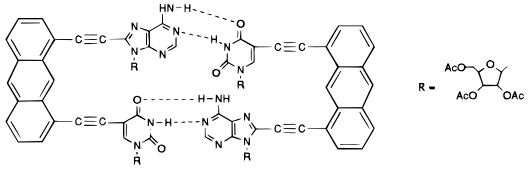
Self-Assembly of an "Artificial Dinucleotide Duplex"

Jonathan L. Sessler* and Ruizheng Wang

Department of Chemistry and Biochemistry
The University of Texas at Austin
Austin, Texas 78712

Received June 3, 1996

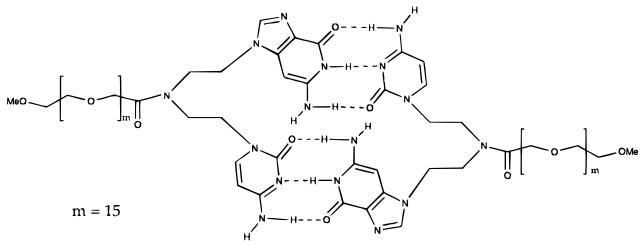
DNA provides a time-honored paradigm for how complementary functionality may be used to induce spontaneous assembly in complex supramolecules. In recent years, many self-complementary structures capable of undergoing assembly[1] have been designed and synthesized. Few of these, however, have made use of DNA-like complementary purine-pyrimidine base pairing to effect the critical recognition processes.[2] This seeming unpopularity could reflect the fact that adenosine (A)-thymidine (T) [or A-uridine (U)] interactions are too weak ($K_a \approx 10^2$ M^{-1} in $CDCl_3$[3]) and that cytosine (C) and guanine (G), which do have high affinities for complementary association ($K_a \approx 10^4$ M^{-1} in $CHCl_3$[2d,3b,4]), are notoriously difficult to work with. In spite of these limitations, we feel that nucleic acid base ("nucleobase") derived systems could provide important chemical and biochemical insight into fundamental base-pairing processes. Accordingly, we report here a "duplex-like" ensemble **I**, that assembles spontaneously from the corresponding "artificial dinucleotide" **1** under appropriate solution phase conditions.



1 **1**

I

Previous attempts[2d] to make duplex-like systems (e.g., **II**) based on the C/G containing dinucleotides failed to produce much in the way of *bona fide* "dimer". While the reasons for this remain obscure, it is likely that this failure reflects both the considerable conformational flexibility of the system and the poor solubility of the starting components.



II

Based on the above, the synthesis of rigid systems was considered desirable. Accordingly, compound **1** was selected as a target. It uses 1,8-diethynylanthracene as the spacer and complementary A/U base pairing entities as the recognition elements. The synthesis of **1** is summarized in Scheme 1.[5] It takes advantage of two successive palladium-catalyzed couplings involving **4**[6] and **5**[7] and **2** and **6**,[7,8] respectively. Two control compounds are also obtained: **2** as an intermediate and **3** as a byproduct.

Proton NMR studies provided initial evidence that the self-complementary molecule **1** associates to give the corresponding dimer (ensemble **I**). For instance, in chloroform-*d*, a solvent expected to favor A-U association[3b] and the formation of ensemble **I**, the uridine N−H protons of **1** are found to resonate at far lower field ($\delta = 14.68$ ppm) than the corresponding signals in **2** and **3** ($\delta = 9.22$ and 9.30 ppm, respectively). On the other hand, this significant downfield shift ($\Delta\delta = 5.4$ ppm) is not seen in DMSO-d_6. In this solvent, which is expected to disrupt any putative hydrogen bonding interactions present in **I**,[2d] the chemical shifts of all three compounds are virtually the same ($\delta = 11.60$, 11.74, and 11.98 ppm for **1**, **2**, and **3**, respectively). Under these DMSO-d_6 solvated conditions, the monomeric species **1**, rather than **I** presumably dominates.

