(37). Hosseinzadeh P; Bhardwaj G; Mulligan VK; Shortridge MD; Craven TW; Pardo-Avila F; Rettie SA; Kim DE; Silva D-A; Ibrahim YM; Webb IK; Cort JR; Adkins JN; Varani G; Baker D Comprehensive computational design of ordered peptide macrocycles. Science 2017, 358, 1461–1466. [PubMed: 29242347]

(38). Frederix PWJM; Scott GG; Abul-Haija YM; Kalafatovic D; Pappas CG; Javid N; Hunt NT; Ulijn RV; Tuttle T Exploring the sequence space for (tri-)peptide self-assembly to design and discover new hydrogels. Nat. Chem 2015, 7, 30. [PubMed: 25515887]

(39). Scott GG; McKnight PJ; Tuttle T; Ulijn RV Tripeptide Emulsifiers. Adv. Mater 2016, 28, 1381–1386. [PubMed: 26639675]

(40). Lampel A; McPhee SA; Park H-A; Scott GG; Humagain S; Hekstra DR; Yoo B; Frederix PWJM; Li T-D; Abzalimov RR; Greenbaum SG; Tuttle T; Hu C; Bettinger CJ; Ulijn RV Polymeric peptide pigments with sequence-encoded properties. Science 2017, 356, 1064–1068. [PubMed: 28596363]

(41). Schneider JA; Craven TW; Kasper AC; Yun C; Haugbro M; Briggs EM; Svetlov V; Nudler E; Knaut H; Bonneau R; Garabedian MJ; Kirshenbaum K; Logan SK Design of Peptoid-peptide Macrocycles to Inhibit the β-catenin TCF Interaction in Prostate Cancer. Nat. Commun 2018, 9, 4396. [PubMed: 30352998]

(42). We learned of an approach to make sequence-defined cavitands at the time of galley proof review, see:Meisel JW; Hu CT; Hamilton AD Heterofunctionalized Cavitands by Macrocyclization of Sequence-Defined Foldamers. Org. Lett 2019, 21, 7763–7767. [PubMed: 31535554]

(43). Woodward RB; Ayer WA; Beaton JM; Bickelhaupt F; Bonnett R; Buchschacher P; Closs GL; Dutler H; Hannah J; Hauck FP; Itô S; Langemann A; Le Goff E; Leimgruber W; Lwowski W; Sauer J; Valenta Z; Volz H The Total Synthesis of Chlorophyll. J. Am. Chem. Soc 1960, 82, 3800–3802.

(44). Woodward RB; Ayer WA; Beaton JM; Bickelhaupt F; Bonnett R; Buchschacher P; Closs GL; Dutler H; Hannah J; Hauck FP; Itô S; Langemann A; Le Goff E; Leimgruber W; Lwowski W; Sauer J; Valenta Z; Volz H The total synthesis of chlorophyll a. Tetrahedron 1990, 46, 7599–7659.

(45). Flídrová K; Tkadlecová M; Lang K; Lhoták P Anion recognition by calix[4]arene-based p-nitrophenyl amides. Tetrahedron Lett. 2012, 53, 678–680.

(46). Flídrová K; Tkadlecová M; Lang K; Lhoták P Anion complexation by calix[4]arene-TTF conjugates. Dyes Pigm. 2012, 92, 668–673.

(47). Gaeta C; Talotta C; Della Sala P; Margarucci L; Casapullo A; Neri P Anion-Induced Dimerization in p-Squaramidocalix[4]-arene Derivatives. J. Org. Chem 2014, 79, 3704–3708. [PubMed: 24645825]

(48). Li Z; Xing H; Huang G; Sun X; Jiang J; Wang L Novel supramolecular organocatalysts of hydroxyprolinamide based on calix[4]arene scaffold for the enantioselective Biginelli reaction. Sci. China: Chem 2011, 54, 1726–1734.

(49). Baldini L; Cacciapaglia R; Casnati A; Mandolini L; Salvio R; Sansone F; Ungaro R Upper Rim Guanidinocalix[4]arenes as Artificial Phosphodiesterases. J. Org. Chem 2012, 77, 3381–3389. [PubMed: 22364173]

(50). Salvio R; Volpi S; Cacciapaglia R; Sansone F; Mandolini L; Casnati A Upper Rim Bifunctional cone-Calix[4]arenes Based on a Ligated Metal Ion and a Guanidinium Unit as DNAase and RNAase Mimics. J. Org. Chem 2016, 81, 4728–4735. [PubMed: 27135962]

(51). Hüggenberg W; Seper A; Oppel IM; Dyker G Multifold Photocyclization Reactions of Styrylcalix[4]arenes. Eur. J. Org. Chem 2010, 2010, 6786–6797.

(52). Sessler JL; Davis JM Sapphyrins: Versatile Anion Binding Agents. Acc. Chem. Res 2001, 34, 989–997. [PubMed: 11747417]

(53). Sessler JL; Seidel D Synthetic Expanded Porphyrin Chemistry. Angew. Chem., Int. Ed 2003, 42, 5134–5175.

(54). Zhu X-J; Fu S-T; Wong W-K; Guo J-P; Wong W-Y A Near-Infrared-Fluorescent Chemodosimeter for Mercuric Ion Based on an Expanded Porphyrin. Angew. Chem., Int. Ed 2006, 45, 3150–3154.

(55). Misra R; Chandrashekar TK Structural Diversity in Expanded Porphyrins. Acc. Chem. Res 2008, 41, 265–279. [PubMed: 18281947]

(56). Saito S; Osuka A Expanded Porphyrins: Intriguing Structures, Electronic Properties, and Reactivities. Angew. Chem., Int. Ed 2011, 50, 4342–4373.

(57). Sarma T; Kim G; Sen S; Cha W-Y; Duan Z; Moore MD; Lynch VM; Zhang Z; Kim D; Sessler JL Proton-Coupled Redox Switching in an Annulated π-Extended Core-Modified Octaphyrin. J. Am. Chem. Soc 2018, 140, 12111–12119. [PubMed: 30180553]

(58). Lash TD; Fosu SC; Smolczyk TJ; AbuSalim DI Synthesis of Expanded Porphyrinoids with Azulene and Indene Subunits and an opp-Dioxadicarbaporphyrin from Fulvene Carbinols and a Dioxacarbatripyrrin. J. Org. Chem 2018, 83, 12619–12631. [PubMed: 30247904]

(59). Kishore MVN; Panda PK Revisiting the intense NIR active bronzaphyrin, a 26-π aromatic expanded porphyrin: synthesis and structural analysis. Chem. Commun 2018, 54, 13135–13138.

(60). Wei X; Zhang G; Shen Y; Zhong Y; Liu R; Yang N; Almkhaizim FY; Kline MA; He L; Li M; Lu Z-L; Shao Z; Gong B Persistent Organic Nanopores Amenable to Structural and Functional Tuning. J. Am. Chem. Soc 2016, 138, 2749–2754. [PubMed: 26877246]

(61). Choi K; Hamilton AD A Dual Channel Fluorescence Chemosensor for Anions Involving Intermolecular Excited State Proton Transfer. Angew. Chem., Int. Ed 2001, 40, 3912–3915.

(62). Leigh DA; Wong JKY; Dehez F; Zerbetto F Unidirectional rotation in a mechanically interlocked molecular rotor. Nature 2003, 424, 174. [PubMed: 12853952]

(63). Zhang W; Moore JS Shape-Persistent Macrocycles: Structures and Synthetic Approaches from Arylene and Ethynylene Building Blocks. Angew. Chem., Int. Ed 2006, 45, 4416–4439.

