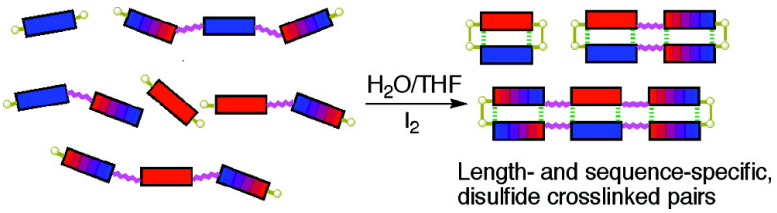
Article

Sequence-Specific, Dynamic Covalent Crosslinking in Aqueous Media

Minfeng Li, Kazuhiro Yamato, Joseph S. Ferguson, Kiran Kumar Singarapu, Thomas Szyperski, and Bing Gong

J. Am. Chem. Soc., **2008**, 130 (2), 491-500 • DOI: 10.1021/ja072567m



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 3 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 12/20/2007

Sequence-Specific, Dynamic Covalent Crosslinking in Aqueous Media

Minfeng Li, Kazuhiro Yamato, Joseph S. Ferguson, Kiran Kumar Singarapu, Thomas Szyperski, and Bing Gong*

Department of Chemistry, University at Buffalo, The State University of New York, Buffalo, New York 14260

Received April 12, 2007; E-mail: bgong@chem.buffalo.edu

Abstract: This article describes an associating system that integrates the specificity of multiple hydrogen bonding and the strength of dynamic covalent interactions. Linear oligoamides that sequence-specifically pair into H-bonded duplexes in nonpolar solvents were modified with *S*-trityl groups, allowing the reversible formation of disulfide bonds. The disulfide-crosslinking reactions of oligoamides capable of pairing via two, four, and six intermolecular H-bonds, along with several control strands, were examined using ESI, MALDI-TOF, reverse phase HPLC, and two-dimensional NMR. Results from these studies demonstrate that this system possesses both the high fidelity of multiply H-bonded assemblies and the high stability of covalent interaction, leading to the sequence-specific crosslinking of complementary oligoamides in not only nonpolar (methylene chloride) solutions but also highly competitive (aqueous) media. Experiments were designed to systematically probe the mechanism behind the specific formation of the sequence-matched products, which revealed a thermodymically controlled process. Multiple pairs in the same solution were crosslinked in a sequence-specific fashion. In addition, a length-dependent selectivity was also observed. Thus, oligoamides with different lengths or sequences did not crosslink into mismatched products. As few as two H-bonds is sufficient to bias the specific formation of the crosslinked product in aqueous media, suggesting that associating units with tunable sizes, high stability, and high specificity can be conveniently designed. The combination of H-bonding and dynamic covalent interactions represents a new, generalizable strategy for developing highly specific molecular associating units that are stable in a wide variety of media. These associating units will greatly facilitate the construction of various structures with many applications.

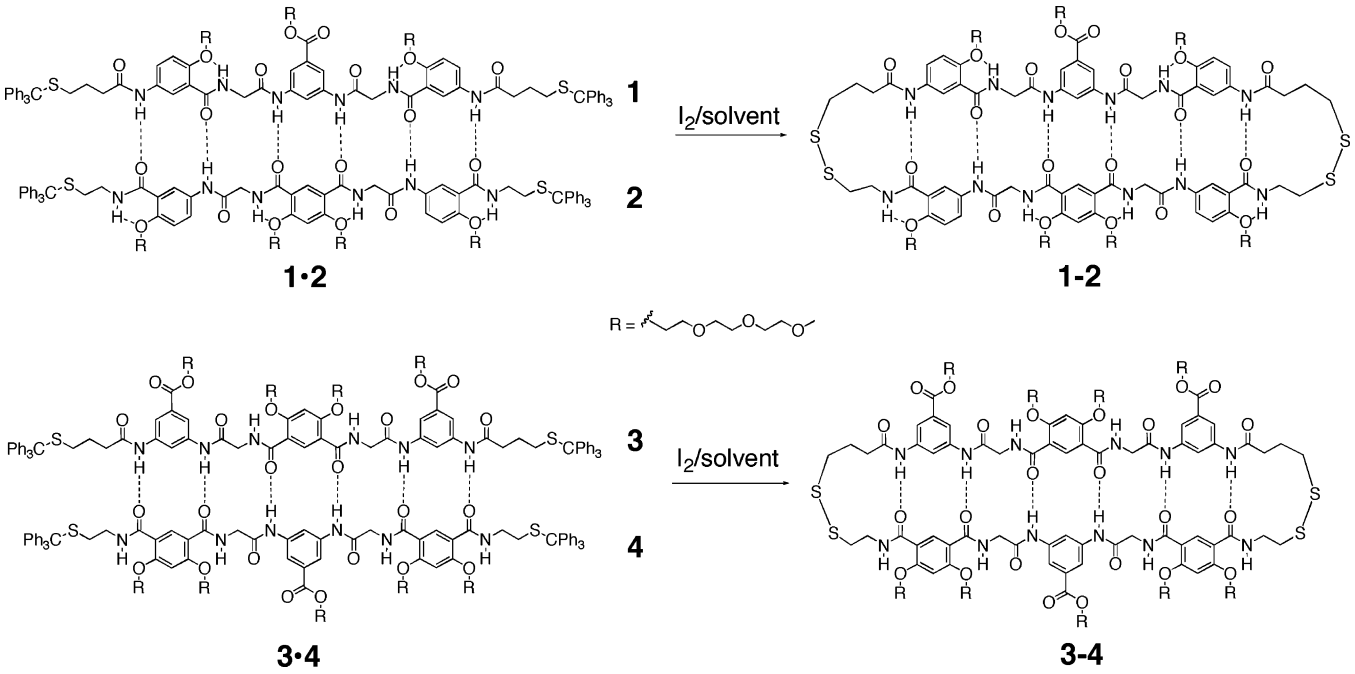
Introduction

The cooperative action of multiple noncovalent forces leads to highly specific recognition events that result in thermodynamically stable assemblies.[1] A systematic approach to controlling self-assembly involves the design of associating units that store and retrieve information.[2] Early investigation on molecular association of nucleobases by Jorgensen and Zimmerman has led to H-bonded pairs for specifying intermolecular interactions.[3] Subsequently, the groups of Zimmerman,[4] Meijer, and Sijbesma[5] reported highly stable heterocyclic complexes with arrays of H-bond donors (D) and acceptors (A). Many other groups[6−18] also reported H-bonded assemblies consisting of multiple components.

Since our first report,[19] we have developed a systematic approach for specifying intermolecular association.[20] Our system involves a class of molecular recognition units with high stability and tunable sequence specificity. This system is based on the recombination of amide O and H atoms of oligoamide strands consisting of readily available building blocks. Oligoamides carrying different combinations (sequences) of H-bond donors and acceptors have been generated. The resultant molecular strands form highly stable duplexes via hydrogen-bonding interactions between backbone amide O and H atoms. The association of these oligoamide strands was found to have high fidelity

