Molecular Boxes Derived from Crown Ethers and Nucleotide Bases: Probes for Hoogsteen vs Watson–Crick H-Bonding and Other Base–Base Interactions in Self-Assembly Processes†

Otto F. Schall and George W. Gokel*

Contribution from the Bioorganic Chemistry Program, Department of Molecular Biology and Pharmacology, Washington University School of Medicine, 660 South Euclid Avenue, Campus Box 8020, St. Louis, Missouri 63110

Received December 20, 1993®

Abstract: Bibracchial lariat ethers based upon 4,13-diaza-18-crown-6 having $-(CH_2)_3-$ sidearms terminated in adenine or thymine have been prepared and characterized. The three structures are as follows: adenine–$(CH_2)_3$–(N18N)–$(CH_2)_3$–adenine (A–O–A), thymine–$(CH_2)_3$–(N18N)–$(CH_2)_3$–thymine (T–O–T), and adenine–$(CH_2)_3$–(N18N)–$(CH_2)_3$–thymine (A–O–T). Association of the nucleotide bases was expected to afford molecular boxes or other aggregates that would be stabilized by interactions between or among the nucleotide bases. These compounds have been studied in solution by ^{1}H-NMR spectroscopy and by vapor pressure osmometry to determine the extent of association as well as what interactions occur between the bases. The ^{1}H-NMR solution studies involved both temperature and concentration dependence and NOE studies. Several lines of evidence make clear that association does occur in $CDCl_3$ with an association constant for A–O–A with T–O–T of 855 M^{-1}. Both intra- and intermolecular H-bonding interactions are detected. Hoogsteen binding modes appear to play a very important role in these flexible systems. The A–O–A·T–O–T box may also comprise a ditopic receptor system in which the sides of the box are Ade::Thy pairs and the ends are crown ethers. We have studied such systems in the presence of decanediyldiammonium and dodecanediyldiammonium salts and report evidence for a ternary induced-fit receptor complex.

Introduction

Numerous proteins have been studied in part to discover how guests are bound by these large and complex hosts.[1] Many model systems have also been explored in an effort to elucidate binding interactions.[2] Despite this enormous effort, relatively few examples exist of simple host–guest molecular (non-metallic) complexes.[3] Even fewer examples are extant of low molecular weight systems that mimic biological induced-fit receptors.[4] It is now clear that binding modes other than B-DNA Watson–Crick interactions are common in nature. Subtle differences in binding modes may significantly alter conformation and have consequences in terms of biological function. Several years ago, we embarked on a program to develop an induced-fit receptor mimic that used the self-assembly properties of nucleotide bases for their formation. Such a system would also present the hitherto rare opportunity to examine nucleotide base interactions in the absence of the normal sugar phosphate structures that usually constrain them to Watson–Crick hydrogen-bonding interactions.

Since our preliminary report of that work,[5] several interesting examples have appeared of receptors utilizing nucleotide base pairs as organizational or recognition elements.[6] Some of those studies have utilized the three hydrogen bond interactions of cytosine (Cyt) and guanine (Gua). The obvious advantage in their use is that assembly involves 50% more hydrogen bonds than are available in the adenine–thymine (Ade::Thy) interaction. The disadvantage is that syntheses involving Cyt and Gua are notoriously difficult although elegant solutions to the synthetic problems have certainly been realized in individual cases.[6c] We recently reported a family of related nucleotide-base-containing systems which displayed highly organized H-bonding and π-stacking cooperativity.[7]

We present here the design, preparation, and detailed solution studies of a series of "molecular boxes" that utilize nucleotide base organizing elements. A detailed examination of their solution behavior allowed us to assess the H-bonding tendencies in flexible systems. Further, in the presence of appropriately sized diammonium salts, a three-component assembly is formed from the self-assembled ditopic receptor that can be accurately called a low molecular weight "induced-fit receptor model".

Results and Discussion

Design of Adenine–Thymine-Based, Self-Assembling Boxes. Our design of an induced-fit receptor system was based on two

† Dedicated to Professor Donald J. Cram on the occasion of his 75th birthday.

* Abstract published in *Advance ACS Abstracts,* June 1, 1994.

