ACCOUNTS
of chemical research

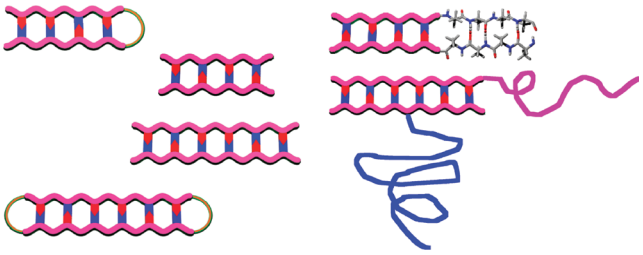
Molecular Duplexes with Encoded Sequences and Stabilities

BING GONG*

Department of Chemistry, University at Buffalo, The State University of New York, Buffalo, New York 14260, United States, and College of Chemistry, Beijing Normal University, Beijing 100875, China

RECEIVED ON JANUARY 6, 2012

CONSPECTUS



Through specific molecular shapes and repeating polymeric sequences, biomacromolecules encode information about both structure and function. Inspired by DNA molecules, we have conceived a strategy to encode linear molecular strands with sequences that specify intermolecular association, and we and our collaborators have supported this idea through our experimental work. This Account summarizes the design and development of a class of molecular duplexes with programmable hydrogen-bonding sequences and adjustable stabilities.

The specific system involves oligoamide strands synthesized from readily available monomeric modules based on standard amide (peptide) chemistry. By covalently linking three types of basic building blocks in different orders, we create oligoamide strands with various arrangements of amide O and H atoms that provide arrays of hydrogen bonding sequences. Because one of the two edges of these molecules presents the sequences of hydrogen-bond donors and acceptors, these oligoamide strands associate via their hydrogen-bonding edges into double-stranded pairs or duplexes. Systematic studies have demonstrated the strict sequence specificity and tunable stability of this system. These structurally simple duplexes exhibit many features associated with DNA sequences such as programmable sequence specificity, shape and hydrogen-bonding complementarity, and cooperativity of multipoint interactions.

Capable of specifying intermolecular associations, these duplexes have formed supramolecular structures such as β-sheets and non-covalent block copolymers and have templated chemical reactions. The incorporation of dynamic covalent interactions into these H-bonded duplexes has created association units that undergo sequence-specific association and covalent ligation in both nonpolar solvents and polar media including water. These new association units may facilitate the development of new dynamic covalent structures, and new properties are emerging from these structures. For example, we discovered hydrogen-bonded duplexes that could gelate different organic solvents, and we could tune the gelatinization by adjusting the multiple side chains attached to the duplexes. In addition, we have recently designed duplexes whose formation and dissociation are controlled by changes in external stimuli such as acidity.

With their programmable specificity and tunable stability, these molecular duplexes have provided a systematic approach for the association of different structural units. Further development of this system could facilitate the creation of many supramolecular and dynamic covalent structures. Because these duplexes are easily modifiable and information is easily encoded and retrieved, this system may address some of the remaining challenges facing information-storing molecules including self-replication.

Information-Carrying Molecules

Information, such as molecular shape, placement of different functionalities, and the type and number of intermolecular forces, encoded into and carried by molecular structures instructs intermolecular association to form various self-assembling architectures. The association of biomacromolecules like proteins results in the generation of numerous supramolecular assemblies.[1] For example, the protein coat of tobacco mosaic virus (TMV) consists of 2130 identical protein molecules that self-assemble into a well-defined tubular entity with an exterior diameter of ~18 nm and an inner channel of ~4 nm.[2] An encapsulated viral RNA

Published on the Web 05/14/2012 www.pubs.acs.org/accounts
10.1021/ar300007k

molecule further stabilizes the tubular aggregate and defines the length (~300 nm) of the resultant protein tube. While the structures of numerous protein assemblies have been determined, elucidating how protein molecules are encoded with the information for forming larger assemblies remains a daunting challenge.

In contrast to the complexities of protein self-assembly, the formation of DNA duplexes involves a powerful, readily predictable strategy that provides the basis for the storage and process of genetic information.[3] DNA strands are encoded by arranging the four nucleobases into different sequences. The sequence-specific pairing of cDNA strands defines highly specific molecular recognition events featuring (1) programmable sequences defined by the different linear arrangement of nucleobases, (2) shape and H-bonding complementarity, (3) cooperativity, that is, the correlation between the stability of a DNA duplex and its number of base pairs, and (4) self-replication, that is, a DNA strand can template the formation of its complementary strand.

Inspired by nucleic acids, the development of H-bonded pairs for specifying intermolecular interactions was started by Jorgensen[4,5] and Zimmerman with their investigation on nucleobases.[6] Highly stable heterocyclic complexes with arrays of H-bond donors (D) and acceptors (A) were subsequently developed by Zimmerman[7,8] and Meijer.[9−11] These H-bonded complexes were used as association units for directing the formation of supramolecular structures such as supramolecular polymers.[12−15] Many other H-bonded multicomponent systems have been studied,[16−28] which have provided invaluable insights into the effects of factors such as the size, shape, and number of H-bonds on the strength, complementarity, and fidelity of various heterocycle-based modules.

The development of DNA-like, information-carrying units with tunable strength and specificity was envisioned as the most likely strategy for realizing the specific control and adjustment of intermolecular association.[29,30] Besides serving as versatile "molecular glues" for linking various structural units, information-carrying molecules may also be equipped with additional interactions, leading to linking units that are compatible with different environments. Incorporating responsive triggers into the association units should allow the control of intermolecular association and dissociation by applying external stimuli. Finally, designing complementary molecular components that template each other's formation may lead to self-replicating molecular and supramolecular systems.

Over the years, my research group has been interested in developing oligoamide strands consisting of readily

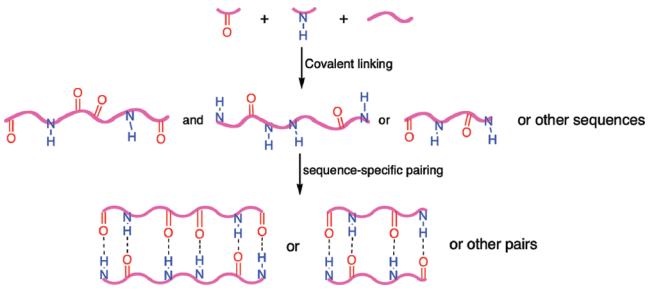


FIGURE 1. Oligoamide strands encoded with arrays of H-bond donors and acceptors. The resultant H-bonding sequences specify the association of two complementary strands.

available building blocks.[30−41] Our system involves encoding oligoamide strands by molecularly programming arrays of amide hydrogens and carbonyl oxygens as H-bond acceptors and donors. We have demonstrated the tunable stability, programmable sequence specificity, and convenient synthetic availability of these duplexes. The H-bonded duplexes that we developed represent a class of evolvable information-carrying molecules that have been employed for the instructed assembly of various self-assembling structures.

This Account discusses (1) the development of sequence-specific H-bonded duplexes, (2) the application of the duplexes as association units for specifying intermolecular association and for templating reactions, (3) the conversion of the duplexes into sequence-specific units in highly competitive media, and (4) our ongoing effort in creating responsive and intelligent information-carrying duplexes. Many other important information-storing or retrieving systems[41−51] are noteworthy but will not be discussed in this Account.

Molecular Strands Bearing Programmable Sequences

Sequence Specificity Based on a Binary Encoding Strategy[30]. Oligoamide strands with backbones bearing amide hydrogen and oxygen atoms as H-bond donors and acceptors were designed by combining residues derived from substituted 3-aminobenzoic acid, isophthalic acid, and *m*-phenylenediamine in different orders, leading to oligoamides having various H-bonding sequences. The resultant molecular strands are expected to form duplexes via H-bonding interactions between the backbone amide O and H atoms (Figure 1).