(64). Li Y; Pink M; Karty JA; Flood AH Dipole-Promoted and Size-Dependent Cooperativity between Pyridyl-Containing Triazolophanes and Halides Leads to Persistent Sandwich Complexes with Iodide. J. Am. Chem. Soc 2008, 130, 17293–17295. [PubMed: 19053220]

(65). Yang Y; Feng W; Hu J; Zou S; Gao R; Yamato K; Kline M; Cai Z; Gao Y; Wang Y; Li Y; Yang Y; Yuan L; Zeng XC; Gong B Strong Aggregation and Directional Assembly of Aromatic Oligoamide Macrocycles. J. Am. Chem. Soc 2011, 133, 18590–18593. [PubMed: 22023016]

(66). Gong B; Shao Z Self-Assembling Organic Nanotubes with Precisely Defined, Sub-nanometer Pores: Formation and Mass Transport Characteristics. Acc. Chem. Res 2013, 46, 2856–2866. [PubMed: 23597055]

(67). Lee S; Hirsch BE; Liu Y; Dobscha JR; Burke DW; Tait SL; Flood AH Multifunctional Tricarbazolo Triazolophane Macrocycles: One-Pot Preparation, Anion Binding, and Hierarchical Self-Organization of Multilayers. Chem. - Eur. J 2016, 22, 560–569. [PubMed: 26593327]

(68). Fatila EM; Twum EB; Karty JA; Flood AH Ion Pairing and Co-facial Stacking Drive High-Fidelity Bisulfate Assembly with Cyanostar Macrocyclic Hosts. Chem. - Eur. J 2017, 23, 10652–10662. [PubMed: 28568775]

(69). Zhang W; Moore JS Arylene Ethynylene Macrocycles Prepared by Precipitation-Driven Alkyne Metathesis. J. Am. Chem. Soc 2004, 126, 12796–12796. [PubMed: 15469273]

(70). Qin B; Ong WQ; Ye R; Du Z; Chen X; Yan Y; Zhang K; Su H; Zeng H Highly selective one-pot synthesis of H-bonded pentagon-shaped circular aromatic pentamers. Chem. Commun 2011, 47, 5419–5421.

(71). Storz C; Badoux M; Hauke CM;Šolomek T; Kühnle A; Bally T; Kilbinger AFM One-Pot Synthesis and AFM Imaging of a Triangular Aramide Macrocycle. J. Am. Chem. Soc 2014, 136, 12832–12835. [PubMed: 25178114]

(72). Zhu D-J; Ding W; Wang D-H; Xue M; Yang Y One-pot, highly efficient, cavity controllable synthesis and binding properties of carbazole-based macrocycles with sulfonamide linkages. Org. Chem. Front 2018, 5, 2345–2348.

(73). Jin Y; Zhang A; Huang Y; Zhang W Shape-persistent arylenevinylene macrocycles (AVMs) prepared via acyclic diene metathesis macrocyclization (ADMAC). Chem. Commun 2009, 46, 8258–8260.

(74). Lindsey JS The Porphyrin Handbook; Academic Press: San Diego, 2000; Vol. 1.

(75). Senge MO; Shaker YM; Pintea M; Ryppa C; Hatscher SS; Ryan A; Sergeeva Y Synthesis of meso-Substituted ABCD-Type Porphyrins by Functionalization Reactions. Eur. J. Org. Chem 2010, 2010, 237–258.

(76). Senge MO Stirring the porphyrin alphabet soup-functionalization reactions for porphyrins. Chem. Commun 2011, 47, 1943–1960.

(77). Okochi KD; Jin Y; Zhang W Highly efficient one-pot synthesis of hetero-sequenced shape-persistent macrocycles through orthogonal dynamic covalent chemistry (ODCC). Chem. Commun 2013, 49, 4418–4420.

(78). Okochi KD; Han GS; Aldridge IM; Liu Y; Zhang W Covalent Assembly of Heterosequenced Macrocycles and Molecular Cages through Orthogonal Dynamic Covalent Chemistry (ODCC). Org. Lett 2013, 15, 4296–4299. [PubMed: 23947530]

(79). De Feyter S; De Schryver FC Self-Assembly at the Liquid/Solid Interface: STM Reveals. J. Phys. Chem. B 2005, 109, 4290–4302. [PubMed: 16851494]

(80). Yang Y; Wang C Hierarchical construction of self-assembled low-dimensional molecular architectures observed by using scanning tunneling microscopy. Chem. Soc. Rev 2009, 38, 2576–2589. [PubMed: 19690738]

(81). Iritani K; Tahara K; De Feyter S; Tobe Y Host-Guest Chemistry in Integrated Porous Space Formed by Molecular Self-Assembly at Liquid-Solid Interfaces. Langmuir 2017, 33, 4601–4618. [PubMed: 28206764]

(82). Goronzy DP; Ebrahimi M; Rosei F; Arramel; Fang; De Feyter S; Tait SL; Wang C; Beton PH; Wee ATS; Weiss PS; Perepichka DF Supramolecular Assemblies on Surfaces: Nanopatterning, Functionality, and Reactivity. ACS Nano 2018, 12, 7445–7481. [PubMed: 30010321]

(83). Mali KS; Adisoejoso J; Ghijsens E; De Cat I; De Feyter S Exploring the Complexity of Supramolecular Interactions for Patterning at the Liquid-Solid Interface. Acc. Chem. Res 2012, 45, 1309–1320. [PubMed: 22612471]

(84). (a) 2,259 results were returned from a Web of Science topic search: “2D assembly on surfaces” (June 5, 2019). (b) Sixty papers detail the assembly of porphyrins, 29 study non-porphyrin macrocycles, and only five of the non-porphyrin papers study sequence information.

(85). Yokoyama T; Yokoyama S; Kamikado T; Okuno Y; Mashiko S Selective assembly on a surface of supramolecular aggregates with controlled size and shape. Nature 2001, 413, 619. [PubMed: 11675782]

(86). Suzuki M; Guo Z; Tahara K; Kotyk JFK; Nguyen H; Gotoda J; Iritani K; Rubin Y; Tobe Y Self-Assembled Dehydro[24]annulene Monolayers at the Liquid/Solid Interface: Toward On-Surface Synthesis of Tubular π-Conjugated Nanowires. Langmuir 2016, 32, 5532–5541. [PubMed: 27183003]

(87). Abb S; Harnau L; Gutzler R; Rauschenbach S; Kern K Two-dimensional honeycomb network through sequence-controlled self-assembly of oligopeptides. Nat. Commun 2016, 7, 10335. [PubMed: 26755352]

(88). Dobscha JR; Debnath S; Fadler RE; Fatila EM; Pink M; Raghavachari K; Flood AH Host-Host Interactions Control Self-assembly and Switching of Triple and Double Decker Stacks of Tricarbazole Macrocycles Co-assembled with anti-Electrostatic Bisulfate Dimers. Chem. - Eur. J 2018, 24, 9841–9852. [PubMed: 29665108]

(89). Mock WL; Irra TA; Wepsiec JP; Adhya M Catalysis by cucurbituril. The significance of bound-substrate destabilization for induced triazole formation. J. Org. Chem 1989, 54, 5302–5308.

(90). Chen J; Rebek J Selectivity in an Encapsulated Cycloaddition Reaction. Org. Lett 2002, 4, 327–329. [PubMed: 11820871]

(91). Ke C; Strutt NL; Li H; Hou X; Hartlieb KJ; McGonigal PR; Ma Z; Iehl J; Stern CL; Cheng C; Zhu Z; Vermeulen NA; Meade TJ; Botros YY; Stoddart JF Pillar[5]arene as a Co-Factor in Templating Rotaxane Formation. J. Am. Chem. Soc 2013, 135, 17019–17030. [PubMed: 24059594]

(92). Kutonova KV; Trusova ME; Postnikov PS; Filimonov VD; Parello J A Simple and Effective Synthesis of Aryl Azides via Arenediazonium Tosylates. Synthesis 2013, 45, 2706–2710.

(93). See the Supporting Information.

(94). Little RG; Anton JA; Loach PA; Ibers JA The synthesis of some substituted tetraarylporphyrins. J. Heterocycl. Chem 1975, 12, 343–349.

(95). Little RG The mixed-aldehyde synthesis of difunctional tetraarylporphyrins. J. Heterocycl. Chem 1981, 18, 129–133.

(96). Hwang KC; Mauzerall D; Wagner RW; Lindsey JS Synthesis Of Amphipathic Porphyrins And Their Photoinduced Electron Transfer Reactions At The Lpid Bilayer-Water Interface. Photochem. Photobiol 1994, 59, 145–151. [PubMed: 8165233]

(97). Rao PD; Dhanalekshmi S; Littler BJ; Lindsey JS Rational Syntheses of Porphyrins Bearing up to Four Different Meso Substituents. J. Org. Chem 2000, 65, 7323–7344. [PubMed: 11076589]

(98). Huisgen R Proc. Chem. Soc 1961, 357–396.

(99). Ni B-B; Wang C; Wu H; Pei J; Ma Y Copper-free cycloaddition of azide and alkyne in crystalline state facilitated by arene-perfluoroarene interactions. Chem. Commun 2010, 46, 782–784.

(100). Chen H; Ni B-B; Gao F; Ma Y Pressure-accelerated copper-free cycloaddition of azide and alkyne groups pre-organized in the crystalline state at room temperature. Green Chem. 2012, 14, 2703–2705.

(101). Ni B-B; Wang K; Yan Q; Chen H; Ma Y; Zou B Pressure accelerated 1,3-dipolar cycloaddition of azide and alkyne groups in crystals. Chem. Commun 2013, 49, 10130–10132.