(1) (a) Evans, P. R.; Hudson, P. J. *Nature* **1979**, *279*, 500. (b) Fersht, A. R. *Enzyme Structure and Mechanism*, 2nd ed.; W. H. Freeman: New York, 1985; pp 263–291. (c) Evans, D. F.; Ninham, B. W. *J. Phys. Chem.* **1986**, *90*, 226. (d) Lau, F. T.-K.; Fersht, A. R. *Nature* **1987**, *326*, 811. (e) Fersht, A. R. *TIBS* **1987**, *12*, 301. (f) Ollis, D. L.; White, S. W. *Chem. Rev.* **1987**, *87*, 981. (g) Hemmingsen, S. M.; Ellis, R. J. *TIBS* **1989**, *14*, 339. (h) Quiocho, F. A. *Pure Appl. Chem.* **1989**, *61*, 1293. (i) Perutz, M. F. In *Mechanisms of Cooperativity and Allosteric Regulation in Proteins*; Cambridge University Press: Cambridge, England, 1989. (j) Laughon, A. *Biochemistry* **1991**, *30*, 11357. (k) Pavletich, N. P.; Pabo, C. O. *Science* **1991**, *242*, 809. (l) Harrison, S. C. *Nature* **1991**, *353*, 715.

(2) Dugas, H. In *Bioorganic Chemistry: A Chemical Approach to Enzyme Action*, 2nd ed.; Springer-Verlag: New York, 1989; Chapter 5.

(3) (a) Askew, B.; Ballester, P.; Buhr, C.; Jeong, K. S.; Jones, S.; Parris, K.; Williams, K.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1989**, *111*, 1082, 1090. (b) Tecilla, P.; Dixon, R. P.; Slobodkin, G.; Alavi, D. S.; Waldeck, D. H.; Hamilton, A. D. *J. Am. Chem. Soc.* **1990**, *112*, 9408. (c) Zimmerman, S. C.; Wu, W.; Zeng, Z. *J. Am. Chem. Soc.* **1991**, *113*, 196. (d) Cochram, J. E.; Parrott, T. J.; Whitlock, B. J.; Whitlock, H. W.; *J. Am. Chem. Soc.* **1992**, *114*, 2269. (e) Diederich, F.; Smithrud, D. B.; Sanford, E. M.; Wyman, T. B.; Ferguson, S. B.; Carcanague, D. R.; Chao, I.; Houk, K. N. *Acta Chim. Scand.* **1992**, *46*, 205.

(4) (a) Van Staveren, C. J.; Van Eerden, J.; Van Veggel, F. C. J. M.; Harkema, S.; Reinhoudt, D. N. *J. Am. Chem. Soc.* **1988**, *110*, 4994. (b) Rebek, J., Jr. *Acc. Chem. Res.* **1990**, *23*, 399. (c) Galan, A.; De Mendoza, J.; Toiron, C.; Bruix, M.; Deslongchamps, G.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 9424. (d) Nowick, J. S.; Feng, Q.; Tjivikua, T.; Ballester, P.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 8831. (e) Inouye, M.; Konishi, T.; Isagawa, T. *J. Am. Chem. Soc.* **1993**, *115*, 8091.

(5) Kim, M. S.; Gokel, G. W. *J. Chem. Soc., Chem. Commun.* **1987**, 1686–1688.

(6) (a) Nagai, K.; Hayakawa, K.; Ukai, S.; Kanematsu, K. *J. Org. Chem.* **1986**, *51*, 3931. (b) Aoyama, Y.; Onishi, H.; Tanaka, Y. *Tetrahedron Lett.* **1990**, *1177*. (c) Furuta, H.; Magda, D.; Sessler, J. L. *J. Am. Chem. Soc.* **1991**, *113*, 978, 4706. (d) Hisatome, M.; Maruyama, N.; Furutera, T.; Ishikawa, T.; Yamakawa, K. *Chem. Lett.* **1990**, 2251. (e) Nowick, J. S.; Chen, J. S.; Noronha, G. *J. Am. Chem. Soc.* **1993**, *115*, 7636.

(7) Schall, O. F.; Gokel, G. W. *J. Chem. Soc., Chem. Commun.* **1992**, 748.