The formation of the corresponding duplex is mediated by multiple H-bonding interactions and thus should be highly specific. Like DNA, a strand may sequence-specifically

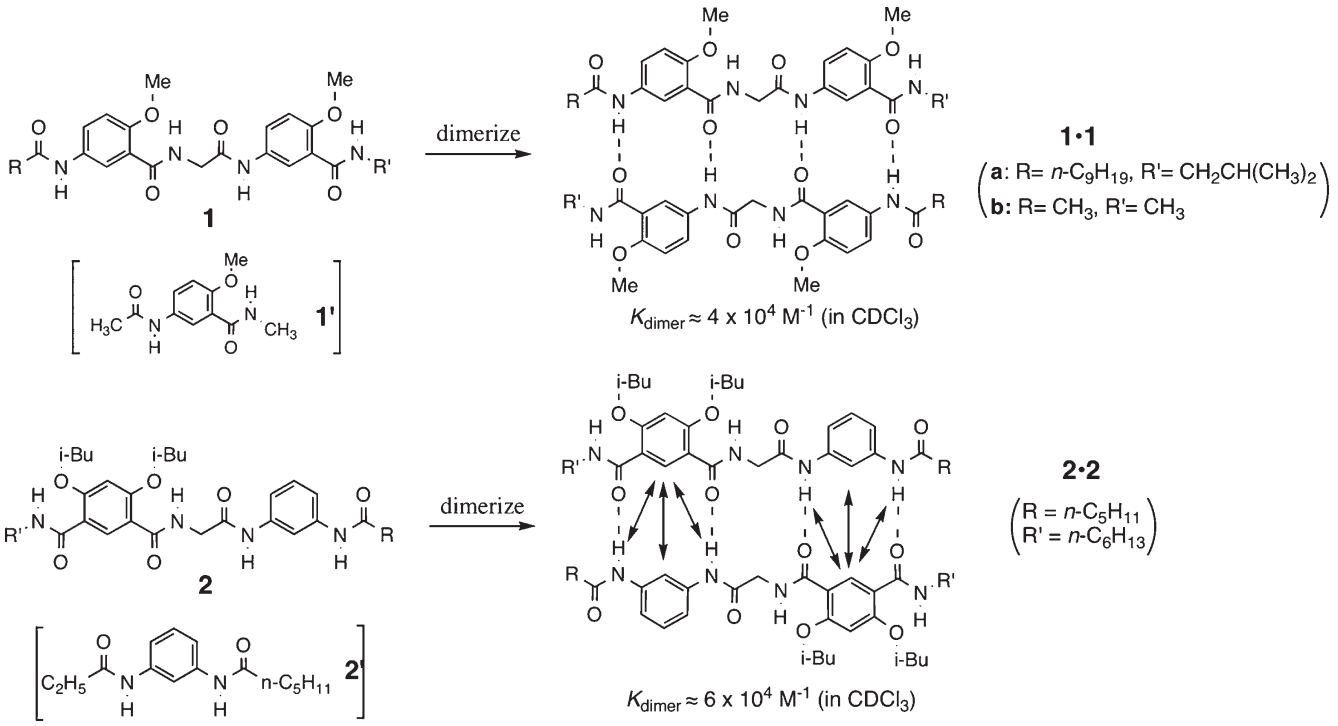


FIGURE 2. Quadruply H-bonded duplexes **1**·**1** and **2**·**2**. Despite their different H-bond donor (D)−acceptor (A) sequences, duplexes **1**·**1** and **2**·**2** have similar stabilities. Cross-strand NOEs revealed by NOESY spectrum is indicated for **2**·**2**.

pair with another complementary strand. The number of unique sequences increases very rapidly as the number of intermolecular H-bonds increases. For example, there should be six different quadruply H-bonded duplexes. For duplexes having six intermolecular H-bonds, the number of different duplexes becomes 20.[32]

Sequence-Independent Stability. Oligoamides **1** and **2** (Figure 2) represent the first examples of the H-bonded duplexes that we designed.[31] These oligoamides are extended because their backbones are preorganized by the intramolecular H-bonds that also block undesired H-bonding interactions. Thus, oligoamides **1** and **2** should dimerize via their self-complementary sequences of DADA and DDAA. In addition, different side chains can be incorporated into the design, leading to duplexes that are compatible with different solvents.

The formation of homoduplexes **1**·**1** and **2**·**2** were confirmed by using multiple analytical techniques. Consistent with the presence of intermolecular H-bonding, ^{1}H NMR spectra measured in $CDCl_3$ revealed significant downfield (>2 ppm) shifts of the aniline NH signals of **1** and **2** relative to those of **1**′ and **2**′. The stability of the duplexes were quantified by ^{1}H NMR binding studies carried out in $CDCl_3$, which revealed dimerization constants of $\geq 4.4 \times 10^4$ M^{-1} for **1a**·**1a** and $\sim 6.5 \times 10^4$ M^{-1} for **2**·**2**. Given the errors of NMR binding experiments (±10%), duplexes **1a**·**1a** and **2**·**2** can be regarded as having similar stabilities. The dimerization of **1a** and **2** was also demonstrated by results from vapor-pressure osmometry (VPO).

The NOESY spectrum of **2** in $CDCl_3$ contains interstrand NOEs (see Figure 2) that are consistent with a H-bonded dimer. The crystal structures of **1a** and **2** both revealed the expected dimeric structures held together by intermolecular H-bonds (Figure 3).

Tunable Stability. The similar dimerization constants of **1**·**1** and **2**·**2** suggest that the stability of these H-bonded complexes is sequence-independent, an observation that contrasts many previously reported H-bonded dimers based on heterocycles.[4−6] The number of interstrand H-bonds of a duplex can be easily increased by adding additional residues, which should lead to an increase in the stability of a duplex. Such an expectation was confirmed by the six-H-bonded **3**·**4** (Figure 4a) consisting of two different strands having complementary H-bonding sequences DADDAD and ADAADA.[32]

The formation of duplex **3**·**4** was first indicated by comparing its solubility with that of **3** or **4**. While the solubility of single strand **3** (<1 mM) or **4** (~10 mM) in chloroform was low or modest, the 1:1 mixture of **3** and **4** was found to be highly soluble (≫100 mM) in the same solvent.

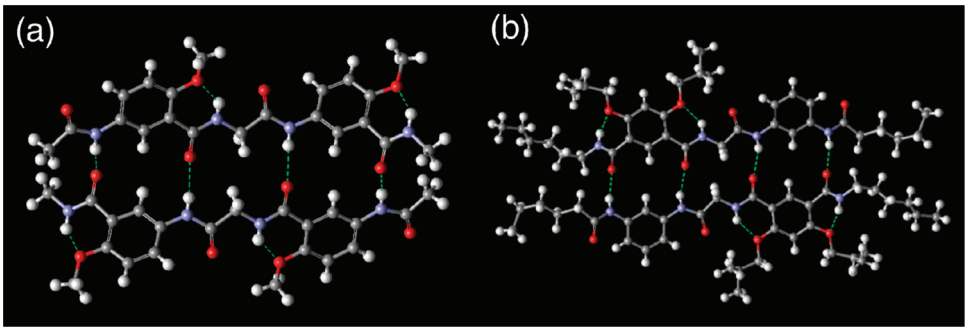
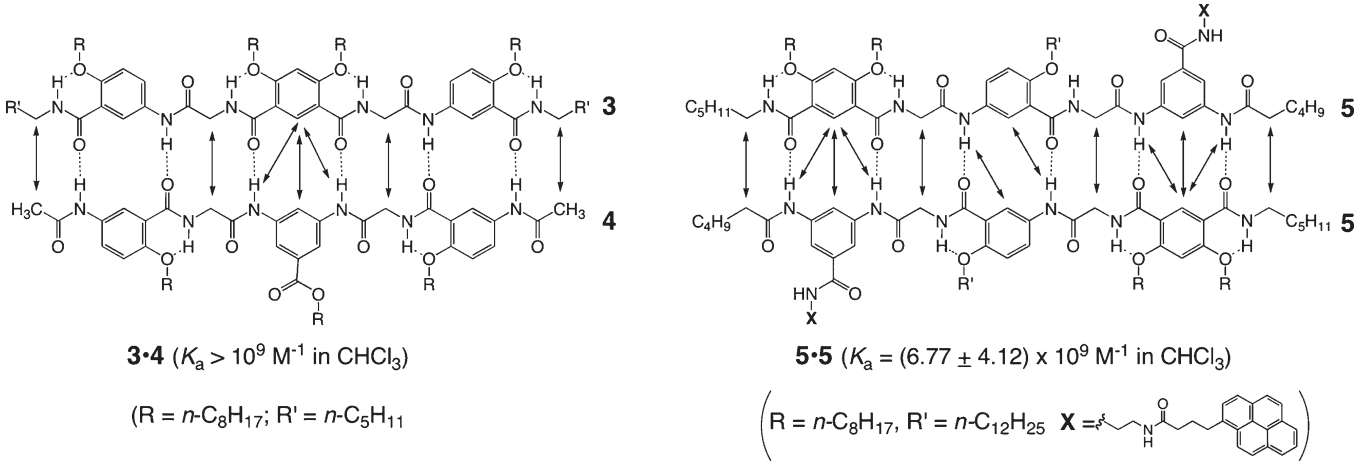