(102). Meng X; Chen H; Xu S; Ma Y Metal-free 1,3-dipolar cycloaddition polymerization via prearrangement of azide and alkyne in the solid state. CrystEngComm 2014, 16, 9983–9986.

(103). Otyepková E; Lazar P;Čépe K; Tomanec O; Otyepka M Organic adsorbates have higher affinities to fluorographene than to graphene. Appl. Mater. Today 2016, 5, 142–149.

(104). Kim S; Castillo HD; Lee M; Mortensen RD; Tait SL; Lee D From Foldable Open Chains to Shape-Persistent Macrocycles: Synthesis, Impact on 2D Ordering, and Stimulated Self-Assembly. J. Am. Chem. Soc 2018, 140, 4726–4735. [PubMed: 29534561]

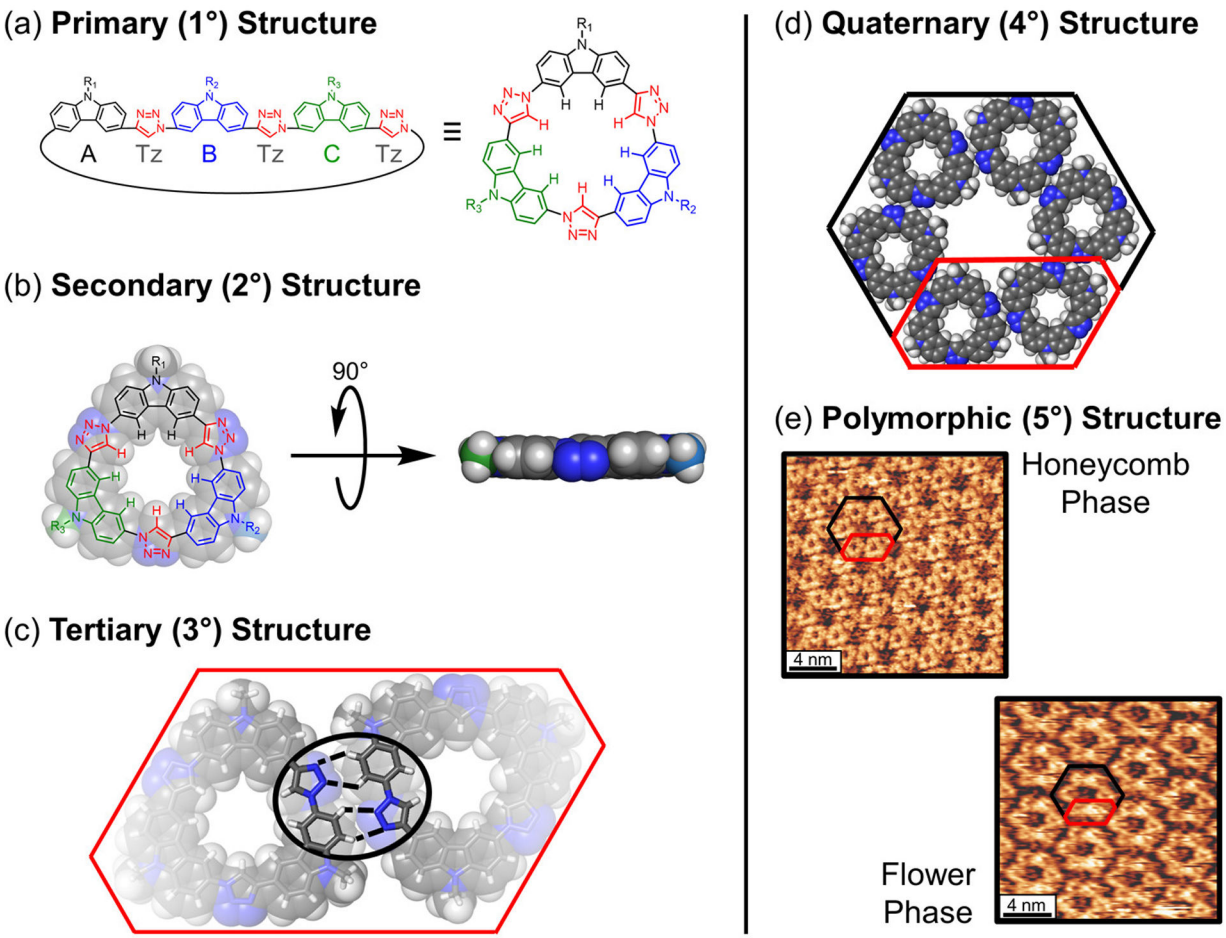


Figure 1.

(a) Primary (1°) structure of a tricarb macrocycle in which A, B, and C are three carbazole monomers linked by triazole (Tz) units. (b) The tricarb macrocycle adopts a planar secondary (2°) structure. (c) The tertiary (3°) structure adopted on graphite surfaces involves macrocycle dimers stabilized by lateral H-bonding between CH donors and N atom acceptors, forming a H-bonding array at the intermacrocycle seam. (d) Propagation of side-on interactions between macrocycles leads to the generation of quaternary (4°) superstructures on graphite surfaces; 6-membered rosettes are preferred. (e) Patterning of rosettes across the surface results in the formation of crystalline 2D polymorphs.

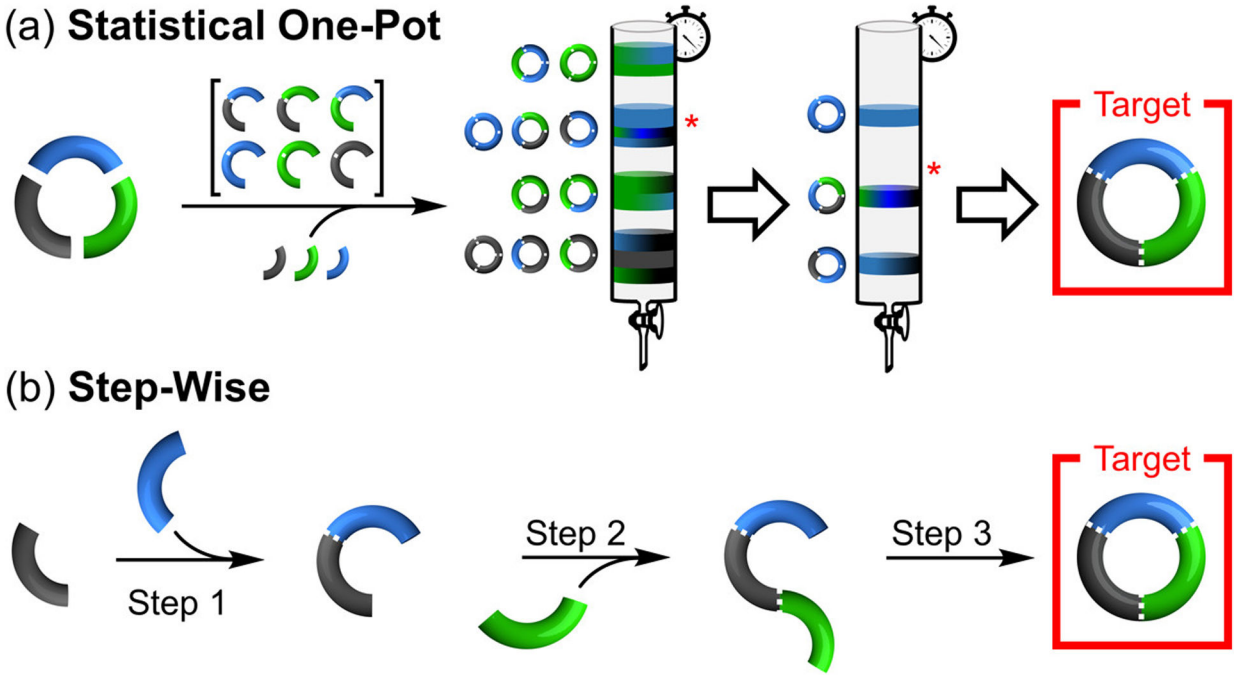


Figure 2.

Syntheses of a three-component macrocycle by means of either (a) a statistical one-pot method in which the desired macrocycle (*) is produced along with a statistical distribution of nine additional macrocycles or (b) a stepwise method in which the desired macrocycle is made over more steps in a targeted manner.