(1) (a) Lindsey, J. S. *New J. Chem.* **1991**, *15*, 153. (b) Leininger, S.; Olenyuk, B.; Stang, P. J. *Chem. Rev.* **2000**, *100*, 853. (c) Prins, L. J.; Reinhoudt, D. N.; Timmerman, P. *Angew. Chem., Int. Ed.* **2001**, *40*, 2383. (d) Fujita, M.; Umemoto, K.; Yoshizawa, M.; Fujita N.; Kusukawa, T.; Biradha, K. *Chem. Commun.* **2001**, 509.
(2) Zimmerman, S. C.; Corbin, P. S. *Struct. Bonding* **2000**, *96*, 63.
(3) (a) Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008. (b) Pranata, J.; Wierschkem, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810. (c) Murray, T. J.; Zimmerman, S. C. *J. Am. Chem. Soc.* **1992**, *114*, 4010.
(4) (a) Corbin, P. S.; Zimmerman, S. C. *J. Am. Chem. Soc.* **1998**, *120*, 9710. (b) Park, T.; Todd, E. M.; Nakashima, S.; Zimmerman, S. C. *J. Am. Chem. Soc.* **2005**, *127*, 18133.
(5) (a) Folmer, B. J. B.; Sijbesma, R. P.; Kooijman, H.; Spek, A. L. Meijer, E. W. *J. Am. Chem. Soc.* **1999**, *121*, 9001. (b) Sijbesma, R. P.; Meijer, E. W. *Chem. Commun.* **2003**, 5.
(6) Lehn, J. M. *Chem.−Eur. J.* **2000**, *6*, 2097.
(7) Whitesides, G. M.; Simanek, E. E.; Mathias, J. P.; Seto, C. T.; Chin, D. N.; Mammen, M.; Gordon, D. M. *Acc. Chem. Res.* **1995**, *28*, 37.
(8) Kolotuchin, S. V.; Zimmerman, S. C. *J. Am. Chem. Soc.* **1998**, *120*, 9092.
(9) Conn, M. M.; Rebek, J. *Chem. Rev.* **1997**, *97*, 1647.
(10) Orr, G. W.; Barbour, L. J.; Atwood, J. L. *Science* **1999**, *285*, 1049.
(11) Percec, V.; Cho, W. D.; Ungar, G. *J. Am. Chem. Soc.* **2000**, *122*, 10273.
(12) Vreekamp, R. H.; van Duynhoven, J. P. M.; Hubert, M.; Verboom, W.; Reinhoudt, D. N. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1215.
(13) Sessler, J. L.; Wang, R. Z. *J. Am. Chem. Soc.* **1996**, *118*, 9808.
(14) Yang, J.; Fan, E. K.; Geib, S. J.; Hamilton, A. D. *J. Am. Chem. Soc.* **1993**, *115*, 5314.
(15) Archer, E. A.; Krische, M. J. *J. Am. Chem. Soc.* **2002**, *124*, 5074.
(16) Hartgerink, J. D.; Clark, T. D.; Ghadiri, M. R. *Chem.−Eur. J.* **1998**, *4*, 1367.
(17) Wu, A. X.; Chakraborty, A.; Fettinger, J. C.; Flowers, R. A.; Isaacs, L. *Angew. Chem., Int. Ed.* **2002**, *41*, 4028.
(18) Fenniri, H.; Deng, B. L.; Ribbe, A. E. *J. Am. Chem. Soc.* **2002**, *124*, 11064.
(19) Gong, B.; Yan, Y.; Zeng, H.; Skrzypczak-Jankunn, E.; Kim, Y. W.; Zhu, J.; Ickes, H. *J. Am. Chem. Soc.* **1999**, *121*, 5607.

Scheme 1



in a nonpolar solvent such as chloroform: a strand specifically pairs with another strand bearing a complementary sequence.

Using the H-bonded duplexes as noncovalent associating units, we developed a system of supramolecular AB diblock copolymers.[21a,b] Incompatible polymer chains (e.g., PS and PEG) end-modified with single H-bonding strands that sequence-specifically paired into hetero-duplexes were mixed in a common solvent. Studies by ^{1}H NMR and GPC revealed that the modified PS and PEG blocks quantitatively associated via H-bonding into the corresponding supramolecular diblock copolymers. In bulk, these block copolymers demonstrated microphase separation, leading to nanodomains typical of covalent diblock copolymers. The versatility of our H-bonded duplexes as highly specific associating units was also demonstrated by their serving as templates for inducing the formation of β-sheets when attached to peptide chains[21c] and for directing the cross metathesis of attached olefin moieties.[21d]

In spite of their proven specificity and tunable stabilities for assembling various molecular components, our H-bonded duplexes, like most other H-bonded complexes, are stable only in nonpolar solvents. Their intolerance to polar media such as water hampers many applications, particularly those involving biological conditions. Although several systems with microenvironments that promote H-bonding in polar media were known,[22] in competitive environments, particularly in aqueous solutions, the association of artificial molecular components based on H-bonding still represents a largely unsolved fundamental problem. In contrast to designed systems, all water-soluble natural self-assembling systems are stabilized by the cooperative interaction of multiple noncovalent forces. One of the best known examples involves DNA duplexes, which are stabilized by H-bonding, aromatic stacking, and other noncovalent interactions.

Using similar strategies found in nature, one possible solution to constructing stable and specific assemblies in competitive media involves the design of molecular components capable of simultaneously specifying H-bonding and other noncovalent forces. However, the successful design of such molecular components still represents a daunting challenge. Instead of contemplating the design of molecular structures that can assemble based on the synergistic action of multiple noncovalent forces, we decided to integrate the superb specificity of H-bond arrays and the strength of covalent interactions into the same duplex. Although most covalent interactions are irreversible and can only lead to structures that are formed under kinetically controlled conditions, a few reversible (dynamic) covalent bond formation reactions do exist[23] and have in fact attracted intense interest in recent years in thermodynamically controlled covalent synthesis and in the creation of dynamic combinatorial libraries.[24] Therefore, the integration of H-bonding and dynamic covalent interaction may lead to a simplified, easily manageable system (Scheme 1).

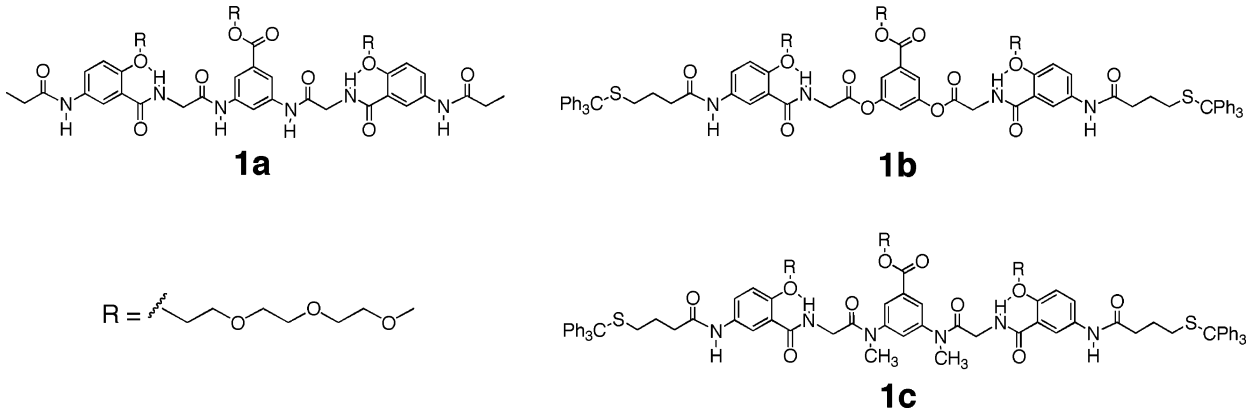
(20) (a) Zeng, H.; Miller, R. S.; Flowers, R. A.; Gong, B. *J. Am. Chem. Soc.* **2000**, *122*, 2635. (b) Zeng, H.; Ickes, H.; Flowers, R. A.; Gong, B. *J. Org. Chem.* **2001**, *66*, 3574. (c) Gong, B. *Synlett* **2001**, 582. (d) Yang, X. W.; Martinovic, S.; Smith, R. D.; Gong, B. *J. Am. Chem. Soc.* **2003**, *125*, 9932. (e) Zeng, H. Q.; Yang, X. W.; Brown, A. L.; Martinovic, S.; Smith, R. D.; Gong, B. *Chem. Commun.* **2003**, 1556−1557. (f) Sanford, A.; Yamato, K.; Yang, X. W.; Yuan, L. H.; Han, Y. H.; Gong, B. *Eur. J. Biochem.* **2004**, *271*, 1416. (g) Li, S. M.; Gong, B.; Guo, H. *Chem. Phys. Lett.* **2005**, *410*, 264. (h) Jiang, H. H.; Dolan, A. R.; Li, M. L.; Moy, M. A.; Gong, B.; Wood, T. D. *Int. J. Mass Spectrom.* **2007**, *261*, 13. (i) Gong, B. *Polym. Int.* **2007**, *56*, 436.

(21) (a) Yang, X. W.; Hua, F. J.; Yamato, K.; Ruckenstein, E.; Gong, B.; Kim, W.; Ryu, C. Y. *Angew. Chem., Intl. Ed.* **2004**, *43*, 6471. (b) Hua, F. J.; Yang, X. W.; Gong, B.; Ruckenstein, E. *J. Polym. Sci., Part A: Polym. Chem.* **2005**, *43*, 1119−1128. (c) Zeng, H. Q.; Yang, X. W.; Flowers, R. A.; Gong, B. *J. Am. Chem. Soc.* **2002**, *124*, 2903−2910. (d) Yang, X. W.; Gong, B. *Angew. Chem., Int. Ed.* **2005**, *44*, 1352.