FIGURE 3. The crystal structures of duplexes (a) **1a**·**1a** and (b) **2**·**2**.



3•4 ($K_a > 10^9$ M^{-1} in $CHCl_3$)

(R = *n*-C_8H_{17}; R' = *n*-C_5H_{11})

5•5 (K_a = (6.77 ± 4.12) x 10^9 M^{-1} in $CHCl_3$)

(R = *n*-C_8H_{17}, R' = *n*-$C_{12}H_{25}$ **X** =)

FIGURE 4. (a) Duplex **3**·**4** consisting of two different strands with complementary H-bonding sequences. (b) Duplex **5**·**5** consisting of two identical single strands with a self-complementary H-bonding sequence. Interstrand NOE contacts revealed by NOESY are indicated by arrows.

The ^{1}H NMR spectrum of the 1:1 mixture of **3** and **4** in $CDCl_3$ revealed very large (2−3 ppm) downfield shifts of the aniline NH signals compared with those of **3** or **4**. No detectable concentration-dependent shifts of the aniline NH signals were observed from 100 mM to 1 μM. A lower limit of 9 × 10^7 M^{-1} for the association constant of **3**·**4** was estimated by assuming a 10% dissociation at 1 μM. Isothermal titration calorimetry (ITC) could only give an estimated value of ≥10^9 M^{-1}. The high stability of **3**·**4** allowed its detection by thin layer chromatography (TLC) under rather polar conditions (silica gel plate, 10% DMF in chloroform). The ^{1}H NMR spectra of mixtures of **3** and **4** in stoichiometries other than 1:1 revealed separate sets of signals corresponding to both duplex **3**·**4** and the excess single strand. This observation suggested that exchange between the assembled and the uncomplexed states was very slow, suggesting a high kinetic stability for **3**·**4**. The presence of **3**·**4** in solution was confirmed by numerous interstrand NOEs in NOESY spectrum of this duplex (Figure 4a).

NOESY spectra the pyrene-labeled homoduplex **5**·**5** (Figure 4b) revealed cross-strand NOEs, confirming the formation of this duplex.[36] The dimerization constant of the duplex **5**·**5** was found to be (6.77 ± 4.12) × 10^9 M^{-1} based on a known method.[53] The studies on **3**·**4** and **5**·**5** clearly demonstrated that the stabilities of these duplexes are indeed only determined by the number of intermolecular H-bonds.

Thus, with two additional H-bonds (i.e., the four- vs six-H-bonded duplexes), stability is enhanced by more than 5 orders of magnitude. Comparing the stabilities of doubly and quadruply H-bonded pairs with that of **3**·**4** indicates that the increase of stabilities is not due to the additive effect of hydrogen bonds but resulted from the cooperativity of multiple H-bonding, which is one of the most important features associated with the self-assembly of DNA.[54]

Strict Sequence Specificity. Sequence-specific pairing of nucleic acids is essential for the storage, transmission, and expression of genetic information. Having probed the factors determining the stabilities of our H-bonded duplex, the sequence specificity of these information-storing molecules was then explored. Incorporating mismatched binding sites into the otherwise fully matched duplex **3**·**4** led to "mutant"

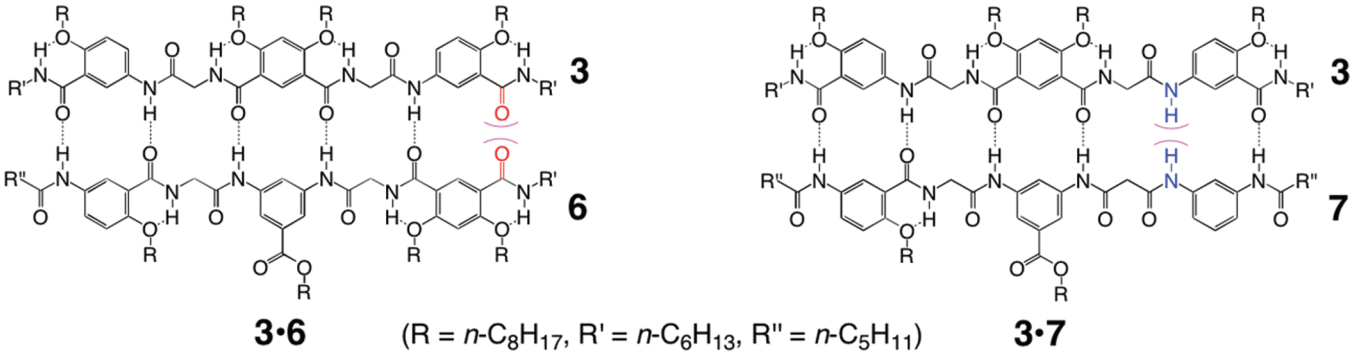


FIGURE 5. "Mutant" duplexes **3·6** and **3·7** derived from **3·4**. The presence of the mismatched binding sites in **3·6** and in **3·7** causes more than 40-fold decrease in the stabilities of these duplexes compared with that of **3·4**.[33]

duplexes (Figure 5).[33] Single strands **6** and **7**, with the H-bonding sequences of DADDAA and DADDDD, were designed and synthesized to pair with **3**. If H-bonded duplexes **3·6** or **3·7** could still form, each of these "mutant" duplexes should contain a mismatched binding site. However, it was not clear that, in the presence of the mismatched site, how single strands **3** and **6** or **3** and **7** would align with each other.

^{1}H NMR showed poorly defined signals for individual strand **6** or **7**, indicating a mixture of many conformations. Upon mixing with one equivalent of **3**, both **6** and **7** gave sharp sets of signals that can be attributed to single conformers. The aniline NH protons of **3**, which appear as one degenerate signal in the ^{1}H NMR spectra of **3** or **3·4**, split into two signals that differed by >0.2 ppm in the spectra of **3·6** or **3·7**. Obviously, the exchange of single strands **3**, **6**, and **7** with the corresponding duplex **3·6** or **3·7** was slow on the NMR time scale. This phenomenon is consistent with the formation of H-bonded **3·6** and **3·7**, which have unsymmetrical structures and relatively high kinetic stabilities.

The stabilities of **3·6** and **3·7** were compared with that of **3·4**. Results from ITC experiments showed that introducing one mismatched site led to >40-fold drop in the association constants of the mismatched **3·6** and **3·7** relative to that of **3·4**.[33]

The NOESY spectrum of **3·6** showed that one end of this pair is locked by intermolecular H-bonds while the other end with the mismatched site was open. On the other hand, NOESY study indicated that both ends of **3·7**, which contains a mismatched site in the middle, were locked, while no or very weak NOEs were detected for protons close to the internal mismatched site. These results indicated that, despite the presence of mismatched sites, these "mutant" duplexes could manage to get around the presence of the repulsive mismatched sites to remain registered. Detailed analysis confirmed that there were five interstrand H-bonds in **3·6** or **3·7**, suggesting that these strands remained associated via the maximum possible number of H-bonds. It was observed that adding one equivalent of **4** into the solution of either **3·6** or **3·7** led to the complete replacement of the "wrong" strand **6** or **7**.