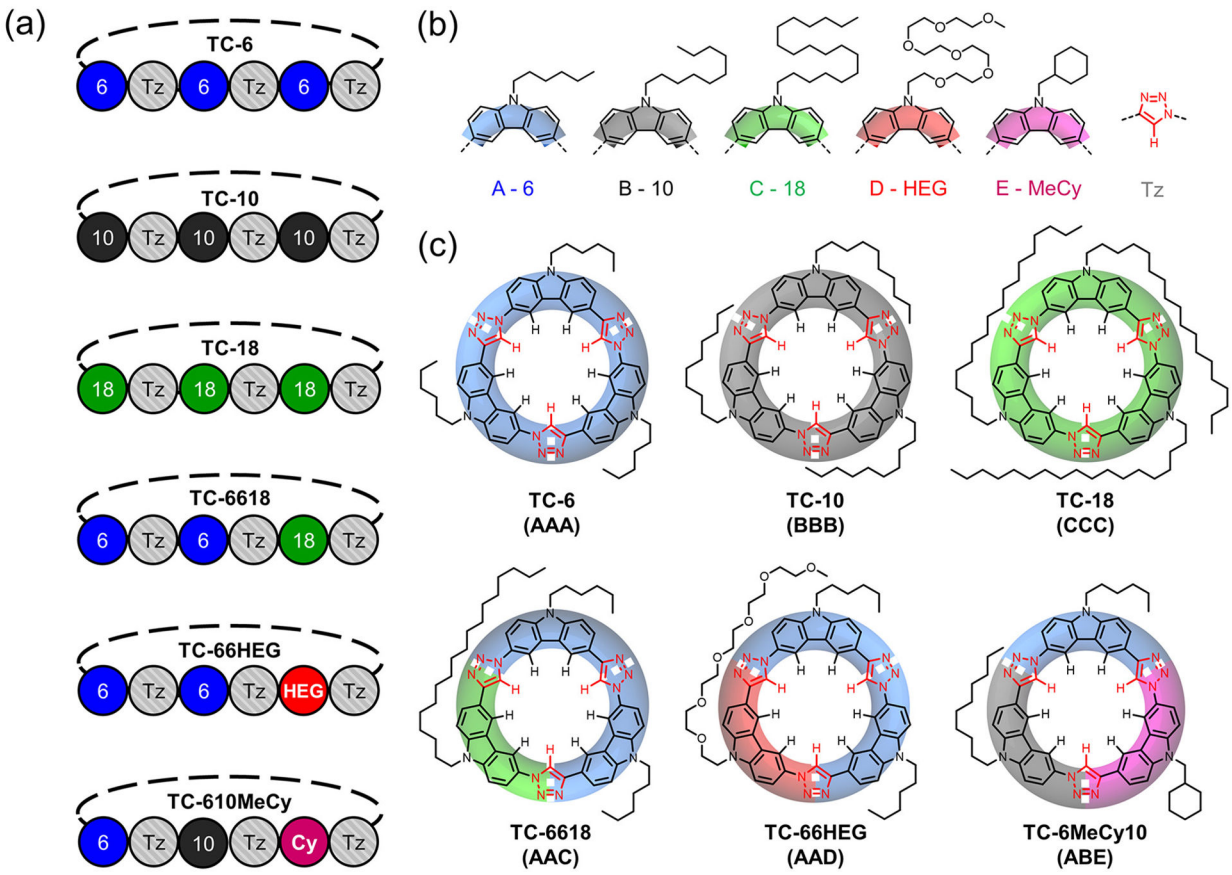


Figure 3.

(a) Primary (1°) sequence of the tricarb macrocycles synthesized in this work: AAA, BBB, CCC, AAC, AAD, ABE. (b) Formulas and color-coded letters of the sequence elements used in tricarb macrocycles. (c) Chemical structures of all the synthesized tricarb macrocycles overlaid onto their cartoon representations.

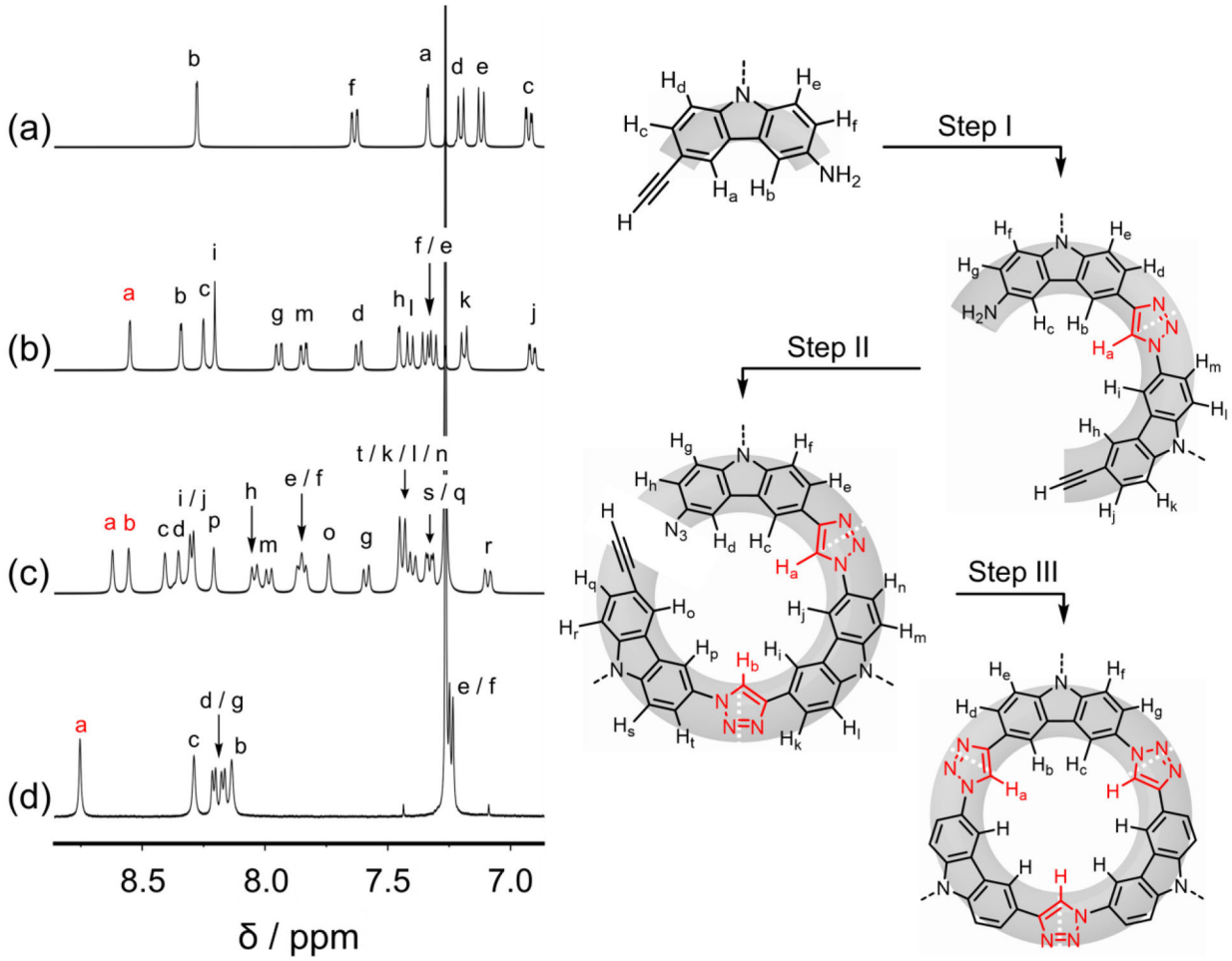


Figure 4.

Aromatic region of the ^{1}H NMR spectra of the key intermediates along the stepwise pathway (500 MHz, 298 K, $CDCl_3$).