The self-associated form of **1** (ensemble **I)** remains intact when up to 30% (v/v) DMSO-d_6 is added to the initial $CDCl_3$ solution. However, it is completely dissociated by the time the relative DMSO-d_6 concentration reaches 60%. In the regime where the DMSO-d_6 concentration in $CDCl_3$ is between 35% and 50% v/v, two different signals of N−H resonance are observed (Figure 1). Based on what is seen in pure $CDCl_3$ and 2:3 $CDCl_3$/DMSO-d_6, one signal at 14.6 ppm is ascribed to the hydrogen bonded ensemble **I**, while the other at 11.6 ppm is attributed to the unbound monomeric species **1**.

The fact that two separate signals are seen in the ^{1}H NMR spectrum is consistent with exchange between monomer **1** and dimer **I** being slow on the NMR time scale. Such slow exchange is usually observed in the case of very tightly bound complexes[9,10] and leads us to infer that self-association of **1** is favorable not only in $CDCl_3$ but also under these mixed solvent conditions.

In 45% (v/v) DMSO-d_6/$CDCl_3$, the two uridine N-H peaks (ascribed to monomer **1** and dimer **I**, respectively) are sufficiently similar in size that their areas may be determined accurately by integration. After accounting for stoichiometry, this gives the relative ratio of species **1** and **I**. With a knowledge of the total amount of starting ligand **1** (3.94×10^{-2} molar), a

(1) (a) Fujita, M.; Yazaki, J.; Ogura, K. *J. Am. Chem. Soc*. **1990**, *112*, 5645. (b) Drain, C. M.; Fischer, R.; Nolen, E. G.; Lehn, J.-M. *J. Chem. Soc., Chem. Commun*. **1993**, 243. (c) Branda, N.; Grotzfeld, R. M.; Valdes, C.; Rebek, Jr. J. *J. Am. Chem. Soc*. **1995**, *117*, 85. (d) Branda, N.; Wyler, R.; Rebek, Jr. J. *Science* **1994**, *263*, 1267. (e) Schwabacher, A. W.; Lee, J.; Lei, H. *J. Am. Chem. Soc*. **1992**, *114*, 7597. (f) Hunter, C. A.; Sarson, L. D. *Angew. Chem., Int. Ed. Engl*. **1994**, *33*, 2313. (g) Wyler, R.; de Mendoza, J.; Rebek, Jr. J. *Angew. Chem., Int. Ed. Engl*. **1993**, *32*, 1699. (h) Brienne, M.-J.; Gbard, J.; Ceclercq, M.; Lehn, J.-M. *Tetrahedron Lett*. **1994**, *35*, 8157. (i) Machias, J. P.; Simanek, E. E.; Whitesides, G. M. *J. Am. Chem. Soc*. **1994**, *116*, 4326. (j) Seto, C. T.; Whitesides, G. M. *J. Am. Chem. Soc*. **1991**, *113*, 712. (k) Bonar-Law, R. P.; Sanders, J. K. M. *Tetrahedron Lett*. **1993**, 1677. (l) Zimmerman, S. C.; Zeng, F.; Reichert, D. E. C.; Kolutuchin, S. V. *Science* **1996**, *271*, 1095. (m) Meissner, R. S.; Rebek, Jr. J.; de Mendoza, J. *Science* **1995**, *270*, 1485.

(2) (a) Wang, K.; Schall, O. F.; Gokel, G. W. *Supramol. Chem*. **1996**, *7*, 85. (b) Schall, O. F.; Gokel, G. W. *J. Am. Chem. Soc*. **1994**, *116*, 6089. (c) Kim, M.; Gokel, G. W.; *J. Chem. Soc., Chem. Commun*. **1987**, 1686. (d) Sessler, J. L.; Magda, D.; Furuta, H. *J. Org. Chem*. **1992**, *57*, 818.

(3) (a) Kyogoku, Y.; Lord, R. C.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A*. **1967**, *57*, 250. (b) Murray, T. J.; Zimmerman, S. C. *J. Am. Chem. Soc*. **1992**, *114*, 4010.