Duplex-Instructed Self-Assembly

Templated Formation of Two-Stranded β-Sheets. When tethered to structural units that associate in more than one way, a H-bonded duplex should specify the intermolecular interaction, leading to a single assembly. Since the interstrand distance of a duplex, based on the crystal structures of four-H-bonded duplexes,[31] is the same as that (~5 Å) in a β-sheet, these H-bonded duplexes may serve as noncovalent templates for directing and nucleating β-sheet structures when attached to natural oligopeptide strands. Because the H-bonding sequence of the duplex template can be designed to be unsymmetrical, two flexible peptide chains may be brought into close proximity and thus be forced to pair with each other, leading to a double-stranded, antiparallel β-sheet (Figure 6).

Our results confirmed the feasibility of this strategy.[34] Peptide segments were attached to an unsymmetrical, four-H-bonded heteroduplex template with the complementary sequences of ADAA/DADD, leading to four hybrid strands **8a**, **8b**, **9a**, and **9b**. Combination of strands **8** with strands **9** resulted in four different pairs. The formation of well-defined β-sheets by paired peptide sequences was confirmed by 1D and 2D NMR, ITC, and VPO studies. For example, 2D NMR (NOESY) showed interstrand NOEs corresponding to protons of duplex templates (in purple) and the peptide strands (in blue). No cross-strand NOEs were observed between the template and peptide segments, suggesting that the hybrid strands were registered as expected. ITC experiments showed that the hybrid strands showed enhanced stability

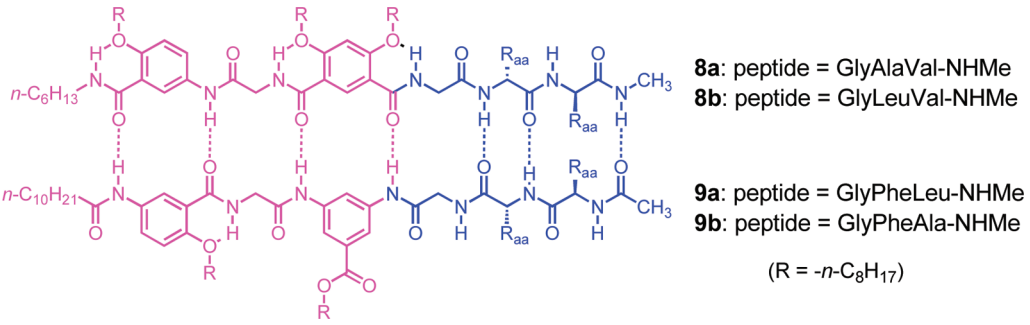


FIGURE 6. When attached to a duplex template, two otherwise flexible peptide chains are directed to form a stably folded β-sheet.

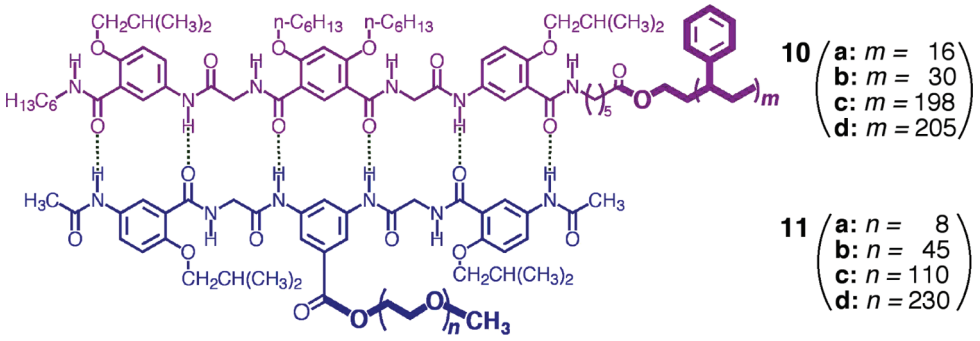


FIGURE 7. Design of supramolecular block copolymers based on the six-H-bonded heteroduplex **10**·**11**.

compared with the duplex template or the peptides alone. VPO measurements showed that strands **8** and **9** formed heterodimers while the corresponding peptide strands form poorly defined, random aggregates.

This supramolecular approach should lead to a rapid generation of a large number of combinations by simple mixing. In contrast to known model systems of β-sheets, most of which involve linear peptide segments linked by turns and are thus strongly dependent on the peptide sequences for their formation, this system may simplify the study of factors that stabilize β-sheet by focusing on the interactions of linear peptide segments without having to rely on turn structures. Furthermore, the directed assembly of peptide chains may offer an effective synthetic strategy when combined with covalent cross-linking.

Supramolecular Block Copolymers. Our H-bonded duplexes also served as units for linking incompatible structures that do not normally associate. This was demonstrated by constructing supramolecular block copolymers (Figure 7).[38,55] Specifically, hydrophobic polystyrene and hydrophilic poly(ethylene glycol) chains were attached to the two strands of a duplex derived from the six-H-bonded **3**·**4**, leading to modified polymers **10** and **11**.

The noncovalent ligation of the PS and PEG chains was confirmed by GPC, which showed quantitative formation of the H-bonded block copolymers.[38] In toluene with 10% DMF, the GPC result showed a single peak corresponding to the H-bonded **10c**·**11c**, which appeared earlier than **10c** and **11c**. When DMF was used as the eluent, **10c**·**11c** gave two peaks that coincided with those of **10c** and **11c**. This result demonstrates that the attached polymer chains did not affect the sequence-specific formation of the H-bonded duplex. NOESY study confirmed that the duplex template was precisely registered. Differential scanning calorimetry (DSC) analysis indicated the thermally reversible formation and dissociation of these supramolecular block copolymers.

For pairs with long PS and PEG chains (**10c**·**11c** and **10d**·**11d**), microphase separated nanodomains, typical of covalent diblock copolymers, were observed by atomic force microscopy (AFM). Hexagonally packed cylinders can be clearly identified (Figure 8a). In contrast, the AFM image of the 1:1 blend of the corresponding PS and PEG chains only revealed macrophase separation. Pairs with short PS and PEG chains (e.g., **10a**·**11a** and **10b**·**11b**) assembled into very different morphology.[56] The AFM image of **10b**·**11b** revealed nanostructures having nothing in common with typical diblock copolymers (Figure 8b), which may be understood by regarding **10b**·**11b** as a supramolecular coil−rod−coil triblock copolymer given that the six-H-bonded duplex is a rigid tape. These results suggest that

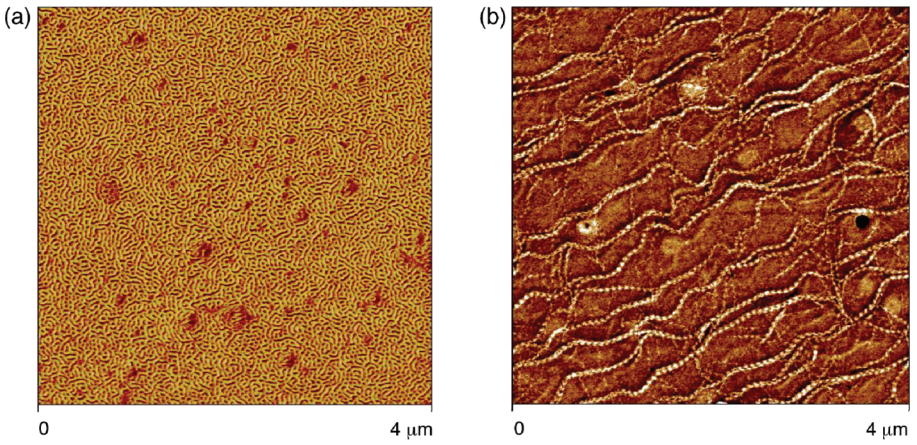


FIGURE 8. (a) The AFM image of spin-cast **10d**·**11d** from benzene showing cylindrical nanodomains from the microphase separation of the supramolecular block copolymer. Hexagonally packed cylinders can be identified in regions where the cylinders are perpendicular to the surface. (b) The AFM image of **10b**·**11b** spin-cast from benzene showing self-assembled fibers.

when the polymer blocks are sufficiently long, the effect of the H-bonded duplex can be neglected. When the polymer chains are short, the H-bonded duplex unit plays a significant role in forming the self-assembling architecture.