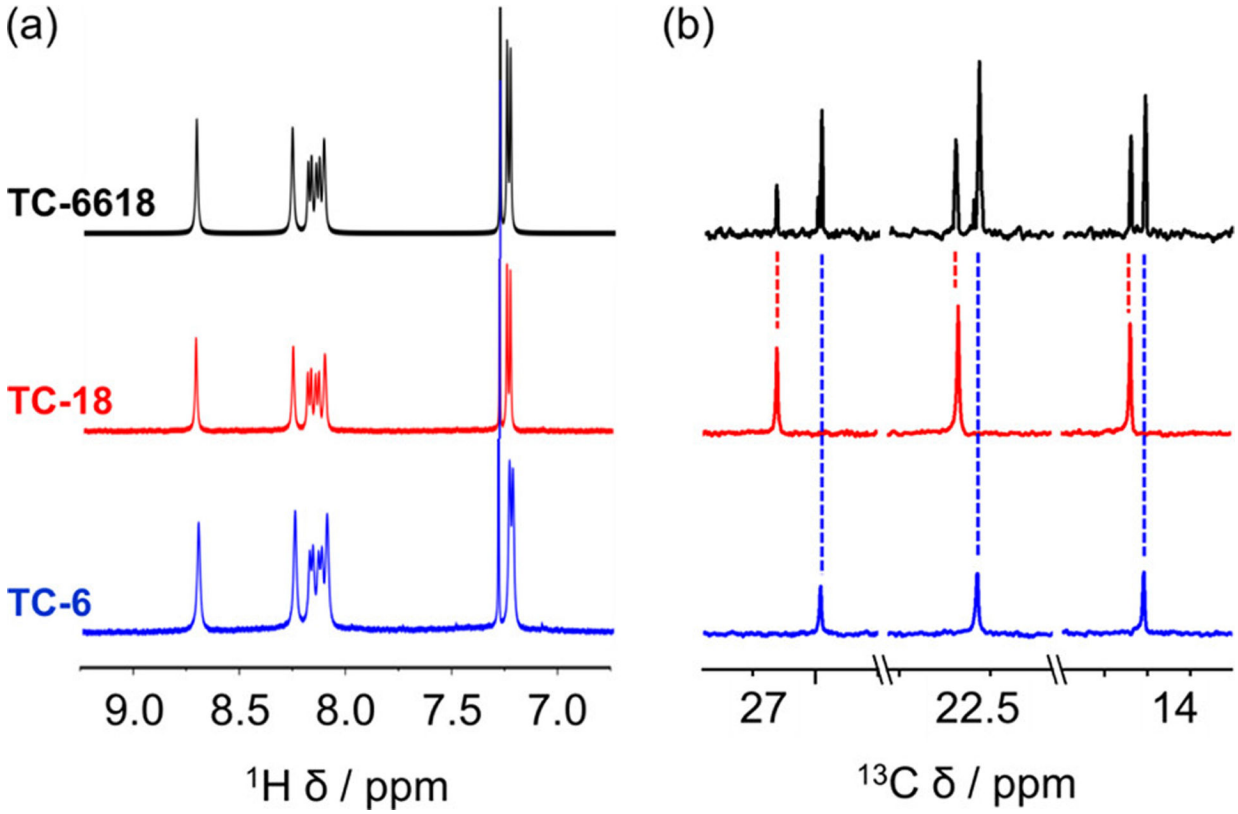


Figure 5.

(a) Aromatic regions of the ^{1}H NMR spectra (1 mM/500 MHz/$CDCl_3$/298 K) of **TC-6618**, **TC-18**, and **TC-6** show identical chemical shifts. (b) ^{13}C NMR spectra (1 mM/125 MHz/$CDCl_3$/298 K) of **TC-6618**, **TC-18**, and **TC-6** can only be distinguished by the unique peaks from the alkyl-chain substituents. ^{13}C NMR peaks arising from octadecyl chains are marked in red and peaks arising from hexyl chains are marked in blue.

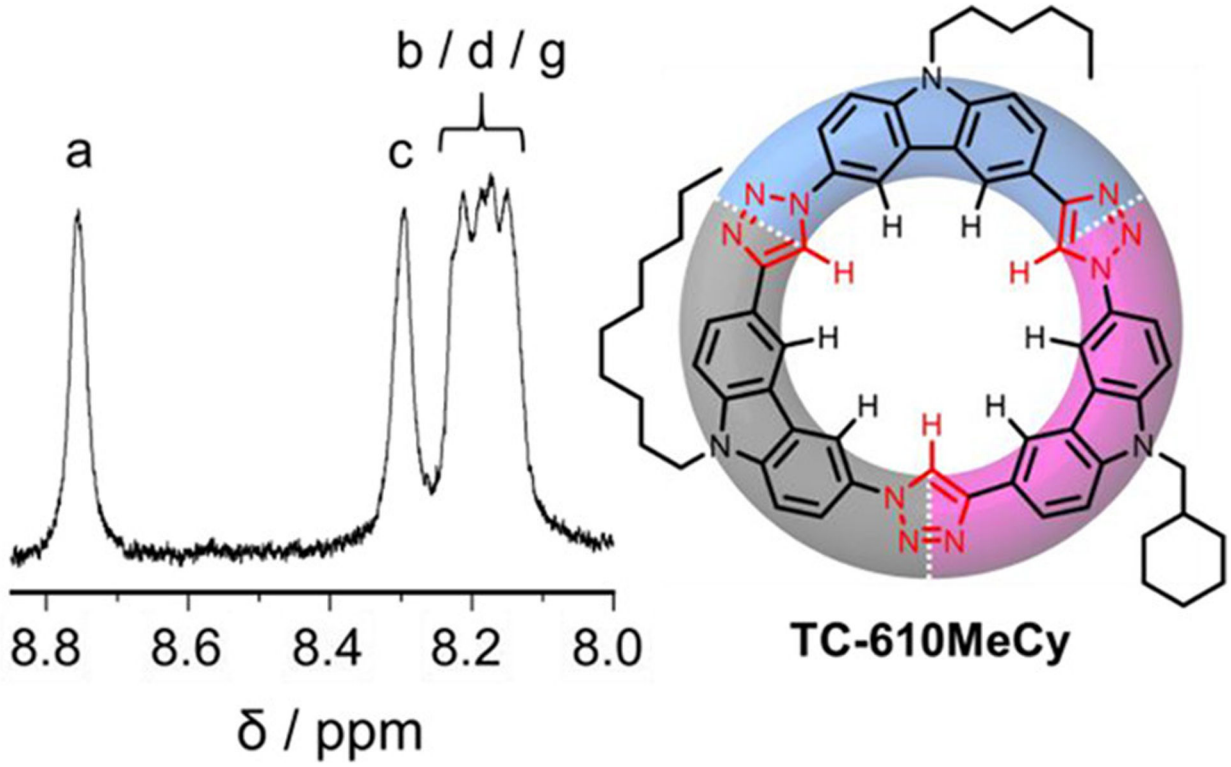


Figure 6.

Aromatic region of the ^{1}H NMR spectrum of **TC-610MeCy** (0.5 mM/600 MHz/$CDCl_3$/298 K).

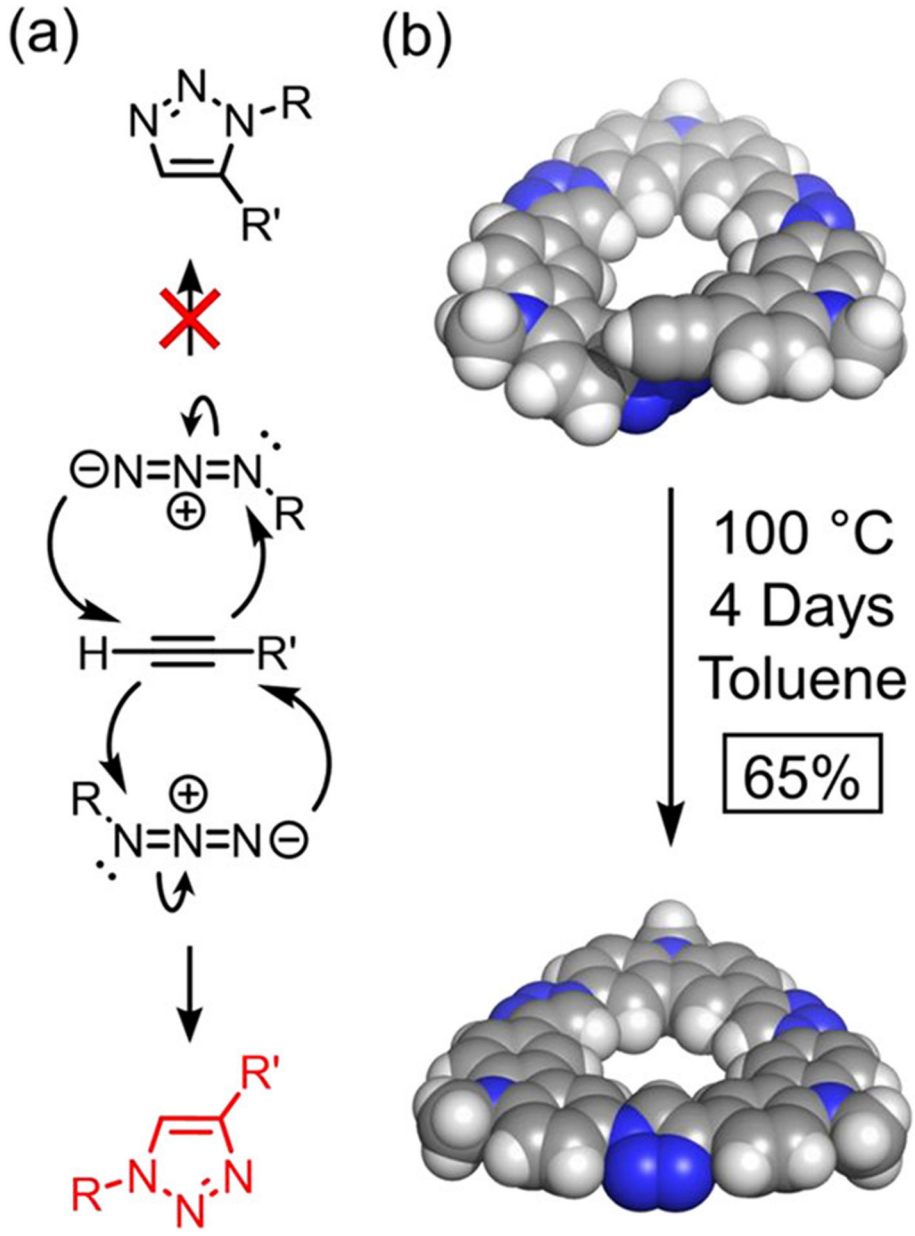


Figure 7.