0002-7863/94/1516-6089$04.50/0

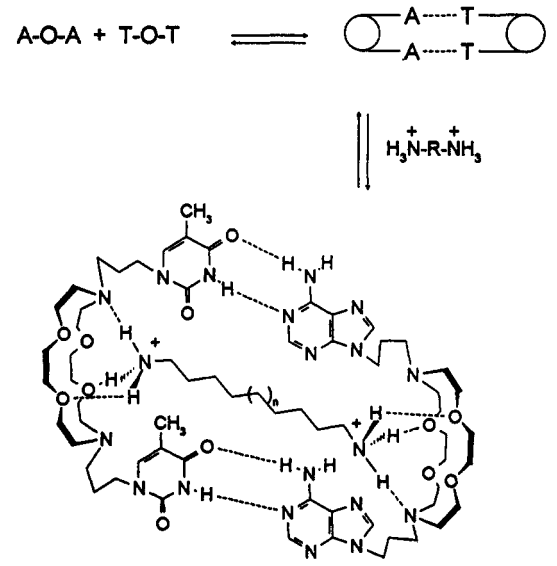
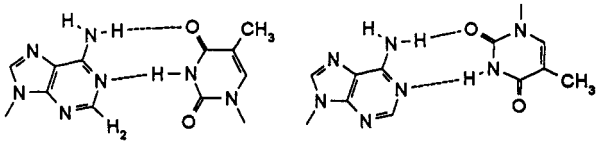
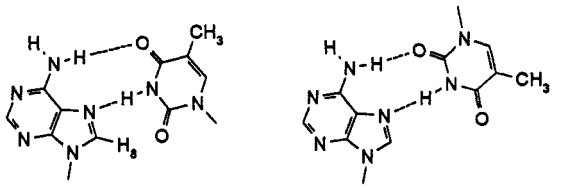


Figure 1. Ternary host–guest complex based on the adenine–thymine "DNA" box.



Normal and Reversed Watson-Crick



Normal and Reversed Hoogsteen

Figure 2. Watson–Crick and Hoogsteen H-bond arrangements.

premises. First, the crown ether unit is known to be an effective complexation partner for primary ammonium salts.[8] Second, a lariat ether[9] having sidearms terminated in a nucleotide such as adenine should pair with its complement. Thus, a bis(ammonium) salt of the type $H_3N^+(CH_2)_nNH_3^+$ is expected to form a tripod of hydrogen bonds between each $-NH_3^+$ group and an 18-crown-6 derivative. If the spacing of the tethers connecting the macroring to the nucleotide base is the appropriate length, a three component assembly could form. This is illustrated schematically in Figure 1. We represent the adenine–crown–adenine system as Ade–$(CH_2)_3$–diaza-18-crown-6–$(CH_2)_3$–Ade (A–O–A); thymine substitution gives Thy–$(CH_2)_3$–diaza-18-crown-6–$(CH_2)_3$–Thy (represented as T–O–T).

We recognized that both Watson–Crick and Hoogsteen hydrogen bond arrangements would be possible in a flexible system that lacks the constraints of a sugar phosphate chain. Four possibilities involving two hydrogen bonds each are well-known[10] and are illustrated in Figure 2. In any one of these, assembly by virtue of hydrogen-bond formation would be possible although the final geometry of the box would obviously differ in each case.

(8) (a) Pedersen, C. J. *J. Am. Chem. Soc.* **1967**, *89*, 7017. (b) Izatt, R. M.; Terry, R. E.; Haymore, B. L.; Hansen, L. D.; Dalley, N. K.; Avondet, A. G.; Christensen, J. J. *J. Am. Chem. Soc.* **1976**, *98*, 7620. (c) Sutherland, I. O. *Chem. Soc. Rev.* **1986**, *15*, 63. (d) Sutherland, I. O. In *Guest-Host Chemistry and Conformational Analysis*; Marchand, A., Ed.; VCH Publishers, Inc.: Deerfield Beach, FL, 1986, Chapter 1.

(9) Gokel, G. W. *Chem. Soc. Rev.* **1992**, *21*, 39.

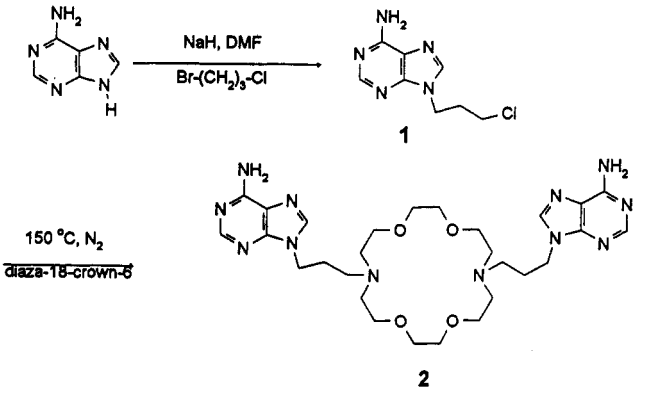


Figure 3. Preparation of A–O–A.