(4) (a) Sessler, J. L.; Wang, B.; Harrimann, A. *J. Am. Chem. Soc*. **1993**, *115*, 10418. (b) Sessler, J. L.; Wang, B.; Harrimann, A. *J. Am. Chem. Soc*. **1995**, *117*, 704. (c) Harrimann, A.; Kubo, Y.; Sessler, J. L. *J. Am. Chem. Soc*. **1992**, *114*, 388.

(5) Characterization data for all new compounds is included in the supporting information.

(6) (a) House, H. O.; Koepsell, D.; Jaeger, W. *J. Org. Chem*. **1973**, *38*, 1167. (b) Katz, H. E. *J. Org. Chem*. **1989**, *54*, 2179.

(7) (a) Matsuda, A. *Synthesis* **1986**, 385. (b) Sy, W.-W. *Synthetic Commun*. **1990**, *20*, 3391.

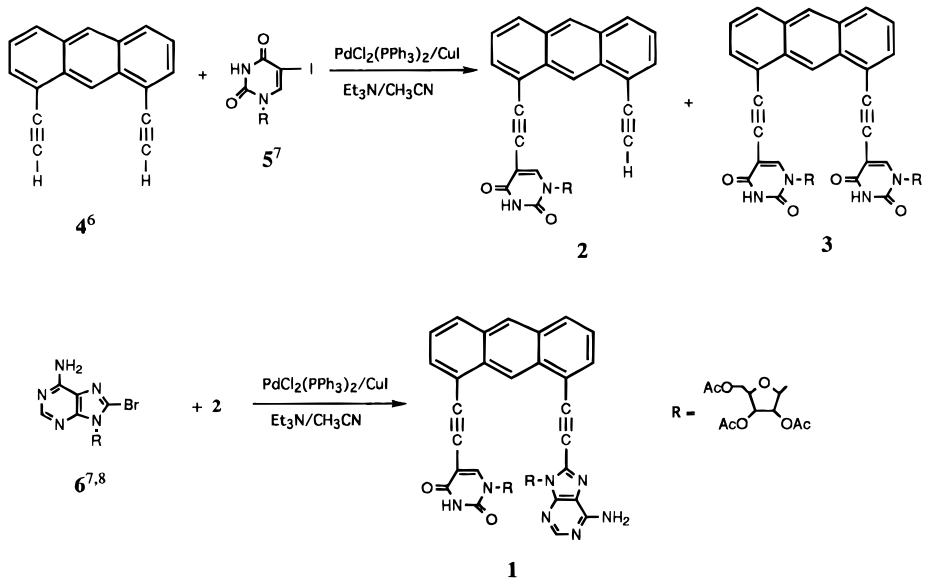
(8) Holmes, R. E.; Robins, R. K. *J. Am. Chem. Soc*. **1964**, *86*, 1242.

(9) (a) Connors, K. A. *Binding Constants;* John Wiley & Sons: New York, 1987. (b) Feeney, J.; Batchelor, J. G.; Albrand, J. P.; Roberts, G. C. K. *J. Magn. Reson*. **1979**, *33*, 519.

(10) The strong self-association of **1** (to give **I**) is also supported by NOESY experiments: Strong couplings are observed between the N−H signal of the uridine moiety and the two nonequivalent protons of the NH_2 substituent; they are also seen between the N−H signal and H^2 of adenosine.

S0002-7863(96)01845-8 CCC: $12.00

Scheme 1. Synthesis of **1**



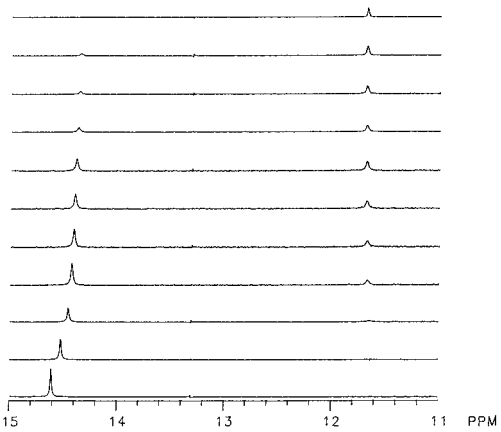


Figure 1. Stacked plots derived from an ^{1}H NMR spectroscopic titration of **I**. The solvent composition from bottom to top: 10% DMSO-d_6 in $CDCl_3$; 20% DMSO-d_6; 30% DMSO-d_6; 35% DMSO-d_6; 38% DMSO-d_6; 40% DMSO-d_6; 42% DMSO-d_6; 45% DMSO-d_6; 48% DMSO-d_6; 50% DMSO-d_6; 60% DMSO-d_6.