(22) (a) Nowick, J. S.; Chen, J. S. *J. Am. Chem. Soc.* **1992**, *114*, 1107−1108. (b) Fan, E.; Vanarman, S. A.; Kincaid, S.; Hamilton, A. D. *J. Am. Chem. Soc.* **1993**, *115*, 369−370. (c) Kato, Y.; Conn, M. M.; Rebek, J. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1208−1212. (d) Torneiro, M.; Still, W. C. *J. Am. Chem. Soc.* **1995**, *117*, 5887−5888. (e) Paleos, C. M.; Tsiourvas, D. *Adv. Mater.* **1997**, *9*, 695−710. (f) Ariga, K.; Kunitake, T. *Acc. Chem. Res.* **1998**, *31*, 371−378. (g) Brunsveld, L.; Vekemans, J.; Hirschberg, J.; Sijbesma, R. P.; Meijer, E. W. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4977−4982. (h) Schmuck, C.; Wienand, W. *J. Am. Chem. Soc.* **2003**, *125*, 452.

(23) Furlan, R. L. E.; Otto, S.; Sanders, J. K. M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4801.

Scheme 2



Indeed, our preliminary study[25] based on MALDI, HPLC, and preparative scale reaction demonstrated that, by introducing *S*-trityl groups capable of reversibly forming disulfide bonds in the presence of iodine,[26] oligoamide strands **1** and **2** (0.5 mM each), with their complementary H-bonding sequences, were specifically crosslinked into **1**−**2** in both organic and aqueous media in nearly quantitative yields. The sequence-specific formation of **1**−**2** was shown by the absence of other crosslinked products when both strands were present in the same solution. Additional evidence supporting the high sequence specificity of this system was the that strand **3**, which has the same length as that of **1** or **2**, but a H-bonding sequence that is not complementary to that of **1** or **2**, could not interfere with the formation of **1**−**2** in aqueous media. The sequence dependence of the crosslinking process was further demonstrated by **3** and its complementary strand **4**. Although **3** failed to crosslink with either **1** or **2**, MALDI detected the crosslinked **3**−**4** as the only product from the 1:1 mixture of **3** and **4** (0.5 mM each) in aqueous solution containing iodine, the same outcome as that observed for **1** and **2**.[25,27] This result indicates that two different strands can be specifically crosslinked only when their H-bonding sequences are complementary to each other.

In this article, we would like to describe a detailed, systematic study on this new associating system. Specifically, we intended to probe: (1) The mechanism of the sequence-specific crosslinking in aqueous media. Is the mechanism based on a kinetically or thermodynamically controlled process? What major interaction is responsible for the observed sequence specificity of the crosslinking reactions? (2) The generality of this system. Can sequence-dependent crosslinking be extended to duplexes with different lengths? (3) Specificity of the duplexes. Do different pairs of complementary oligoamides interfere with each other's crosslinking? Answers to these questions should not only establish the mechanism of the H-bond-assisted crosslinking of these duplexes but should also prove the generality and efficiency of this associating system.

Results and Discussion

Mechanism of Sequence-Specific Crosslinking in Aqueous Media. The crosslinking of strands **1**, **2**, **3**, and **4** demonstrated an apparent correlation to H-bonding sequences in the formation of the crosslinked products. What is the mechanism behind the observed sequence-specificity? Does H-bond-mediated intermolecular association play any role for the observed sequence-specificity in aqueous media?

The sequence-specific crosslinking of two complementary oligoamide strands may involve one of two possible mechanisms. The first is a kinetic one, involving the intermolecular association of the two stands, which accelerates the rate of disulfide bond formation by increasing the effective molarity of the *S*-trityl groups. The other is a thermodynamic mechanism, based on the selective stabilization of the crosslinked product consisting of the sequence-matched strands. To distinguish these two mechanisms, control strands **1a**, **1b**, and **1c** were designed and prepared (Scheme 2).

To probe whether intermolecular association of complementary strands have any effect on the final product distribution, the crosslinking of **1** and **2** was examined in the presence of control strand **1a**. Sharing identical backbone and H-bonding sequence, strands **1a** and **1** should compete for H-bonding or aromatic stacking interactions with **2**. The only difference between **1** and **1a** is that the latter is incapable of forming crosslinked product with **2**. If the selective formation of **1**−**2** depended on the specific association of **1** and **2** via H-bonding or aromatic stacking interactions, the presence of **1a** should interfere with such association, which would slow down the rate-determining step of the disulfide bond formation reaction and compromise the sequence selectivity. Thus, a mixture containing **1**, **2**, and **1a** (0.5 mM each) in aqueous solution in the presence of iodine was examined using MALDI. As shown in Figure 1a, two peaks were detected in the MALDI spectrum, corresponding to the crosslinked **1**−**2** and strand **1c**. No other potential products from the self-cyclization or homo-dimerization of **1** or **2** were detected. Therefore, the formation of the crosslinked **1**−**2** was not affected to any detectable extent by **1a**. Based on this result, it can be concluded that the sequence

(24) (a) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J. L.; Sanders, J. K. M.; Otto, S. *Chem. Rev.* **2006**, *106*, 3652. (b) Saur, B.; Scopelliti, R.; Severin, K. *Chem.*−*Eur. J.* **2006**, *12*, 1058. (c) Rowan, S. J.; Cantrill, S. J.; Cousins, G. R. L.; Sanders, J. K. M.; Stoddart, J. F. *Angew. Chem., Int. Ed.* **2002**, *41*, 898−952. (d) Huc, I.; Lehn, J.-M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2106. (e) Cousins, G. R. L.; Poulsen, S.-A.; Sanders, J. K. M. *Chem. Commun.* **1999**, 1575. (f) Lehn, J.-M. *Chem.*−*Eur. J.* **1999**, *5*, 2455. (g) Polyakov, V. A.; Nelen, M. I.; Nazarpack-Kandlousy, N.; Ryabov, A. D.; Eliseev, A. V. *J. Phys. Org. Chem.* **1999**, *12*, 357. (h) Giger, T.; Wigger, M.; Audetat, S.; Benner, S. A. *Synlett* **1998**, 688. (i) Swann, P. G.; Casanova, R. A.; Desai, A.; Frauenhoff, M. M.; Urbancic, M.; Slomczynska, U.; Hopfinger, A. J.; Le, Breton, G. C.; Venton, D. L. *Biopolymers* **1996**, *40*, 617.

(25) Li, M. F.; Yamato, K.; Ferguson, J. S.; Gong, B. *J. Am. Chem. Soc.* **2006**, *128*, 12628.

(26) Kamber, B.; Hartmann, A.; Eisler, K.; Riniker, B.; Rink, H.; Sieber, P.; Rittel, W. *Helv. Chim. Acta* **1980**, *63*, 899.

(27) See Supporting Information for additional data.