This system has provided a platform for creating a wide variety of block copolymers with tunable properties. In addition to linear block copolymers, the multiple side chains of the duplex units offer sites for attaching different oligomer and polymer chains, which should lead to brush- or star-like block copolymers that otherwise would be difficult to prepare. Oligomer and polymer chains that are difficult to link together based on traditional covalent methods can be modified separately and then linked via H-bonds. Furthermore, the simplicity of this method allows the convenient construction of supramolecular combinatorial libraries, which should lead to the rapid generation and screening of various materials.

Instructed Chemical Reaction: Templated Olefin Metathesis. When attached to our H-bonded duplexes, reactive species could be brought together, leading to enhancement of reaction rates. Disulfide exchange and Pd-catalyzed olefin metathesis are two reversible covalent reactions that have attracted intense recent interest.[57] We first probed the incorporation of olefin units into our H-bonded duplexes. Based on a duplex template with the same unsymmetrical H-bonding sequence used for directing the formation of β-sheets, we prepared two groups (strands **12** and **13**) of five olefins covalently linked to the two template strands respectively (Figure 9). Mixing each of **12** with each of **13** in a 1:1 fashion results in a small library of 25 (5×5) members.

Formation of the ADAA/DADD duplex should bring the olefin moieties into close proximity, which lowers the $\Delta S^{\ddagger}$ of the reaction and thus facilitates the metathesis reactions. Indeed, in the presence of Grubb's catalyst, most of the templated pairs were cross-linked in very high yields. In cases when the tethers were too short, such as pair **12e** and **13e**, metathesis reaction did not occur. With strands **12** or **13** alone, no homo-cross-linked product was observed. It was also found that under the same conditions (r.t., 1−2 mM in $CDCl_3$) the same reactions did not happen with the corresponding free olefins. These results indicate that the H-bonded duplex template promoted the metathesis reactions by increasing the effective molarity of the reacting olefins. The duplex template instructed the otherwise symmetrical olefin metathesis reaction to proceed in an unsymmetrical fashion. In principle, this strategy should be applicable to instructing other bimolecular reactions.

Assembly Dependent Function: Gelation of Organic Solvent. Recently, oligoamide duplexes derived from the quadruply H-bonded **1**·**1** or **2**·**2** that consist of strands bearing self-complementary H-bonding sequences were found to gelate a range of organic solvents.[58] The observed gelation was found to be dependent on the presence of the H-bonded duplexes. The relatively flat duplexes, each consisting of two oligoamide strands (see Figure 3), undergo further assembly mediated by strong $\pi-\pi$ aromatic stacking and van der Waals interactions, which leads to the formation of fibrous networks that gelate solvents. The multiple side and end chains allowed convenient structural modification and the tuning of the gelling behavior of the duplexes. These findings may lead to a new strategy for preparing a novel class of functional soft materials with tunable properties.

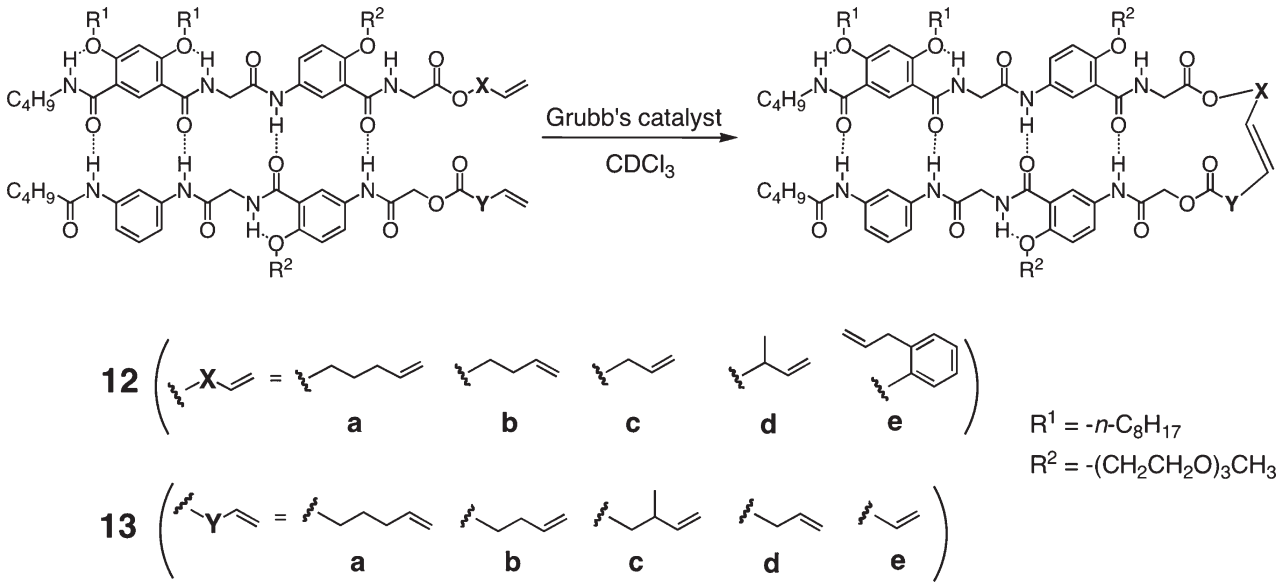


FIGURE 9. Templated olefin cross-metathesis of olefins tethered to the two duplex strands.

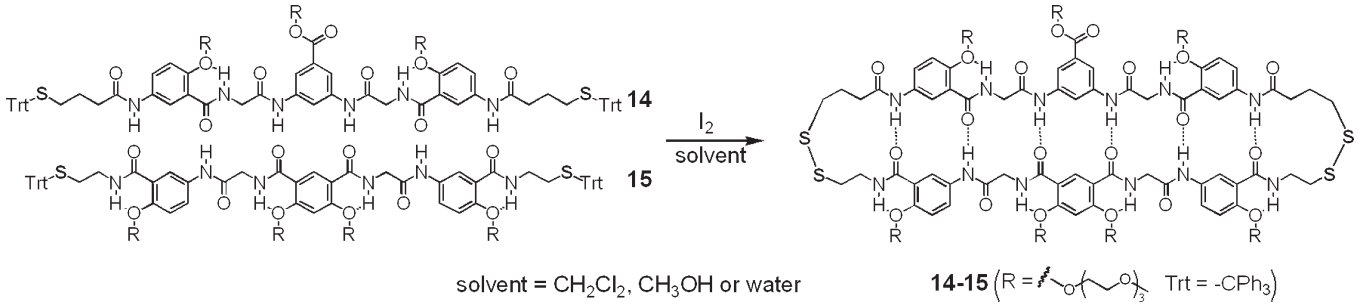


FIGURE 10. Complementary strands **14** and **15** carrying protected thiol groups are sequence-specifically cross-linked when subjected to reversible redox conditions in both nonpolar and polar media.[59,60]

Sequence-Specific Association in Aqueous Media: H-Bonded Duplexes Equipped with Disulfide Linkages

An unsolved challenge in the application of H-bonded assemblies is the instability of H-bonding in competitive media. Can the strategy of combining the sequence specificity of H-bond arrays with the strength of reversible covalent interactions be extended into developing stably associated structures in polar media? To answer this question, strands **3** and **4** were modified with protected thiol end groups, leading to complementary strands **14** and **15** (Figure 10).[59,60]

The 1:1 mixtures of strands **14** and were prepared in CH_2Cl_2, methanol, and water. The samples were subjected to I_2 and then examined by MALDI. It was found that **14** and **15** sequence-specifically associated with each other not only in CH_2Cl_2, but also in methanol and water, leading to disulfide-cross-linked **14**–**15**. The sequence dependence of the cross-linking of **14** and **15** in water became obvious when a control strand (not shown) with two mismatched binding sites was mixed with **14** and **15**. MALDI revealed that the control strand did not interfere with the formation of **14**–**15**. Strand **14** or **15** alone formed self-cyclized product and only a very small amount of cross-linked self-dimers, further demonstrating the H-bond-dependence of the disulfide cross-linking reaction. Thus, similar to duplex-templated olefin cross-metathesis, the otherwise symmetrical disulfide formation has also been directed into an unsymmetrical process by a duplex consisting of two different strands.