(a) General mechanism in which (top) 1,5 and (bottom) 1,4 regioisomers can form in thermal Huisgen's azide-alkyne cycloaddition. (b) Molecular mechanics modeling (MMFF94) showing the orientation of the terminal azido and alkynyl functionalities within a tricarbazole crescent and the reaction conditions for the formation of **TC-10** under thermal conditions.

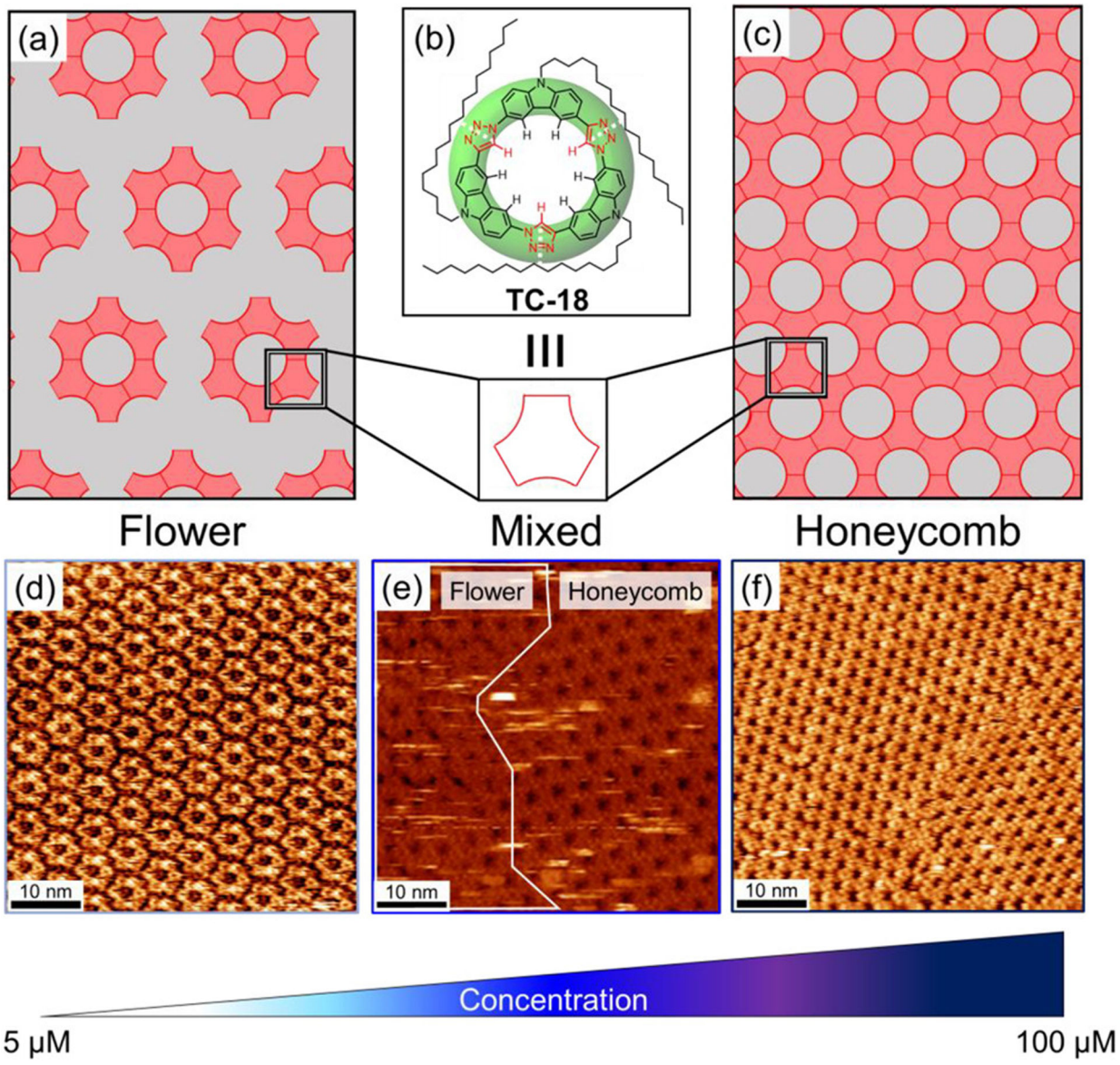


Figure 8.

Models of the (a) flower and (c) honeycomb phases formed by (b) **TC-18**. High-resolution STM images of **TC-18** at the TCB/graphite interface showing the (d) flower phase (5 μM, I_t = 0.3 nA, V_{sample} = −0.8 V), (e) mixed phases (75 μM, I_t = 0.03 nA, V_{sample} = −0.4 V), and (f) honeycomb phase (100 μM, I_t = 0.03 nA, V_{sample} = −0.4 V; unit cell: $a = b = 2.9 \pm 0.1$ nm).

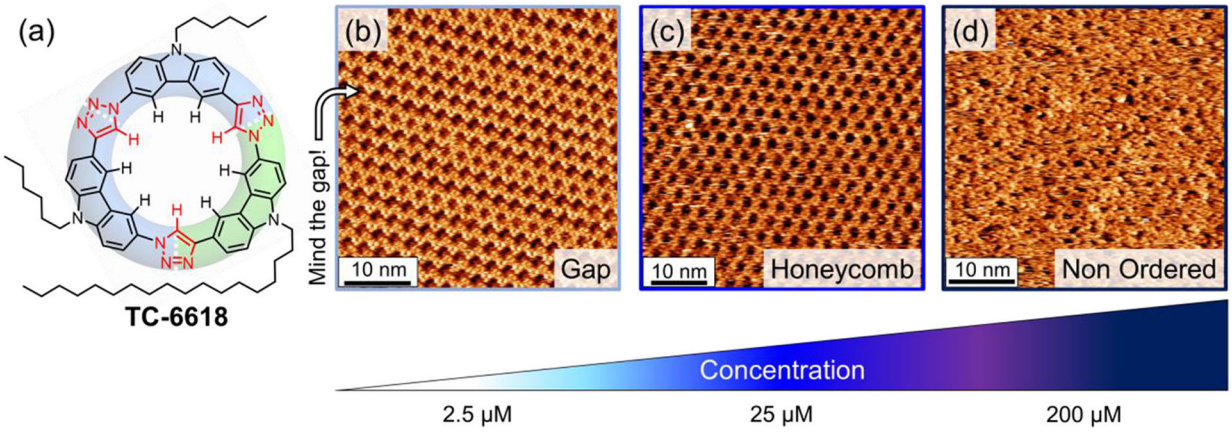


Figure 9.

High-resolution STM images of (a) **TC-6618** at the TCB/graphite interface showing (b) gap phase (25 μM, I_t = 0.15 nA, V_{sample} = −1 V), (c) honeycomb (unit cell: $a = b = 2.9 \pm 0.1$ nm; 100 μM, I_t = 0.55 nA, V_{sample} = −1 V), and (d) nonordered packing states (300 μM, I_t = 0.5 nA, V_{sample} = −0.8 V).