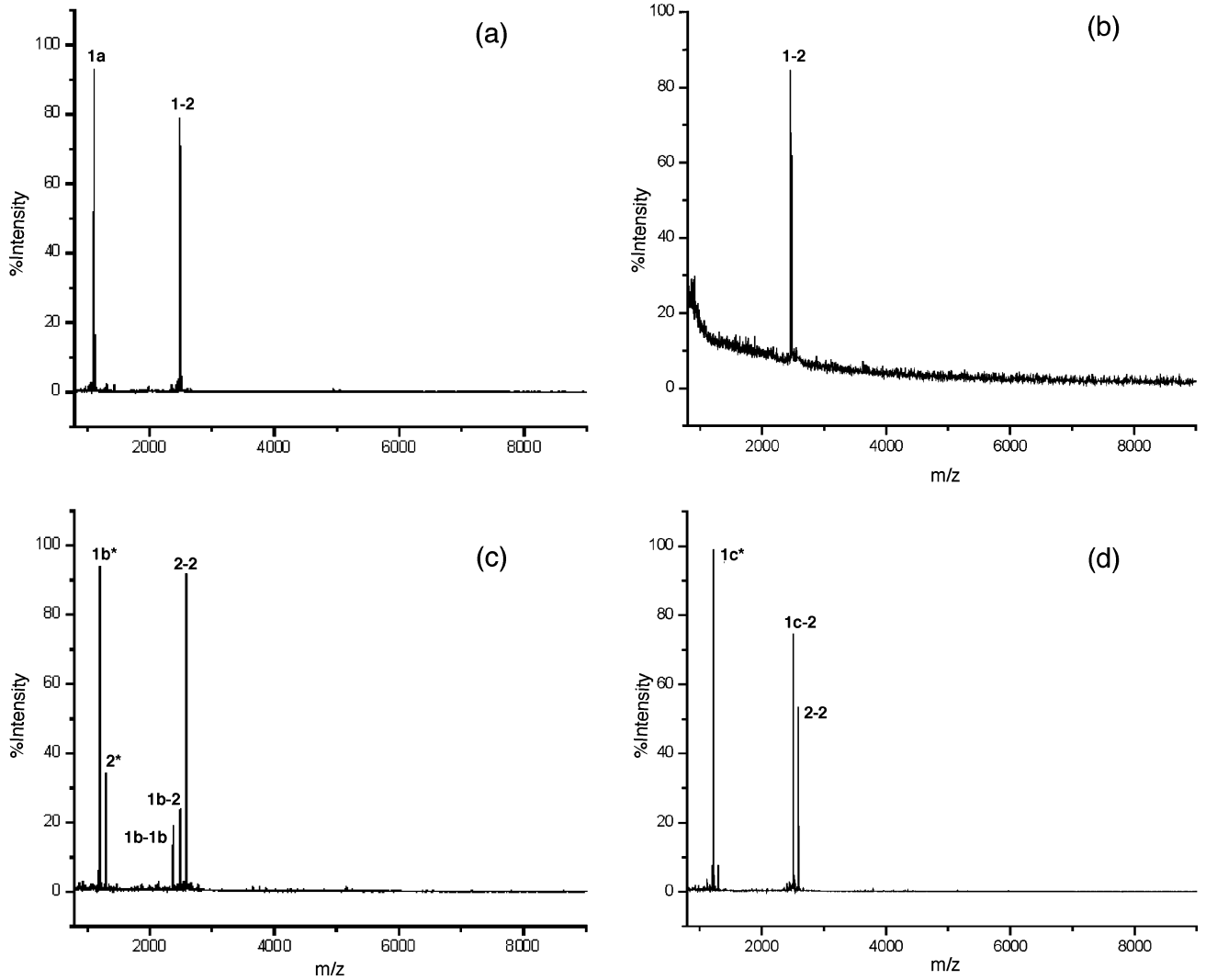


Figure 1. MALDI spectra of (a) the 1:1:1 mixture of **1**, **1a**, and **2** (0.5 mM each); (b) the 1:1 mixture of **1** and **2** (0.05 mM each); (c) the 1:1 mixture of **1b** and **2** (0.5 mM each); and (d) the 1:1 mixture of **1c** and **2** (0.5 mM each) in H_2O/THF (9/1, v/v)] in the presence of iodine. (*m/z* values: 1108.9 [**1a** + Na^+], 2482.5 [**1**−**2** + Na^+], 1201.2 [**1b*** + Na^+], 1305.3 [**2*** + Na^+], 2380.2 [**1b**−**1b** + Na^+], 2484.1 [**1b**−**2** + H^+], 2588.3 [**2**−**2** + Na^+], 1227.5 [**1c*** + Na^+], 2509.4 [**1c**−**2** + Na^+]).

specificity shown in the formation of **1**−**2** was not due to the intermolecular association of **1** and **2**.

The conclusion that intermolecular association of complementary strands has no effect on the observed sequence specificity was further verified by investigating the crossinking of **1** and **2** at different concentrations. If the formation of **1**−**2** were kinetically controlled, depending on the specific association of **1** and **2**, then product distribution would change with concentration. In this case, higher selectivity would be observed at higher concentrations. The crosslinking reactions of **1** and **2** were carried out in aqueous media at concentrations of 0.05 and 5 mM for each strand. These concentrations were 1/10 or 10 times of that adopted in the crosslinking reaction of **1** and **2** we reported.[25] At 0.05 mM, MALDI experiment indicated that **1**−**2** formed as the only product, and no other self-cyclized or homodimerized products were detected (Figure 1b). At a concentration (5 mM for each strand) that was 100 times higher, the crosslinked **1**−**2** again appeared in the MALDI spectrum as the only product.[27] Thus, within the concentration range from 0.5 to 5 mM, the formation of **1**−**2** did not show any concentration dependence.

Results from the above experiments have ruled out the possibility of a kinetic mechanism for the sequence-specific crosslinking of **1** and **2**, and similarly **3** and **4**, in aqueous media. Instead, these results are consistent with a thermodynamically controlled process, which involves the selective stabilization of the crosslinked product.

The observed H-bonding sequence specificity suggests that the selective formation of a crosslinked product is very likely to be resulted from the stabilization by intramolecular H-bonds between two complementary strands. These fully matched interstrand intramolecular H-bonds came into existence upon the formation of the disulfide bonds. However, given the strong competition from water molecules, whether intramolecular H-bonds could have any effect on the putative stabilization of the crosslinked products remains unclear. Thus, another possibility is that the observed sequence specificity does not have anything to do with H-bonding but instead arises from π-stacking interactions between the complementary strands of the crosslinked products. It is possible that, with their complementary H-bonding sequences, strands **1** and **2**, and **3** and **4**, also happen to have complementary surfaces that facilitate specific

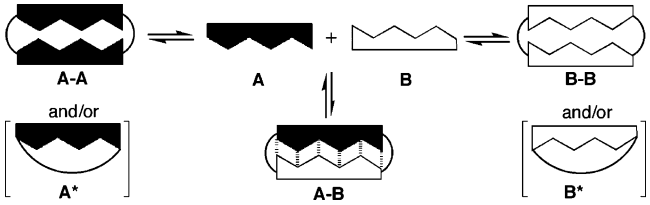


Figure 2. Two components bearing complementary H-bonding sequences can reversibly form disulfide-crosslinked products. **A**−**B** represents the most stable product due to the stabilization from its fully matched, interstrand (now intramolecular) H-bonds.

π-stacking interactions,[28] leading to the selective stabilization of the disulfide corsslinked products.

The role of H-bonding or aromatic stacking were investigated by examining the crosslinking of oligomers **1b** and **1c** with **2**. Strands **1b** and **1c** can be regarded as being derived from **1** by replacing the two central aniline NH groups of the latter with O atoms and N−Me groups respectively. These two control strands have similar electronic structures as that of **1** but carry fewer (four) H-bonding sites. Therefore, in aqueous media, if aromatic stacking, rather than H-bonding, played the major role in determining the relative stability of the crosslinked product, strand **1b** or **1c** should follow the behavior of **1** by pairing with **2**, leading to the corresponding disulfide-crosslinked **1b**−**2** or **1c**−**2** as the major product.

The 1:1 mixture of **1b** and **2** in aqueous media (0.5 mM each) was subjected to the same conditions for forming **1**−**2**. MALDI results showed that neither **1a** nor **1b** selectively crosslinked with **2**. In sharp contrast to the exclusive formation of **1**−**2** or **3**−**4**, multiple products were detected from the 1:1 mixture of **1b** and **2**, among which **1b***, formed from the self-cyclization of **1b**, and **2**−**2**, from the homodimerization of **2**, represented the major products (Figure 1a). Product **1b**−**2**, from the crosslinking of **1b** and **2**, only appeared as a minor peak. Similarly, when mixed together, **1c** and **2** could not be crosslinked exclusively into **1c**−**2** (Figure 1b). The observed distribution of products, that is, the self-cyclized **1c***, heterodimer **1c**−**2**, and homodimer **2**−**2**, can be regarded as being a statistical one.

Results from these experiments indicate that aromatic stacking interaction between complementary strands is unlikely the rate-determining factor responsible for the stabilization of the crosslinked product in aqueous media. Instead, the critical role played by interstrand H-bonding in the stabilization of products **1**−**2** (and similarly **3**−**4**) was supported by the high sequence specificity of the crosslinking reactions, as well as by the behavior of control strands **1b** and **1c** in the presence of **2**.