Subsequent mechanistic studies revealed a thermodynamically controlled process.[60] When two strands with complementary sequences undergo reversible cross-linking, possible products include those corresponding to the self-cyclized strands **A***, **A**–**A**, **B***, **B**–**B** and the sequence-matched duplex **A**–**B** (Figure 11). Among these products, **A**–**B** gains the most stabilization from the newly generated, complementary intramolecular H-bonds due to

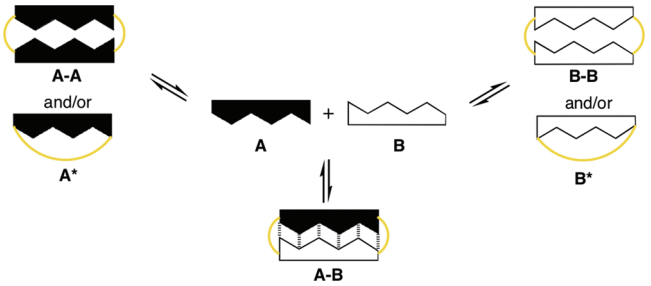


FIGURE 11. Two components bearing complementary H-bonding sequences reversibly form disulfide-cross-linked products. **A**–**B** represents the most stable product due to the stabilization from its fully matched, interstrand (intramolecular) H-bonds.

the formation of the disulfide bonds. The reversibility of dynamic covalent cross-linking allows the equilibrium to shift toward the most stable product. Therefore, the observed sequence specificity in the formation of the cross-linked **14**–**15** was due to the stabilization provided by the six intramolecular H-bonds. The thermodynamic nature of this system was further verified by the fact that as few as two H-bonds were sufficient to shift the equilibria toward the formation of sequence-matched products in aqueous solutions.

Toward Responsive Systems

One strategy to control the association of H-bond-mediated association is to add a complementary strand as a "chain stopper" that competes for H-bonding with the association unit. Such an approach was adopted by Meijer et al. to reverse the formation of H-bonded supramolecular polymers.[12] The complete interruption of intermolecular association mediated by H-bonded pairs of very high affinity requires the presence of a sufficiently high concentration of the chain stopper, which may not be a desirable feature for many applications.

"Mutant" duplexes such as pair **3**·**6** and **3**·**7** (see Figure 5) may allow the development of association units with significantly enhanced sensitivity and responsiveness. For example, the supramolecular polymerization of difunctional unimers consisting of a spacer end-functionalized with strands **3** and **6** or **3** and **7** should result in a reasonably high degree of polymerization given the high strength of **3**·**6** or **3**·**7** ($K_a > 10^5$ in chloroform).[33] However, when a monofunctional chain stopper forms a fully matched pair with either one of the strands of the mutant duplex, the mutant duplex can be completely interrupted, leaving monomeric species in solution.

Besides directing the association of structural units, our duplexes allow the introduction of additional levels of controllability. Triggers that are responsive to external stimuli could be incorporated into the duplexes. As shown in Figure 12, the presence of the acid-responsive 2,6-diamidopyridine moiety[61] in **16** that is derived from oligoamide **4** should lead to a strand that is responsive to the presence or absence of acid. The addition of an acid would convert **16** from a linear to a curved conformation, which should interrupt the ability of such an oligoamide strand to pair with its complementary strand (Figure 12). The ^{1}H NMR spectra of duplex **3**·**16** in the presence of absence of acid in $CDCl_3$ demonstrated the expected conformational change and the association and dissociation of the corresponding H-bonded duplex.[52] Similarly, the presence of metal ion or the incorporation of photoisomerizable moiety into the design of our duplexes should lead to association units, chain stoppers, and analogous linking units allowing different means of controllability and with responsiveness toward various external stimuli.

Conclusions and Future Perspective

Information-encoded molecular duplexes with programmable H-bonding sequences and tunable binding strength have successfully reproduced some of the most important features of duplex DNAs. Mimicking some aspect of the encoding capability of DNAs, these duplexes possess instructing ability as demonstrated by the noncovalent conjugation of otherwise noncompatible polymer chains, the directed association of peptide strands, and the templation of chemical

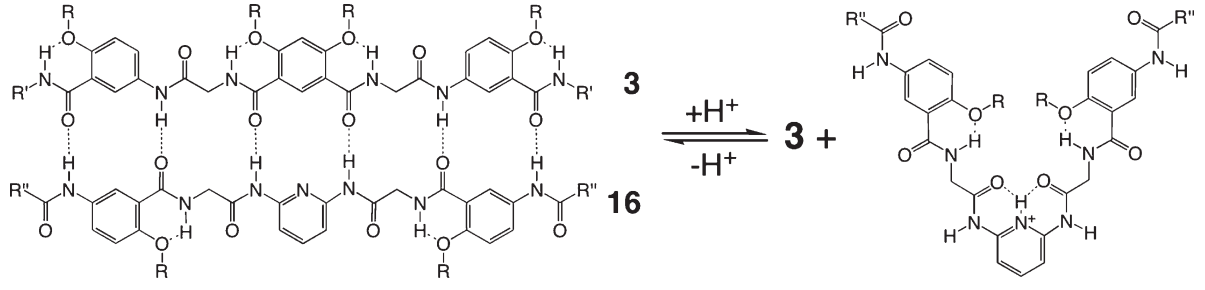


FIGURE 12. Incorporating an acid-responsive diamidopyridine unit leads to **16** that pairs with **3** into duplex **3**·**16**.

reactions. The incorporation of dynamic covalent bonding interactions has led to sequence-specifically formed, covalently linked structures under thermodynamic conditions. These specific dynamic covalent ligation units should open a new avenue to previously unavailable architectures. In the long run, such information-carrying systems, with convenient structural and functional modifiability, may provide a platform allowing constant optimization, leading to intelligent behavior, and finally, may reproduce the most important feature of DNAs, that is, self-replication and amplification.[62]

I would like to thank previous and current students, postdoctoral fellows and collaborators who have made this research possible. This research is supported by the National Science Foundation (Grants CBET-1036171 and CBET-1066947) and the Petroleum Research Fund, administered by the ACS (Grant 51048-ND7). Previous support provided by the National Institutes of Health and the Office of Naval Research are also acknowledged.

BIOGRAPHICAL INFORMATION

Bing Gong received his bachelor's degree in chemistry from Sichuan University in China in 1984. He attended the University of Chicago under the supervision of Professor David Lynn and received his Ph.D. in 1990. He then joined the laboratory of Professor Peter Schultz as a Damon Runyon-Walter Winchell Fund Postdoctoral Fellow at the University of California, Berkeley. He began his independent academic career in 1994 and is currently Professor of Chemistry at the State University of New York at Buffalo. His research interests include directed self-assembly, the folding of biomimetic unnatural molecules and the creation of nanoporous organic structures.

FOOTNOTES

*E-mail: bgong@buffalo.edu.
The authors declare no competing financial interest.

REFERENCES

1 Branden, C.; Tooze, J. *Introduction to Protein Structure*, 2nd ed.; Garland Science: New York, 1999.

2 Namba, K; Stubbs, G. Structure of Tobacco Mosaic Virus at 3.6-Å Resolution: Implications for Assembly. *Science* **1986**, *231*, 1401–1406.