self-association constant of 35 ± 5 M^{-1}, corresponding to the formation of **I**, could thus be calculated directly.[11]

Further evidence that the self-assembled form of **I** is stable came from fast atom bombardment mass spectrometric (FABMS) experiments. In these studies, not only was a peak at $m/z =$ 985 amu (corresponding to the monomeric species **1**) observed, but also a peak at $m/z = 1970$ amu attributable to the dimer **I** was seen.[12]

The average molecular weight of **I** was also determined in solution using vapor pressure osmometry (VPO). Here, 1,2-dichloroethane was selected as the solvent and 4,5-bis(5′-ethynyl-2″,3″,5″tri-(O-acetyl)uridine)dibenzofuran[13] was used as the molecular weight standard. From these measurements, a M_w value of 2210 ± 45 (from three independent measurements) was obtained for **I**. This value agrees with the one of 2168 calculated for a molecular equivalent of dimer **I** containing two $ClCH_2CH_2Cl$ molecules within its boxlike core.

That **I** can bind 2 equiv of 1,2-dichloroethane either by inclusion or in the form of a clathrate complex (the available data do not permit a distinction) was further established *via* NMR spectroscopic studies and elemental analysis. When compound **1** (**I**) was recrystallized from 1,2-dichloroethane, it gave the following analysis: C, 56.51; H, 4.43; N, 8.94; Cl, 6.59. This corresponds to the bis-1,2-dichloroethane adduct of **I** (calcd for $(C_{49}H_{43}N_7O_{16})_2\cdot(C_2H_4Cl_2)_2$: C, 56.5; H, 4.37; N, 9.05; Cl, 6.46). The ^{1}H NMR spectrum (in $CDCl_3$) of this same recrystallized material also showed the exact integral ratio expected for a bis-1,2-dichloroethane complex.[14]

Currently we are working to define further the nature of the interactions between ensemble **I** and various small molecule guests.[15] We are also working to extend the chemistry of these "artificial oligonucleotides" by preparing systems with other rigid spacers (e.g., dibenzofuran) and other recognition units (e.g., C and G). Progress along these lines will be reported when warranted.

Acknowledgment. This work was supported by the Robert A. Welch foundation. We thank Prof. Grant Willson and Mr. Allen Gardiner for their assistance in carrying out VPO experiments.

Supporting Information Available: Details of synthesis of **1**, **2**, **3**; binding constant calculations and quantification of DMSO solvation; and NOESY spectrum of **I** in $CDCl_3$ (6 pages). See any current masthead page for ordering and Internet access instructions.

JA9618457

(11) Four equiv of DMSO are involved in the equilibrium (cf. supporting information). The calculated equilibrium constant is thus an effective one relevant under these particular solvation conditions.

(12) High resolution mass spectrometric analysis of this signal proved consistent with the proposed chemical formula (calcd for **I** $C_{98}H_{86}N_{14}O_{32}$: $M_w = 1970.5533$; found 1970.5520).

(13) Sessler, J. L.; Wang, R., unpublished results.

(14) Peak assignments for $ClCH_2CH_2Cl$ were made by adding extra 1,2-dichloroethane into the sample solution.

(15) Further support for the idea that ensemble **I** can interact strongly with small organic molecules came from a microanalysis of compound **1** recrystallized from a mixture of C_6H_6/CH_2Cl_2. The following specifically was found: C, 62.07; H, 4.65; N, 9.17. This corresponds to the bis-benzene adduct of **I** (calcd for $(C_{49}H_{43}N_7O_{16})_2\cdot(C_6H_6)_2$: C, 62.09; H, 4.64; N, 9.21).