The above observations are consistent with a thermodynamically controlled process shown in Figure 2. Thus, when two strands, **A** and **B**, with complementary H-bonding sequences and termini that can be reversibly crosslinked, are present in the same solution under redox condition, products **A**−**A** (or self-cyclized **A***), **B**−**B** (or self-cyclized **B***), and **A**−**B**, may be generated. Among these products, **A**−**B** gains the most stabilization from the newly generated, complementary intramolecular H-bonds due to the formation of the two disulfide bonds. Thus, **A**−**B** represents the most stable combination, that is, the one with the lowest relative Gibbs free energy. The reversibility of the dynamic covalent crosslinking allows the equilibria as shown in Figure 2 to shift toward the formation of the most stable **A**−**B**. Therefore, the observed sequence specificity in the formation of the crosslinked pairs was due to the extra stabilization from the H-bonding interactions, which in turn were the result of the disulfide bond formation.

Generality of the Sequence-Dependent Crosslinking Process. In aqueous media, is the observed sequence-specific crosslinking limited to duplexes with six or more interstrand H-bonds? Could duplexes with shorter lengths, that is, fewer interstrand H-bonds, also be sequence-specifically crosslinked in polar solvents? To answer these questions, the crosslinking of pairs **5** and **6**, and **7** and **8**, which may associate with four and two H-bonds respectively, were treated with iodine in methlyene chloride and in water containing 10% THF.

ESI indicated that, in methylene chloride, both H-bonded pairs were crosslinked, forming the corresponding disulfide-linked covalent structures **5**−**6** and **7**−**8** as the overwhelmingly major products.[27] Given its relatively high stability in chloroform[20a] and, similarly, in methylene chloride, the presence of H-bonded **5**•**6** should facilitate the formation of the disulfide-crosslinked **5**−**6** by bringing the terminal *S*-trityl groups into close proximity. Interestingly, in spite of its rather low stability ($K_a \approx 25$ M^{-1} to 50 M^{-1} in chloroform),[20a] the doubly H-bonded **7**•**8** also showed sequence-selectivity, forming the crosslinked **7**−**8** as the dominant product. In this case, the self-dimerized **7**−**7** and **8**−**8** appeared only as minor products. The crosslinking of **5** and **6**, and **7** and **8**, was then carried out with iodine in aqueous media. Examining the reaction mixtures by ESI revealed the same results as those from methylene chloride.[27] In aqueous media, the crosslinked **5**−**6** and **7**−**8** were again observed as the clearly dominant products. The homo-dimerized products appeared as weak peaks in the ESI spectra of both pairs (Scheme 3).

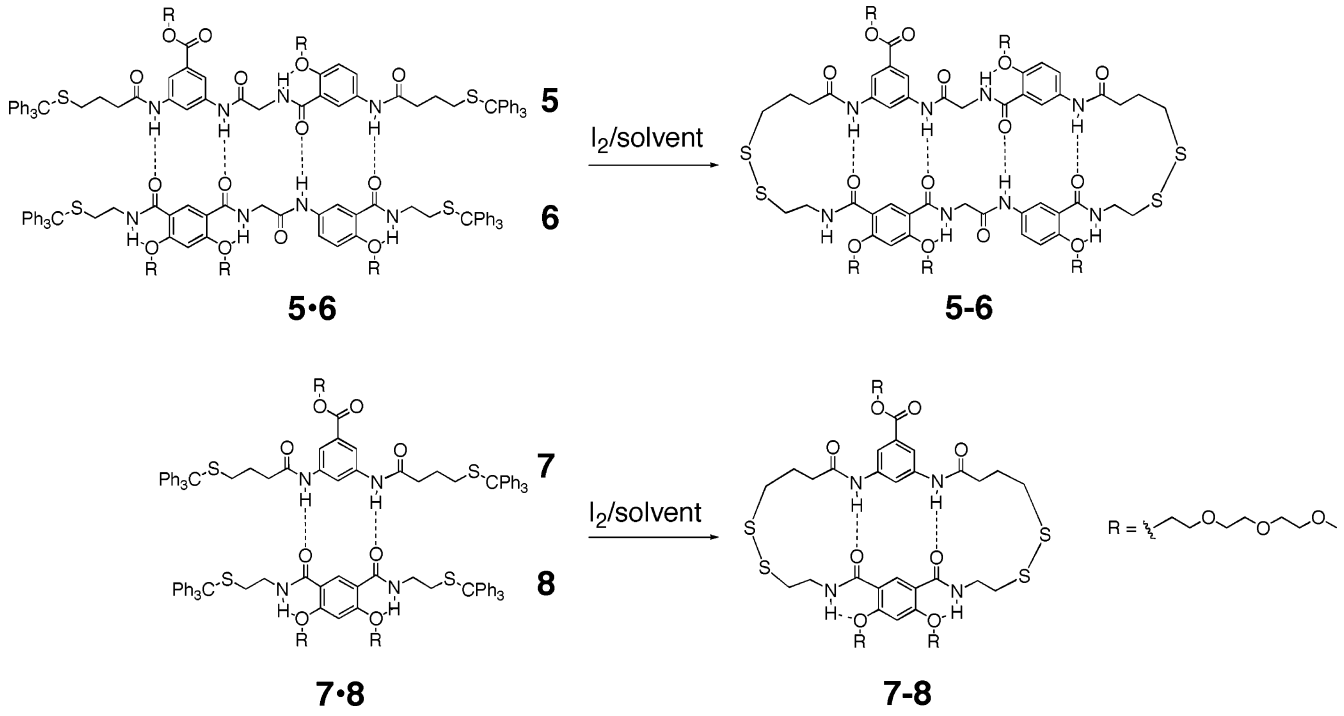
The dominance of the sequence-specifically paired **5**−**6** and **7**−**8** from aqueous media was confirmed by reverse-phase HPLC measurements. As shown in Figure 3, the sequence-matched **5**−**6** and **7**−**8** were detected as the dominant products. The self-dimerized **5**−**5** and **6**−**6** could barely be detected (Figure 3a) from the 1:1 mixture of **5** and **6**. In the 1:1 mixture of **7** and **8**, the self-dimerized **7**−**7** (Figure 3b), which appeared as a minor product, was probably due to a slight excess amount of **7** in the reaction mixture. The HPLC results are fully consistent with those from the ESI experiments, which demonstrated the reliability of the mass spectrometric method. These results indicated that as few as two intramolecular H-bonds could sufficiently stabilize the corresponding crosslinked product **7**−**8**, leading to the observed selectivity and efficiency.

It needs to be pointed out that, in the absence of partners with complementary H-bonding sequences, the individual oligoamides were found to self-cyclize or dimerize or both.[25,27] In these cases, the observed self-cyclized or dimerized products were obtained because of the enthalpic (i.e., more disulfide bonds per unit) and entropic (i.e., formation of more particles) advantages associated with their formation over that of longer oligomers and polymers.[29] When complementary strands were present, these products were completely or largely replaced by the ones with fully matched H-bonding sequences. These self-

(28) Cubberley, M. S.; Iverson, B. L. *J. Am. Chem. Soc*. **2001**, *123*, 7560.

(29) (a) Ercolani, G.; Mandolini, L.; Mencarelli, P.; Roelens, S. *J. Am. Chem. Soc*. **1993**, *115*, 3901. (b) Ercolani, G.; Mandolini, L.; Mencarelli, P.; Roelens, S. *J. Am. Chem. Soc*. **1993**, *115*, 3901. (c) ten Cate, A. T.; Sijbesma, R. P. *Macromol. Rapid Commun*. **2002**, *23*, 1094.

Scheme 3



cyclized or homo-dimerized products could not compete with the sequence-matched pairs because they do not possess fully matched H-bonds and thus could not gain the same level of stabilization available to the latter. This observation has provided additional evidence supporting the role of intramolecular H-bonding in stabilizing the sequence-matched products.