3 Bruce, A.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walters, P. *Molecular Biology of the Cell*, 5th ed.; Garland Science: New York and London, 2007.

4 Jorgensen, W. L.; Pranata, J. Importance of Secondary Interactions in Triply Hydrogen Bonded Complexes: Guanine-Cytosine vs Uracil-2,6-Diaminopyridine. *J. Am. Chem. Soc.* **1990**, *112*, 2008–2010.

5 Pranata, J.; Wierschkem, S. G.; Jorgensen, W. L. OPLS Potential Functions for Nucleotide Bases - Relative Association Constants of Hydrogen-Bonded Base-Pairs in Chloroform. *J. Am. Chem. Soc.* **1991**, *113*, 2810–2819.

6 Murray, T. J.; Zimmerman, S. C. New Triply Hydrogen-Bonded Complexes with Highly Variable Stabilities. *J. Am. Chem. Soc.* **1992**, *114*, 4010–4011.

7 Fenlon, E. E.; Murray, T. J.; Baloga, M. H.; Zimmerman, S. C. Convenient Synthesis of 2-Amino-1,8-Naphthyridines, Building-Blocks for Host-Guest and Self-Assembling Systems. *J. Org. Chem.* **1993**, *58*, 6625–6628.

8 Corbin, P. S.; Zimmerman, S. C. Self-Association without Regard to Prototropy. A Heterocycle That Forms Extremely Stable Quadruply Hydrogen-Bonded Dimmers. *J. Am. Chem. Soc.* **1998**, *120*, 9710–9711.

9 Beijer, F. H.; Kooijman, H.; Spek, A. L.; Sijbesma, R. P.; Meijer, E. W. Self-Complementarity Achieved through Quadruple Hydrogen Bonding. *Angew. Chem., Int. Ed.* **1998**, *37*, 75–78.

10 Beijer, F. H.; Sijbesma, R. P.; Kooijman, H.; Spek, A. L.; Meijer, E. W. Strong Dimerization of Ureidopyrimidones via Quadruple Hydrogen Bonding. *J. Am. Chem. Soc.* **1998**, *120*, 6761–6769.

11 Folmer, B. J. B.; Sijbesma, R. P.; Kooijman, H.; Spek, A. L.; Meijer, E. W. Cooperative Dynamics in Duplexes of Stacked Hydrogen-Bonded Moieties. *J. Am. Chem. Soc.* **1999**, *121*, 9001–9007.

12 Sijbesma, R. P.; Beijer, F. H.; Brunsveld, L.; Folmer, B. J. B.; Hirschberg, J. H. K. K.; Lange, R. F. M.; Lowe, J. K. L.; Meijer, E. W. Reversible Polymers Formed from Self-Complementary Monomers Using Quadruple Hydrogen Bonding. *Science* **1997**, *278*, 1601–1604.

13 de Greef, T. F. A.; E. W. Meijer, E. W. Supramolecular Polymers. *Nature* **2008**, *434*, 171–173.

14 Todd, E. M.; Zimmerman, S. C. Supramolecular Star Polymers. Increased Molecular Weight with Decreased Polydispersity Through Self-Assembly. *J. Am. Chem. Soc.* **2007**, *129*, 14534–14535.

15 Li, Y.; Park, T. H.; Quansah, J. K.; Zimmerman, S. C. Synthesis of a Redox-Responsive Quadruple Hydrogen-Bonding Unit for Applications in Supramolecular Chemistry. *J. Am. Chem. Soc.* **2011**, *133*, 17118–17121.

16 Lehn, J. M. Programmed Chemical Systems: Multiple Subprograms and Multiple Processing/Expression of Molecular Information. *Chem.—Eur. J.* **2000**, *6*, 2097–2102.

17 Whitesides, G. M.; Simanek, E. E.; Mathias, J. P.; Seto, C. T.; Chin, D. N.; Mammen, M.; Gordon, D. M. Noncovalent Synthesis-Using Physical-Organic Chemistry to Make Aggregates. *Acc. Chem. Res.* **1995**, *28*, 37–44.

18 Kolotuchin, S. V.; Zimmerman, S. C. Self-Assembly Mediated by the Donor—Donor—Acceptor·Acceptor—Acceptor—Donor (DDA·AAD) Hydrogen-Bonding Motif: Formation of A Robust Hexameric Aggregate. *J. Am. Chem. Soc.* **1998**, *120*, 9092–9093.

19 Conn, M. M.; Rebek, J. Self-Assembling Capsules. *Chem. Rev.* **1997**, *97*, 1647–1668.

20 Orr, G. W.; Barbour, L. J.; Atwood, J. L. Controlling Molecular Self-Organization: Formation of Nanometer-Scale Spheres and Tubules. *Science* **1999**, *285*, 1049–1052.

21 Percec, V.; Cho, W. D.; Ungar, G. Increasing the Diameter of Cylindrical and Spherical Supramolecular Dendrimers by Decreasing the Solid Angle of Their Monodendrons via Periphery Functionalization. *J. Am. Chem. Soc.* **2000**, *122*, 10273–10281.

22 Vreekamp, R. H.; van Duynhoven, J. P. M.; Hubert, M.; Verboom, W.; Reinhoudt, D. N. Molecular Boxes Based on Calix[4]arene Double Rosettes. *Angew. Chem., Int. Ed.* **1996**, *35*, 1215–1218.

23 Sessler, J. L.; Wang, R. Z. Self-Assembly of an "Artificial Dinucleotide Duplex". *J. Am. Chem. Soc.* **1996**, *118*, 9808–9809.

24 Yang, J.; Fan, E. K.; Geib, S. J.; Hamilton, A. D. Hydrogen-Bonding Control of Molecular Self-Assembly - Formation of a 2 + 2 Complex in Solution and the Solid-State. *J. Am. Chem. Soc.* **1993**, *115*, 5314–5315.

25 Hartgerink, J. D.; Clark, T. D.; Ghadiri, M. R. Peptide Nanotubes and Beyond. *Chem.—Eur. J.* **1998**, *4*, 1367–1372.

26 Wu, A. X.; Chakraborty, A.; Fettinger, J. C.; Flowers, R. A.; Isaacs, L. Molecular Clips That Undergo Heterochiral Aggregation and Self-Sorting. *Angew. Chem., Int. Ed.* **2002**, *41*, 4028–4031.

27 Fenniri, H.; Deng, B. L.; Ribbe, A. E. Helical Rosette Nanotubes with Tunable Chiroptical Properties. *J. Am. Chem. Soc.* **2002**, *124*, 11064–11072.

28 Gooch, A.; McGhee, A. M.; Pellizzaro, M. L.; Lindsay, C. I.; Wilson, A. J. Substituent Control over Dimerization Affinity of Triply Hydrogen Bonded Heterodimers. *Org. Lett.* **2011**, *13*, 240–243.

29 Zimmerman, S. C.; Corbin, P. S. Heteroaromatic Modules for Self-Assembly Using Multiple Hydrogen Bonds. *Struct. Bonding (Berlin)* **2000**, *96*, 63–94.

30 Gong, B. Specifying Non-Covalent Interactions: Sequence-Specific Assembly of Hydrogen-Bonded Molecular Duplexes. *Synlett* **2001**, 582–589.

31 Gong, B.; Yan, Y. F.; Zeng, H. Q.; Skrzypczak-Jankunn, E.; Kim, Y. W.; Zhu, J.; Ickes, H. A New Approach for the Design of Supramolecular Recognition Units: Hydrogen-Bonded Molecular Duplexes. *J. Am. Chem. Soc.* **1999**, *121*, 5607–5608.

32 Zeng, H. Q.; Miller, R. S.; Flowers, R. A.; Gong, B. A Highly Stable, Six-Hydrogen-Bonded Molecular Duplex. *J. Am. Chem. Soc.* **2000**, *122*, 2635–2644.