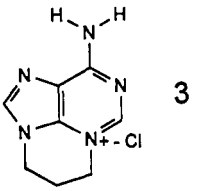


Figure 4. Structure of tricyclic compound **3**.

Successful methodology for the alkylation of nucleic acid bases has been known since the early 1960s.[11] Of the five nucleotide bases, adenine, thymine, uracil, cytosine and guanine, derivatization of the latter has presented the greatest synthetic difficulty because of its multifunctionalities. It was decided to tether the nucleotide bases Ade and Thy to the nitrogen atoms of 4,13-diaza-18-crown-6. The 18-membered ring system was selected because of its complementarity to the primary ammonium cation.[12] Attachment at nitrogen imparts flexibility to the system, and we had previously developed synthetic methods for related compounds.[13]

The Adenine–Crown–Adenine (A–O–A) Monomer Unit. The monomer unit A–O–A (**2**) has been prepared and characterized.[5,14] We obtained 9-(3-chloropropyl)adenine (**1**), after deprotonating adenine (NaH, DMF) and treating its sodium salt with excess 1-bromo-3-chloropropane, as white crystals (35 ± 5%).

N,N'-Bis[3-(9-adeninyl)propyl]-4,13-diaza-18-crown-6 (**2**) was obtained after reaction of the alkylated adenine and 4,13-diaza-18-crown-6 (1:1 ratio), in a melt (*ca.* 150 °C), as a white powder (30 ± 5%). The formation of the tricyclic compound (**3**) and the use of the diazacrown as nucleophile and base may account for the low yield of this reaction. Intramolecular displacement appears to occur preferentially. Nucleophilic attack of adenine's N3 at the 3-position of the propyl chain displaces a chloride ion and results in the formation of a stable tricyclic structure.[15] The displacement also occurs when **3** is heated until it liquifies.

The Thymine–Crown–Thymine (T–O–T) System. Several repetitions of our previously reported successful synthesis of T–O–T proved this approach to be somewhat capricious. The

(10) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer Verlag: Berlin, 1991; p 260.

(11) (a) Leonard, N. J.; Scott, T. G.; Huang, P. C. *J. Am. Chem. Soc.* **1967**, *89*, 7137. (b) Takemoto, K.; Kondo, K.; Kono, M.; Ueda, N. *Makromol. Chem.* **1968**, *120*, 13. (c) Carraway, K. L.; Huang, P. C.; Scott, T. G. In *Synthetic Procedures in Nucleic Acid Chemistry*; J. Wiley & Sons: New York, 1969; Vol. 1, pp 3–5. (d) Browne, D. T. In *Synthetic Procedures in Nucleic Acid Chemistry*; J. Wiley & Sons: New York, 1969; Vol. 1, pp 96–97.

(12) Schultz, R. A.; Schlegel, E.; Dishong, D. M.; Gokel, G. W. *J. Chem. Soc., Chem. Commun.* **1982**, 242.

(13) (a) Gatto, V. J.; Gokel, G. W. *J. Am. Chem. Soc.* **1984**, *106*, 8240. (b) Schultz, R. A.; White, B. D.; Dishong, D. M.; Arnold, K. A.; Gokel, G. W. *J. Am. Chem. Soc.* **1985**, *107*, 6659. (c) Gatto, V. J.; Arnold, K. A.; Viscariello, A. M.; Miller, S. R.; Morgan, C. R.; Gokel, G. W. *J. Org. Chem.* **1986**, *51*, 5373. (d) Gatto, V. J.; Arnold, K. A.; Viscariello, A. M.; Miller, S. R.; Gokel, G. W. *Tetrahedron Lett.* **1986**, 327.

(14) We previously reported (ref 5) this monomer as a monohydrate having mp 154–157 °C (see the Experimental Section).

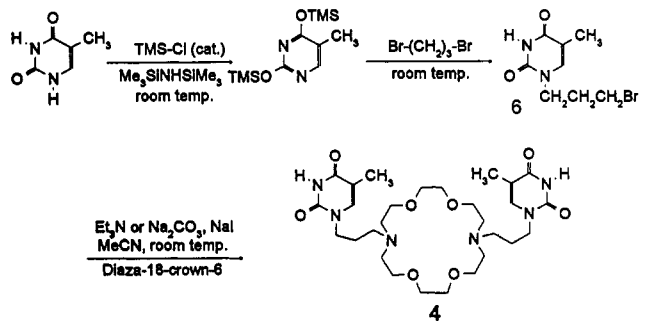


Figure 5. Preparation of T–O–T.