Crosslinking of Multiple Strands in the Same Solution. The above finding suggests that highly specific associating units with significantly reduced sizes (masses) can be easily designed based on this series of molecules. The availability of crosslinked duplexes of different lengths should further enhance the diversity of these molecular associating units. However, one potential problem is that a shorter duplex usually contain H-bonding sequences that overlap with part of those in a longer one. Will the selectivity of the crosslinking reactions be compromised if multiple strands with different lengths but partially overlapping H-bonding sequences are present in the same solution?

Figure 4 shows the results from MALDI-TOF experiments on four reaction mixtures containing two pairs, **1** and **2**, and **5** and **6** (Figure 4a), **1** and **2**, and **7** and **8** (Figure 4b), **5** and **6**, and **7** and **8** (Figure 4c), and three pairs, **1** and **2**, **5** and **6**, and **7** and **8** (Figure 4d), in aqueous media. In all four combinations examined, only crosslinked pairs consisting of components with matched lengths were detected.

The reaction mixtures were also examined using reverse phase HPLC, which provided results that paralleled those from the mass spectrometric experiments. The HPLC traces of two of the combinations are shown in Figure 5. Analyzing the solution containing **1**, **2**, **5**, and **6** in the presence of iodine led to the detection of the two length-matched products **1**−**2** and **5**−**6**

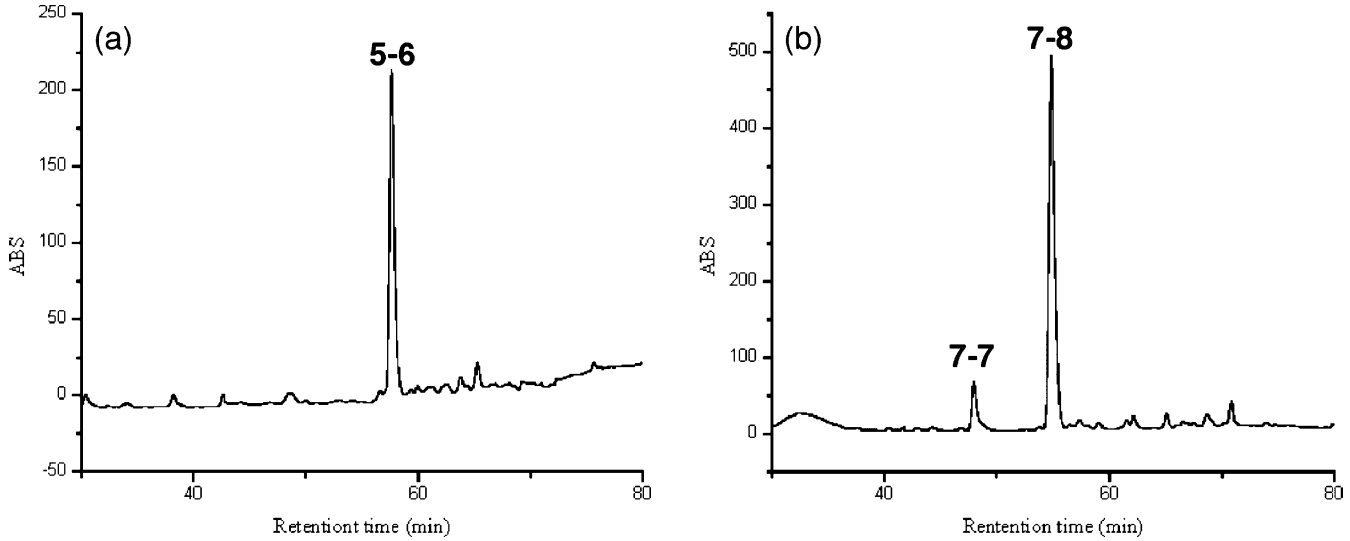


Figure 3. HPLC traces of an aliquot (15 μL) of the solutions of (a) the 1:1 mixture of **5** and **6**, and (b) the 1:1 mixture of **7** and **8** in H_2O/THF (9/1, v/v) in the presence of iodine (6 mM). Retention time: **5**−**6** (57.6 min); **7**−**7** (46.5 min); **7**−**8** (54.9 min).

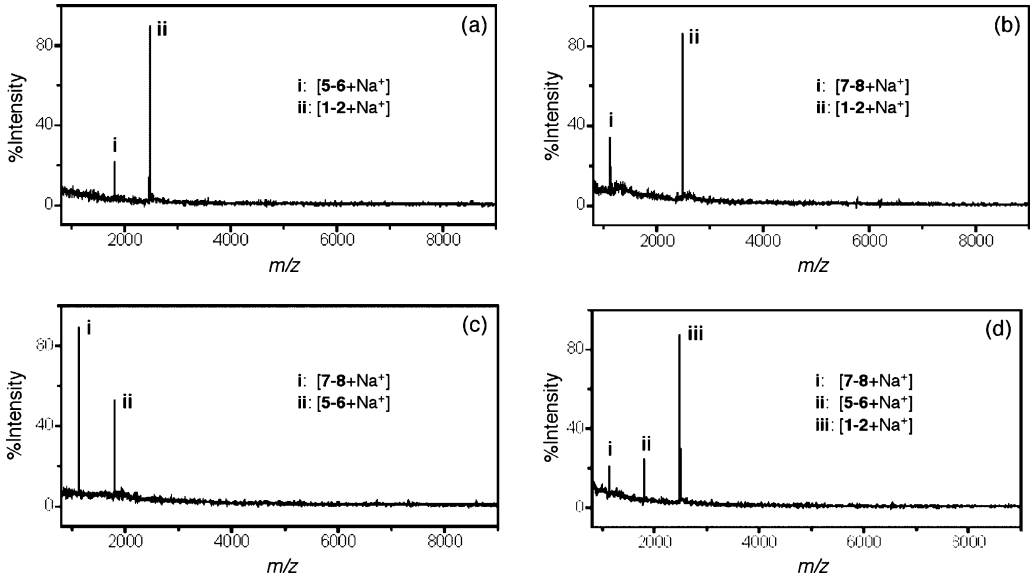


Figure 4. MALDI-TOF MS spectra of (a) the 1:1:1:1 mixture of **1**, **2**, **5**, and **6**; (b) the 1:1:1:1 mixture of **1**, **2**, **7**, and **8**; (c) the 1:1:1:1 mixture of **5**, **6**, **7**, and **8**; and (d) the 1:1:1:1:1:1 mixture of **1**, **2**, **5**, **6**, **7**, and **8** in H_2O/THF (9/1, v/v) in the presence of iodine. (*m*/*z* values: 2483.0 [**1**−**2** + Na^+], 1806.1 [**5**−**6** + Na^+], and 1129.0 [**7**−**8** + Na^+].)

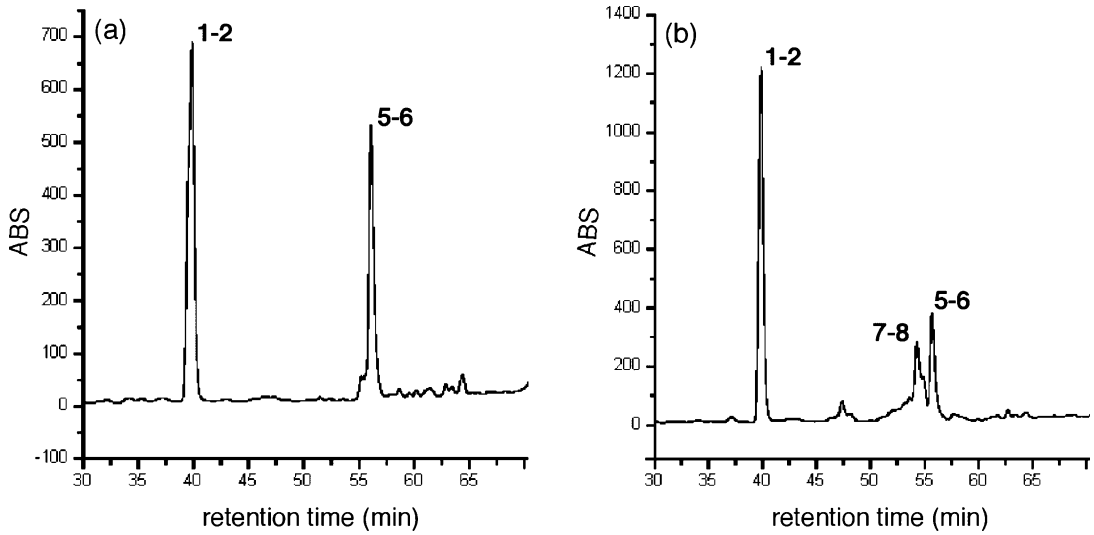


Figure 5. HPLC traces of an aliquot (15 μL) of (a) the 1:1:1:1 mixture of **1**, **2**, **5**, and **6** and (b) the 1:1:1:1:1:1 mixture of **1**, **2**, **5**, **6**, **7**, and **8** in H_2O/THF (9/1, v/v) in the presence of iodine (6 mM). Retention time: **1**−**2** (40.9 min in (a) and 40.6 in (b)); **5**−**6** (57.6 min); **7**−**8** (54.5 min).