33 Zeng, H. Q.; Ickes, H.; Flowers, R. A.; Gong, B. Sequence Specificity of Hydrogen-Bonded Molecular Duplexes. *J. Org. Chem.* **2001**, *66*, 3574–3583.

34 Zeng, H. Q.; Yang, X. W.; Flowers, R. A.; Gong, B. A Noncovalent Approach to Antiparallel β-Sheet Formation. *J. Am. Chem. Soc.* **2002**, *124*, 2903–2910.

35 Yang, X. W.; Martinovic, S.; Smith, R. D.; Gong, B. Duplex Foldamers from Assembly-Induced Folding. *J. Am. Chem. Soc.* **2003**, *125*, 9932–9933.

36 Zeng, H. Q.; Yang, X. W.; Brown, A. L.; Martinovic, S.; Smith, R. D.; Gong, B. An Extremely Stable, Self-Complementary Hydrogen-Bonded Duplex. *Chem. Commun.* **2003**, 1556–1557.

37 Sanford, A. R.; Yamato, K.; Yang, X. W.; Yuan, L. H.; Han, Y. H.; Gong, B. Well-Defined Secondary Structures: Information-Storing Molecular Duplexes and Helical Foldamers Based on Unnatural Peptide Backbones. *Eur. J. Biochem.* **2004**, *271*, 1416–1425.

38 Yang, X. W.; Hua, F. J.; Yamato, K.; Ruckenstein, E.; Gong, B.; Kim, W.; Ryu, C. Y. Supramolecular AB Diblock Copolymers. *Angew. Chem., Int. Ed.* **2004**, *43*, 6471–6474.

39 Yang, X. W.; Gong, B. Template-Assisted Cross Olefin Metathesis. *Angew. Chem., Int. Ed.* **2005**, *44*, 1352–1356.

40 Bialecki, J. B.; Yuan, L. H.; Gong, B. A Branched, Hydrogen-Bonded Heterodimer: A Novel System for Achieving High Stability and Specificity. *Tetrahedron* **2007**, *63*, 5460–5469.

41 Zhang, P. H.; Chu, H. Z.; Li, X. H.; Feng, W.; Deng, P. C.; Yuan, L. H.; Gong, B. Alternative Strategy for Adjusting the Association Specificity of Hydrogen-Bonded Duplexes. *Org. Lett.* **2011**, *13*, 54–57.

42 Goodwin, J. T.; Lynn, D. G. Template-Directed Synthesis - Use of a Reversible-Reaction. *J. Am. Chem. Soc.* **1992**, *114*, 9197–9198.

43 Berl, V.; Huc, I.; Khoury, R. G.; Krische, M. J.; Lehn, J. M. Interconversion of Single and Double Helices Formed from Synthetic Molecular Strands. *Nature* **2000**, *407*, 720–723.

44 Bisson, A. P.; Carver, F. J.; Eggleston, D. S.; Haltiwanger, R. C.; Hunter, C. A.; Livingstone, D. L.; McCabe, J. F.; Rotger, C.; Rowan, A. E. Synthesis and Recognition Properties of Aromatic Amide Oligomers: Molecular Zippers. *J. Am. Chem. Soc.* **2000**, *122*, 8856–8868.

45 Gabriel, G. J.; Iverson, B. L. Aromatic Oligomers That Form Hetero Duplexes in Aqueous Solution. *J. Am. Chem. Soc.* **2002**, *124*, 15174–15175.

46 Zhao, X.; Wang, X. Z.; Jiang, X. K.; Chen, Y. Q.; Li, Z. T.; Chen, G. J. Hydrazide-Based Quadruply Hydrogen-Bonded Heterodimers, Structure, Assembling Selectivity, and Supramolecular Substitution. *J. Am. Chem. Soc.* **2003**, *125*, 15128–15139.

47 Jiang, H.; Maurizot, V.; Huc, I. Double versus Single Helical Structures of Oligopyridine-Dicarboxamide Strands. Part 1: Effect of Oligomer Length. *Tetrahedron* **2004**, *60*, 10029–10038.

48 Gong, H. G.; Krische, M. J. Duplex Molecular Strands Based on the 3,6-Diaminopyridazine Hydrogen Bonding Motif: Amplifying Small-Molecule Self-Assembly Preferences through Preorganization and Iterative Arrangement of Binding Residues. *J. Am. Chem. Soc.* **2005**, *127*, 1719–1725.

49 Tanaka, Y.; Katagiri, H.; Furusho, Y.; Yashima, E. A Modular Strategy to Artificial Double Helices. *Angew. Chem., Int. Ed.* **2005**, *44*, 3867–3870.

50 Yang, Y.; Yang, Z. Y.; Yi, Y. P.; Xiang, J. F.; Chen, C. F.; Wan, L. J.; Shuai, Z. G. Helical Molecular Duplex Strands: Multiple Hydrogen-Bond-Mediated Assembly of Self-Complementary Oligomeric Hydrazide Derivatives. *J. Org. Chem.* **2007**, *72*, 4936–4946.

51 Levin, S.; Nowick, J. S. An Artificial β-Sheet That Dimerizes through Parallel β-Sheet Interactions. *J. Am. Chem. Soc.* **2007**, *129*, 13043–13048.

52 Unpublished data.

53 Sontjens, S. H. M.; Sijbesma, R. P.; van Genderen, M. H. P.; Meijer, E. W. Stability and Lifetime of Quadruply Hydrogen Bonded 2-Ureido-4[1H]-Pyrimidinone Dimers. *J. Am. Chem. Soc.* **2000**, *122*, 7487–7493.

54 Lindsey, J. S. Self-Assembly in Synthetic Routes to Molecular Devices - Biological Principles and Chemical Perspectives - A Review. *New J. Chem.* **1991**, *15*, 153–180.

55 Hua, F. J.; Yang, X. W.; Gong, B.; Ruckenstein, E. Preparation of Oligoamide-Ended Poly(Ethylene Glycol) and Hydrogen-Bonding-Assisted Formation of Aggregates and Nanoscale Fibers. *J. Polym. Sci., Part A: Polym. Chem.* **2005**, *43*, 1119–1128.

56 Li, M. F.; Yang, X. W.; Ryu, C. R.; Gong, B. Molecular Zippers for Preparing Supramolecular and Dynamic Covalent Block Copolymers. *Polym. Prepr.* **2005**, *46*, 1128–1129.

57 Rowan, S. J.; Cantrill, S. J.; Cousins, G. R. L.; Sanders, J. K. M.; Stoddart, J. F. Dynamic Covalent Chemistry. *Angew. Chem., Int. Ed.* **2002**, *41*, 898–952.

58 Cao, R. K.; Zhou, J. J.; Wang, W.; Feng, W.; Li, X. H.; Zhang, P. H.; Deng, P. C.; Yuan, L. H.; Gong, B. Oligoamide Duplexes as Organogelators. *Org. Lett.* **2010**, *12*, 2958–2961.

59 Li, M. F.; Yamato, K.; Ferguson, J. S.; Gong, B. Sequence-Specific Association in Aqueous Media by Integrating Hydrogen Bonding and Dynamic Covalent Interactions. *J. Am. Chem. Soc.* **2006**, *128*, 12628–12629.

60 Li, M. F.; Yamato, K.; Ferguson, J. S.; Singarapu, K. K.; Szyperski, T.; Gong, B. Sequence-Specific, Dynamic Covalent Crosslinking in Aqueous Media. *J. Am. Chem. Soc.* **2008**, *130*, 491–500.

61 Geib, S. J.; Hirst, S. C.; Vicent, C.; Hamilton, A. D. Molecular Recognition and the Design of Solid-State Structures - Protonation-Induced Conformational Change and Self-Assembly of 2,6-Diamidopyridinium Phosphates. *Chem. Commun.* **1991**, 1283–1285.

62 Rebek, J. Molecular Recognition and Self-Replication. *J. Mol. Recognit.* **1992**, *5*, 83–88.