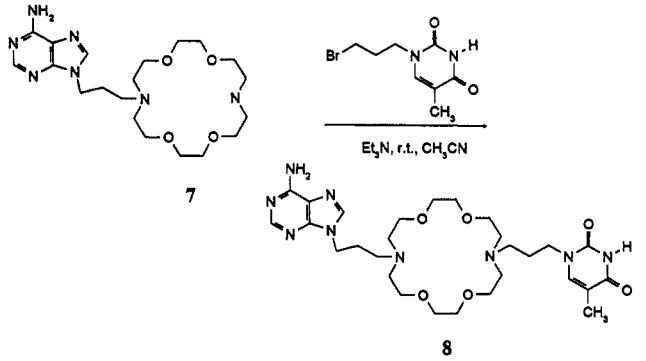


Figure 6. Preparation of A–O–T.

approach using alkali-metal carbonates and NaI (Figure 5) yielded *N,N'*-bis[3-(1-thyminyl)propyl]-4,13-diaza-18-crown-6 (**4**, T–O–T), as a white solid[16] (13 ± 2%), as well as *N*-[3-(1-thyminyl)-propyl]-4,13-diaza-18-crown-6 (**5**, T–O, colorless crystals, yield 35 ± 5%). Thymine was bis(silylated) using hexamethyldisilazane and catalytic Me_3SiCl and then treated with excess 1,3-dibromopropane to give 1-(3-bromopropyl)thymine (**6**) as white crystals (63 ± 2%). Alternatively, treatment of **6** with diaza-18-crown-6 and Et_3N in MeCN at ambient temperature for 7 d gave T–O–T monomer **4** (65 ± 3%), which crystallized from the reaction mixture. One recrystallization from MeCN afforded analytically pure T–O–T.

The previous reproducibility problems with this reaction were due to the inherent acidity of the N–H bond in **6** (pK_a = 9.7 in thymidine)[17] and concomitant competing intramolecular cyclization.[18]

The T–O Monomer. It seemed possible that the slow rate of T–O–T formation might be due to cation complexation. Varying the base ($Na_2CO_3 \rightarrow Li_2CO_3$, K_2CO_3) did not significantly alter the T–O–T yield. Examination of Corey–Pauling–Koltun (CPK) molecular models suggested to us that the thymine residue in monosubstituted **5** might coordinate, through its carbonyl O2,[19] to a ring-bound cation. The presumed cation complexation by **5** (T–O) would compromise alkylation of the crown's unsubstituted nitrogen atom. The preparation of T–O–T was thus successfully accomplished using Et_3N as the base (see the Experimental Section for details).

The Adenine–Crown–Thymine (A–O–T) Monomer. A monomer having both adenine- and thymine-terminated sidearms was very appealing to us because a single precursor could, in principle, form a dimeric "molecular box". Thus, *N*-[3-(9-adeninyl)propyl]-4,13-diaza-18-crown-6 (**7**, A–O, obtained as a byproduct of A–O–A synthesis) was treated with 1-(3-bromopropyl)thymine and Et_3N in MeCN at ambient temperature for 7 d (Figure 6). The product ("mixed monomer", A–O–T, **8**) was obtained (53%) as a white foam. Although pure by all normal criteria, attempts to crystallize the foam failed.

The Cytosine–Crown–Cytosine Monomer (C–O–C). The C–O–C monomer corresponding to T–O–T was prepared[20] as the complement for G–O–G. Unfortunately, the latter remains unrealized.

Self-Association of the A–O–A, T–O–T, and A–O–T Monomers. Analysis by ^{1}H-NMR of the A–O–A and T–O–T monomers in $CDCl_3$ (1:1 ratio, 5 mM each at 22 °C) gave results qualitatively