(Figure 5a). Similarly, the three length-matched products **1**−**2**, **5**−**6**, and **7**−**8** were detected in the reaction mixture of **1**, **2**, **5**, **6**, **7**, and **8** in the presence of iodine. In these cases, no other product consisting of components with mismatched lengths was detected.

Given that the shorter strands have H-bonding sequences that partially overlap with those of the longer ones (e.g., **5** and **6** vs **1** and **2**, and **7** and **8** vs **5** and **6**), the high selectivity demonstrated by this associating system is rather dramatic, which indicated that the formation of the crosslinked products required component strands with both complementary sequences and matched lengths. This newly discovered length selectivity provides an additional strategy for designing molecular associating units with orthogonal specificities.

Persistence of Intramolecular H-Bonds in Aqueous Media. An inference based on the thermodynamically controlled mechanism is that the fully matched intramolecular H-bonds in the sequence-matched products should prevail in both nonpolar and aqueous media. NOESY and ROESY experiments were carried out on **7**−**8** in $CDCl_3$ and H_2O containing 20% THF-d_8. As shown in Figure 6a, in $CDCl_3$ in which the interstrand H-bonds are favored, strong NOEs were detected between protons *a* and *d*, and *a* and *f*, supporting a disulfide-crosslinked, intramolecularly H-bonded structure for the crosslinked **7**−**8**. Examining the 1:1 mixture of **7** and **8** in $CDCl_3$ under the same condition failed to reveal any NOEs between **7** and **8**, which suggested that, in the absence of the disulfide crosslinks, the weakly H-bonded **7**•**8** did not have a sufficient stability and a long enough lifetime to allow its detection on the time scale of the NMR experiment. Although the presence of intramolecular H-bonds in **7**−**8** is expected in the nonpolar $CDCl_3$, the same interstrand NOEs, that is, those between *a* and *d*, and *a* and *f*, were also detected in H_2O containing 20% THF-d_8 (Figure 6b). That the same NOEs were detected in both

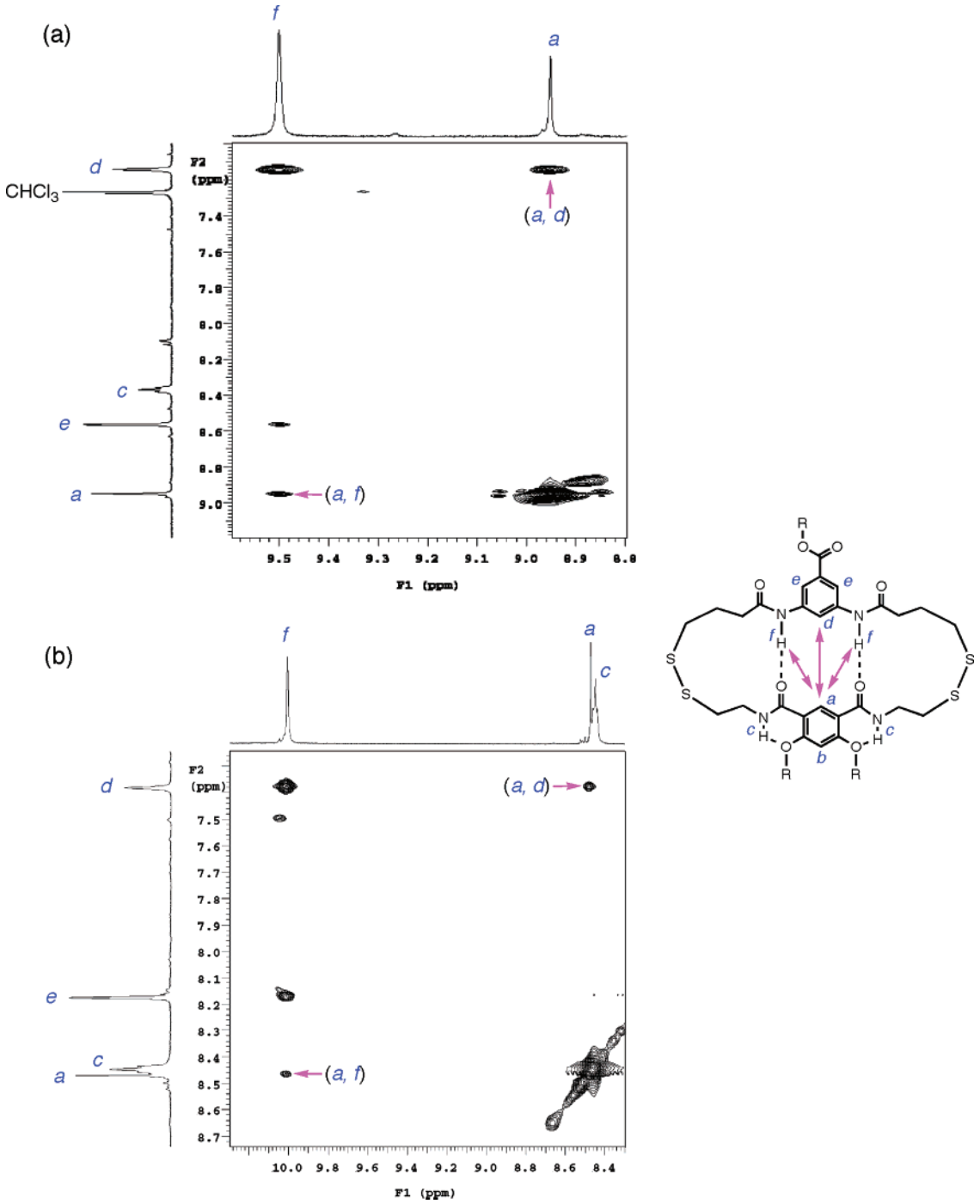
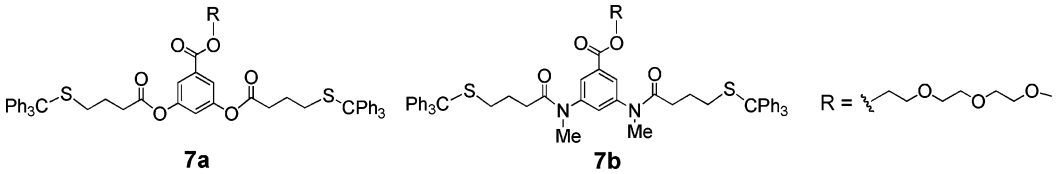


Figure 6. Partial 2D ^{1}H NMR spectra of disulfide-crosslinked **7**−**8** in (a) $CDCl_3$ (ROESY, 278 K, mixing time: 0.3 s) and (b) H_2O/THF-d_8 (80/20, v/v) (NOESY, 278 K, mixing time: 0.3 s). The same interstrand NOEs were detected in both $CDCl_3$ and in H_2O/THF-d_8.

Scheme 4



$CDCl_3$ and aqueous solution, along with the fact that no interstrand NOEs involving other amide and aromatic protons, that is, protons *b*, *c*, and *e*, argues for a H-bonded, edge-to-edge arrangement of the two crosslinked components of **7**−**8** in aqueous solution. The determining role of the two interstrand H-bonds in the selective formation of **7**−**8** in aqueous media was demonstrated by control compounds **7a** and **7b**, which are very similar to **7** but lack H-bond donors. Crosslinking of

compound **7a** or **7b** with **8** in aqueous media failed to lead to **7a**−**8** or **7b**−**8** as the dominant product.[27] The observed product distribution in either case can best be regarded as being statistical.

All factors considered, the best explanation is that intramolecular H-bonding still persists in aqueous media, which effectively stabilize the crosslinked products with matched H-bonding sequences (Scheme 4).