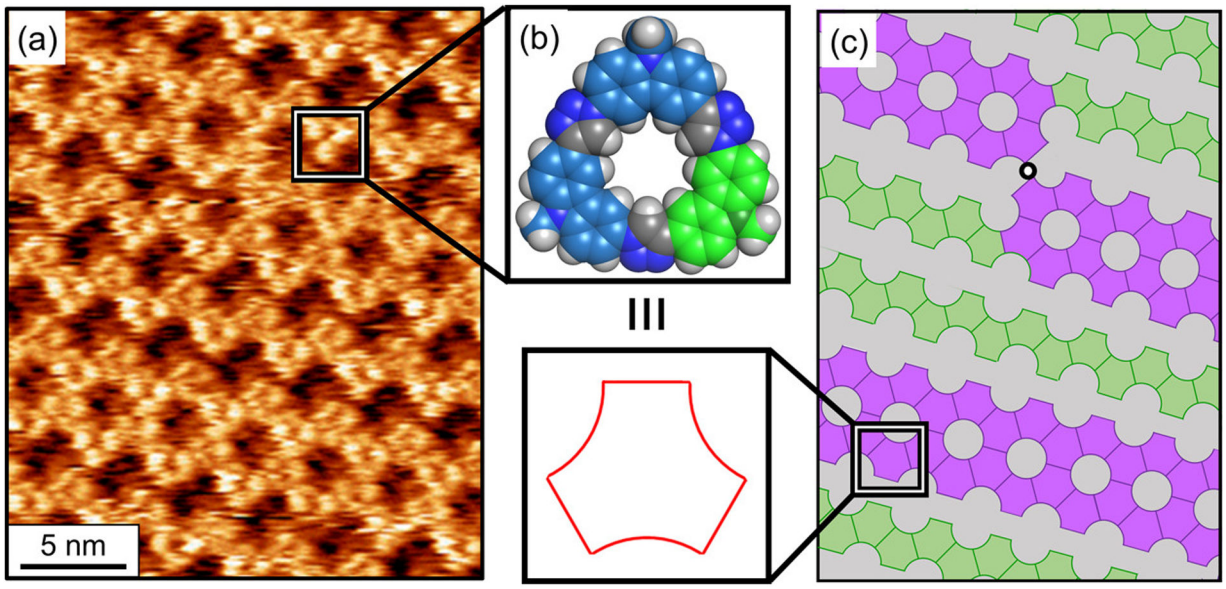


Figure 10.

(a) High-resolution STM imaging of the gap phase formed by (b) **TC-6618**. (c) Model of the gap phase showing the zigzag (green), honeycomb (purple), and the line defect transition between the two structures (black dot).

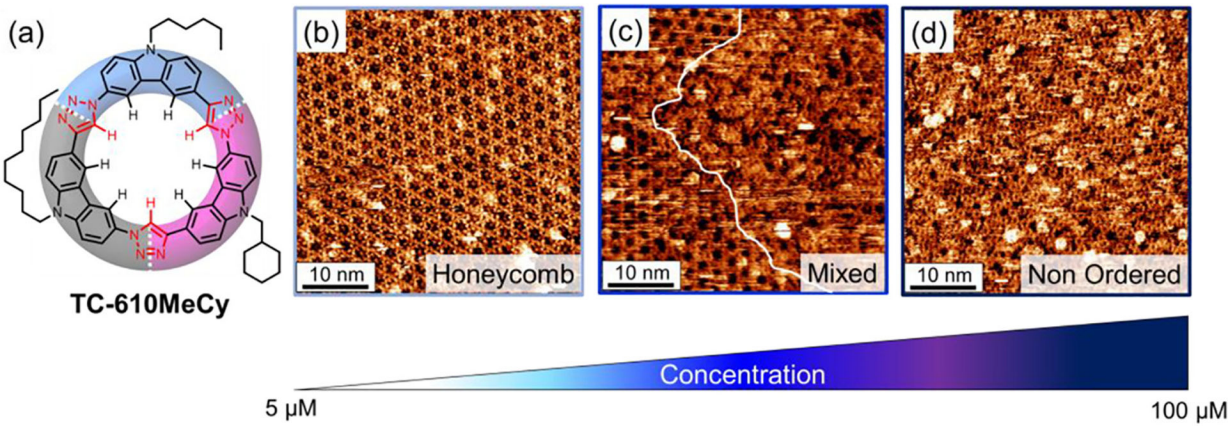


Figure 11.

(a) Macrocycle **TC-610MeCy**. High-resolution STM images of **TC-610MeCy** at the TCB/graphite interface showing (b) honeycomb ordering (unit cell: $a = b = 2.9 \pm 0.1$ nm 300 μM, $I_t = 0.13$ nA, $V_{sample} = -0.8$ V), (c) the coexistence of honeycomb and nonordered phases (300 μM, $I_t = 0.6$ nA, $V_{sample} = -0.5$ V), and (d) the nonordered packing state (300 μM, $I_t = 1.1$ nA, $V_{sample} = -0.6$ V).

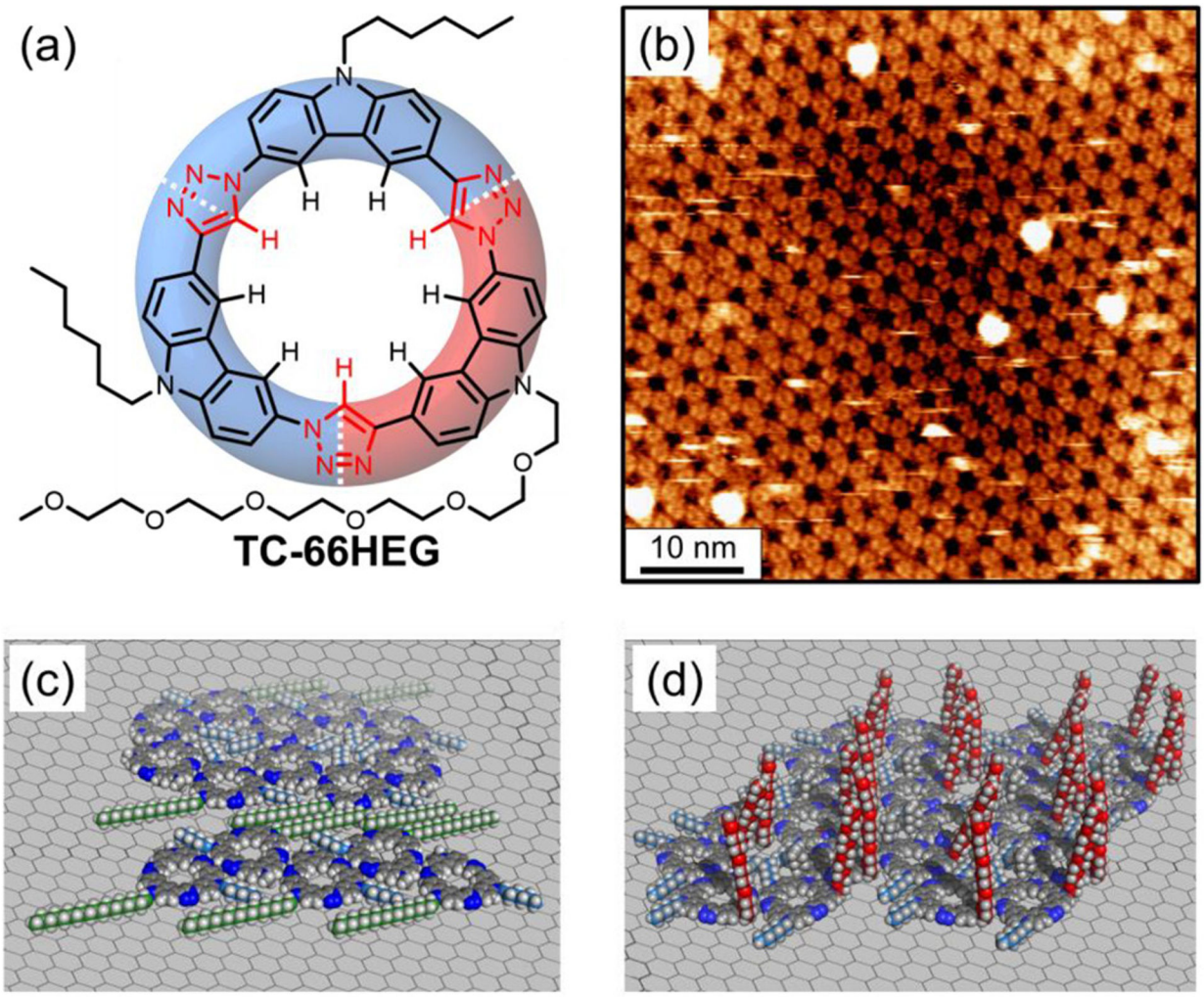
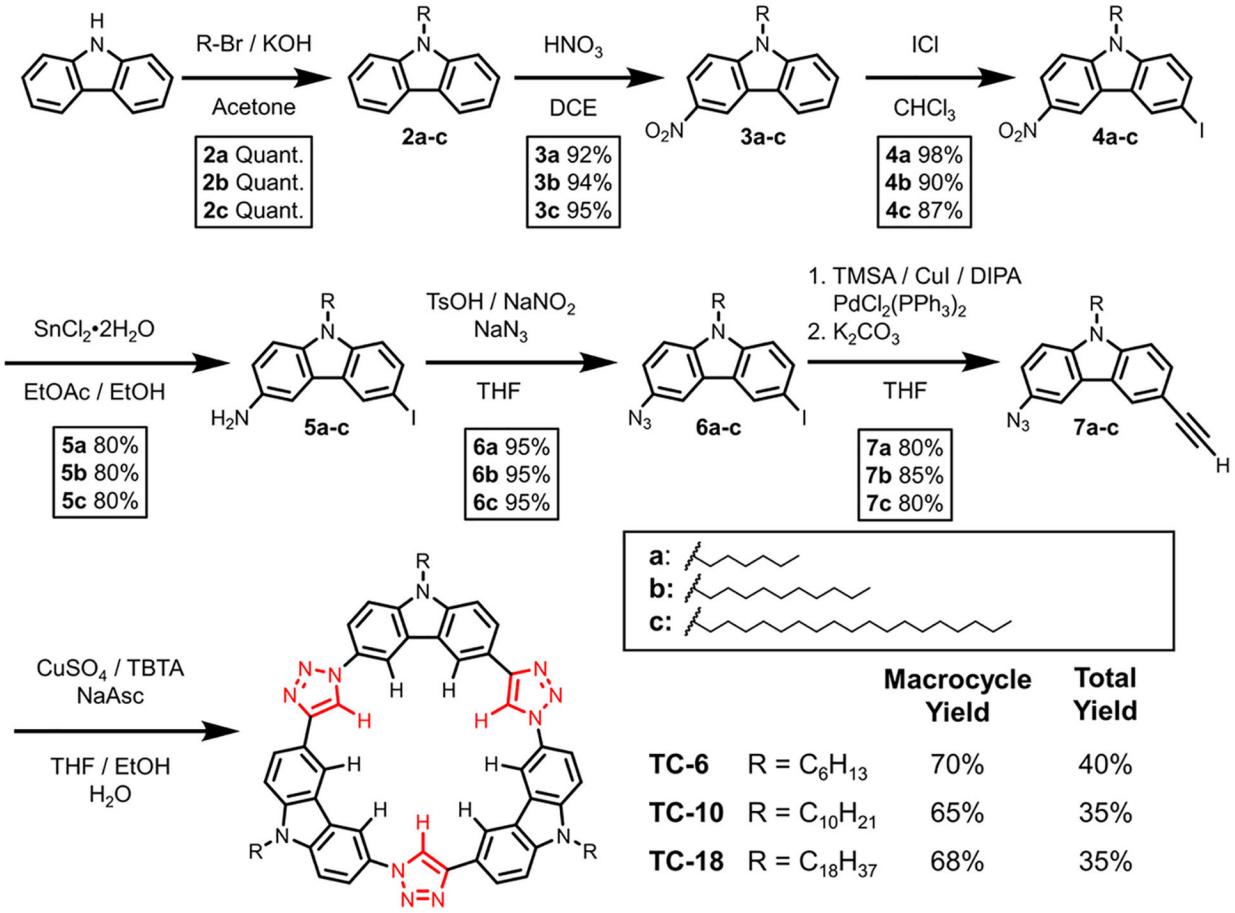