Conclusions

In summary, the reversible crosslinking of oligoamides carrying various numbers and arrangement of H-bond donors and acceptors was systematically examined in both nonpolar and aqueous media. Results from ESI, MALDI-TOF, reverse phase HPLC, and 2D NMR studies clearly demonstrated the high fidelity of this class of new associating units. The sequence specificity observed in aqueous media for complementary oligoamides is consistent with a thermodynamically controlled mechanism that involves the stabilization of the disulfide-crosslinked products by fully matched arrays of H-bond donors and acceptors. The interstrand H-bonds, which seem not to play any role in associating two complementary strands in aqueous media, are turned into intramolecular ones upon the formation of the disulfide crosslinks and serve to stabilize the crosslinked products. Results from ^{1}H NMR experiments are consistent with the presence of the intramolecular H-bonds in not only chloroform but also aqueous solution. In addition, the observed length-dependent selectivity offers another means that further enhances the diversity of this series of molecular associating units. This information-storing molecular system is characterized by (1) reversibility and thus error-correcting ability typical of noncovalent, self-assembling systems; (2) high stability resulted from the covalent, i.e., disulfide, crosslinks; (3) high fidelity as evidenced by the specific pair of complementary partners and the lack of any "cross talk" between noncomplementary oligoamide strands. The thermodynamic, instead of kinetic, nature of the disulfide-crosslinking process is consistent with results from previous studies, which indicated a significantly diminished energetic contribution by H-bonding to molecular association in aqueous media due to strong competition from water molecules.[30] The effectiveness of intramolecular H-bonds, which appear upon the formation of disulfide bonds, in stabilizing the corresponding products was clearly desmonstrated.[31] The fact that specific crosslinking can be biased by as few as two H-bonds indicates that, even in aqueous media, H-bonds, when properly placed, can effectively direct and shift the equilibria involving multiple products related by dynamic covalent interactions. This system represents one of the few examples[32] of dynamic covalent processes biased by H-bonding in the highly competitive aqueous media. The thermodynamic mechanism revealed by this study should be general to the design of systems combining H-bonding and other reversible covalent interactions. On the basis of the highly efficient crosslinking of both long and short oligoamide strands, in nonpolar and particularly in aqueous media, a variety of sequence- and length-specific molecular recognition units should become available, which points to the exciting possibility of programming the self-assembly of various structural components in aqueous as well as organic media. This system, which integrates highly specific multiple H-bond arrays with dynamic covalent bonds, also offers an easily tunable platform for studying the potential role of H-bonding interaction in water.

Experimental Procedures

All chemicals were purchased from Aldrich or Fisher (Acros) and were used as received unless otherwise indicated. All reactions were followed by thin-layer chromatography (precoated 0.25 mm silica gel plates from Aldrich), and silica gel column chromatography was carried out with silica gel 60 (mesh 230−400). ^{1}H NMR spectra were collected on at 400 MHz and ^{13}C NMR spectra were recorded at 75 MHz on Varian INOVA instruments. ^{1}H and ^{13}C NMR chemical shifts are reported as δ values in ppm relative to TMS or residual solvent. The ^{1}H chemical shifts (δ) were reported in parts per million downfield from TMS (tetramethylsilane). ^{13}C NMR spectra were recorded in ppm relative to the solvent signals: $CDCl_3$ (77.23 ppm) and $(CD_3)_2SO$ (39.52 ppm). Melting points were determined using a MEL-TEMP II apparatus. MALDI-TOF MS spectra were recorded on a Bruker Bifrex IV MS spectrometer using 2,5-Dihydroxybenzoic acid (DHB) or dithranol (DTH) as the matrices. Electrospray ionization mass spectra (ESI-MS) were on a Thermo Finnigan LCQ Advantage MS spectrometer.

General Procedures of the Disulfide-Crosslinking Reactions. A portion (0.5 mL, or 1 mL for reactions involving an individual strand) of an oligoamide was taken from its stock solutions (2.0 mM in CH_2Cl_2) and mixed in a 2-dram vial. Solvent was evaporated *in vacuo*, and the residue was redissolved into 2 mL of iodine solution (6 mM) in the corresponding solvent (CH_2Cl_2 or H_2O−THF). In all of the reaction mixtures, the concentration of each oligoamide component was maintained at 0.5 mM. The resulting mixture was stirred at room temperature.

MALDI-TOF Experiments. MALDI-TOF MS spectra of the reaction mixtures were recorded on a Bruker Bifrex IV MS spectrometer. Sampling and matrix: 2,5-dihydroxybenzoic acid (DHB) in THF was used as matrix. Two sampling methods, **A** and **B** (see below), were compared. No significant difference in the results obtained by these tow methods was observed. If there is no specific notice, all the MALDI-TOF MS spectra presented are obtained by sampling method **A** (direct sampling). **Sampling Method A:** Reaction solution was directly spotted on the layer of the matrix. **Sampling Method B:** Reaction mixture (0.5 mL) was taken and cooled in an ice bath, and 3 mM of aqueous sodium thiosulfate was added until the color of iodine disappeared. The mixture then was extracted with CH_2Cl_2. The organic layer was washed with saturated aqueous NaCl solution and dried over anhydrous Na_2SO_4, which was then spotted on the layer of the matrix on the target.

Reverse Phase HPLC Measurements. The HPLC experiments were carried out on a Shimadzu 10AV VP. All traces were recorded with a UV detector at 254 nm. The column was a Microsorb 100 C8 (4.6 × 150 mm, 5 μm) of VARIAN. An aliquot (15 μL) drawn from a reaction mixture was directly injected for analysis at ambient temperature. HPLC conditions: 1.0 mL/min; methanol−water gradient elution; methanol 20% in 2 min followed by methanol 20% to 40% in 5 min, methanol 40% to 70% in 45 min, finally, methanol 70−95% in 30 min. The identity of each peak was confirmed by collecting the corresponding fraction that was then subjected to analysis by mass spectrometry.

(30) (a) Fersht, A. R.; Shi, J.-P.; Knill-Jones, J.; Lowe, D. M.; Wilkinson, A. J.; Blow, D. M.; Brick, P.; Carter, P.; Waye, M. M. Y.; Winter, G. *Nature* **1985**, *314*, 235−238. (b) Klotz, I. M. *Protein Sci.* **1993**, *2*, 1992. (c) Williams, D. H.; Searle, M. S.; Mackay, J. P.; Gerhard, U.; Maplestone, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1172. (d) BenTal, N.; Sitkoff, D.; Topol, I. A.; Yang, A. S.; Burt, S. K.; Honig, B. *J. Phys. Chem. B* **1997**, *101*, 450. (e) Sheu, S. Y.; Yang, D. Y.; Selzle, H. L.; Schlag, E. W. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12683.

(31) A previous FT-IR study indicated that, in water, interstrand H-bonding played an insignificant role in stabilizing β-hairpin: Colley, C. S.; Griffiths-Jones, S. R.; George, M. W.; Searle, M. S. *Chem. Commun.* **2000**, 593.

(32) Krishnan-Ghosh, Y.; Balasubramanian, S. *Angew. Chem., Int. Ed.* **2003**, *42*, 2171.

NMR Spectroscopy. 2D [^{1}H,^{1}H]-NOESY/ROESY spectra were recorded on a VARIAN INOVA spectrometer operating at ^{1}H resonance frequency of 500 MHz at $T = 278$ K using ~5 mM solutions of **7•8** and **7**−**8** in $CDCl_3$ (measurement time 11.5 h) and **7**−**8** in 80% H_2O/20% THF-d_8 (18.4 h). Chemical shifts were referenced relative to TMS. The 2D spectra were acquired with 2048 complex points in the direct dimension and 256 complex points in the indirect dimension, with 16 transients per FID. The relaxation delay between transients was set to of 2 s, and the mixing time was set to 300 milliseconds.

Acknowledgment. This work was supported by the Petroleum Research Fund, administered by the ACS (37200-AC4). ONR and NSF are acknowledged for partial funding.

Supporting Information Available: Synthetic procedures, additional ESI, 2D NMR spectra, and HPLC traces. This material is available free of charge via the Internet at http://pubs.acs.org.

JA072567M