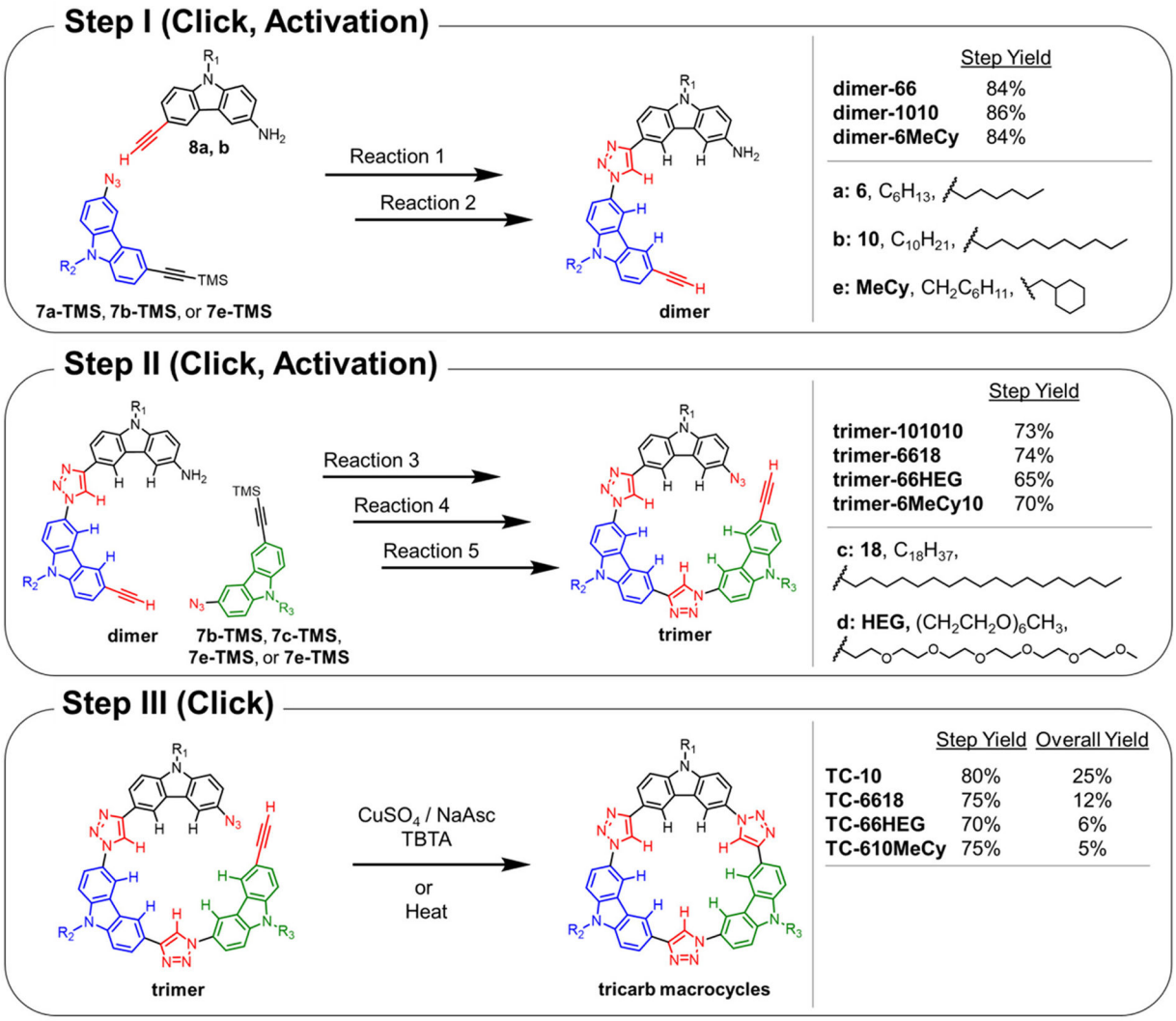
Figure 12.

(a) Macrocycle **TC-66HEG**. (b) STM images of **TC-66HEG** at the TCB/graphite interface with honeycomb ordering (10 μM, I_t = 0.15 nA, V_{sample} = −0.7 V). (c) Models of the octadecyl chains (green) of **TC-6618** surface adsorbed to form the gap and (d) HEG chains (red) of **TC-66HEG** directed away into solution leading to the formation of the honeycomb phase.



Scheme 1.

Synthetic Sequence Leading to the One-Pot Preparation of C_3-Symmetric Macrocycles (TC-6, TC-10, TC-18) from the Azide-alkyne Building Block 7 Bearing the Appropriate Alkyl Chain



Scheme 2. General Stepwise Synthesis of Sequence-Defined Tricarb Macrocycles[a]

[a]Step I: TMS-protected carbazoles (**7a-TMS**, **7b-TMS**, or **7e-TMS**) are coupled with amine-substituted carbazoles (**8a** or **8b**) and, after deprotection, give the corresponding crescent **dimer**. Step II: The crescent **dimer** is coupled with an additional TMS-protected carbazole building block (**7b-TMS** to **7d-TMS**). Step III: Following installation of the azide and deprotection of the alkyne, the crescent **trimer** can be closed to produce the tricarb macrocycle. Reactions 1 and 3: $CuSO_4$/NaAsc/TBTA/2:1:1 THF/EtOH/H_2O/55 °C/Ar. Reactions 2 and 5: K_2CO_3/1:1 MeOH/THF. Reaction 4: TsOH/$NaNO_2$/NaN_3/THF/0 °C.