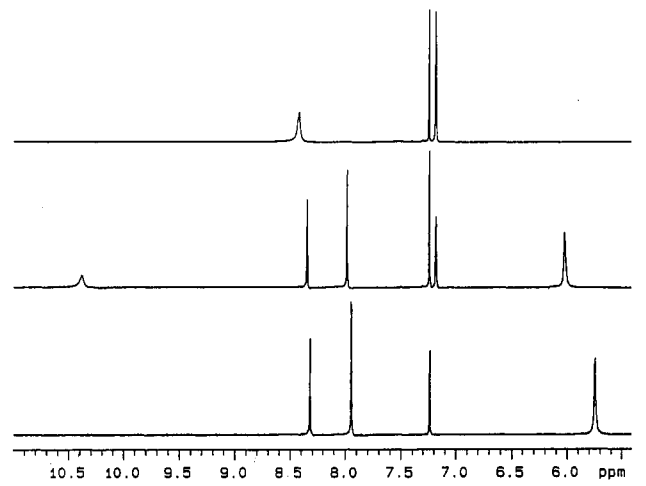


Figure 7. Partial ^{1}H-NMR spectra of A–O–A (lower trace) and T–O–T (upper trace) and their 1:1 mixture: [A–O–A] = [T–O–T] = [A–O–T + A–O–T] = 10 mM.

similar to those obtained originally by Katz and Penman[21] and others.[22] Downfield shifts were observed for the thymine imido-H (1.964 ppm), the adenine amino-H's (0.270 ppm), and the C2-H (0.026 ppm) and C8-H (0.037 ppm) of adenine, clearly indicating the formation of hydrogen-bonded dimers (Figure 7). Others[22] have inferred that a significant C2-H downfield shift implies the formation of Watson–Crick hydrogen bonds and that a C8-H downfield shift implies interactions of the Hoogsteen type. Our system thus appears to give a mixture of H-bond interactions.

Upon A–O–A·T–O–T complex formation, the C8-H (Hoogsteen) or C2-H (Watson–Crick) of adenine would be deshielded (weakly H-bonded) by the thymine C2-carbonyl group (shown in Figure 8 by →). Indeed, in cyclic dimers of guanosine and cytidine, under conditions where triply hydrogen-bonded Watson–Crick dimers predominate, the guanine C8-H shifted less than 0.005 ppm upon mixing.[22a] It was shown in the same work that equilibrium constants for either Watson–Crick or Hoogsteen H-bond formation were nearly equal for adenine and uracil.

It is interesting to note that at concentrations <5 mM the monomers showed unexpected concentration-dependent behavior.

(15) Carraway, K. L.; Huang, P. C.; Scott, T. G. In *Synthetic Procedures in Nucleic Acid Chemistry*; J. Wiley & Sons: New York, 1969; Vol. 1, pp 3–5.

(16) The previously reported approach gave a colorless oil. The procedure described here gave a solid. A corrigendum has appeared: Kim, M. S.; Gokel, G. W. *J. Chem. Soc., Chem. Commun.* **1990**, 948.

(17) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984; p 111.

(18) (a) Browne, D. T.; Eisinger, J.; Leonard, N. J. *J. Am. Chem. Soc.* **1968**, *90*, 7302. (b) Skaric, D.; Skaric, V.; Cizmek, A. *J. Chem. Soc., Perkin Trans. 1* **1984**, 2221.

(19) See ref 17, pp 201–219.

(20) The cytosine amino group was acetylated, followed by protection–activation as its bis(trimethylsilyl) derivative. Further reaction with 1,3-dibromopropane afforded the halogenated sidearm N^4-acetyl-N^1-(3-bromopropyl)cytosine (39%, mp 149–152 °C) as yellow needles. Treatment of 4 equiv of the latter with 1 equiv of the diazacrown in MeCN (4.2 equiv of Et_3N, room temperature) for 14 d, followed by deprotection (NH_3/MeOH, room temperature, 16 h) yielded fully-characterized C–O–C as a white amorphous powder (25% overall, mp 237–240 °C, dec). ^{1}H-NMR (DMSO-d_6): 1.649 (m, 4H), 2.367 (t, 4H), 2.589 (t, 8H), 3.455 (t, 8H), 3.490 (s, 8H), 3.616 (t, 4H), 5.594 (d, 2H), 6.913 (broad d, 4H, exchange D_2O), 7.539 (d, 2H). Anal. Calcd for $C_{26}H_{44}N_8O_6$: C, 55.30; H, 7.85; N, 19.84%. Found: C, 55.02; H, 7.82; N, 19.75%.

(21) Katz, L.; Penman, P. *J. Mol. Biol.* **1966**, *15*, 220. Katz, L. *J. Mol. Biol.* **1969**, *44*, 279.

(22) (a) Newmark, R. A.; Cantor, C. R. *J. Am. Chem. Soc.* **1968**, *90*, 5010. (b) Engel, J. D.; Von Hippel, P. H. *Biochemistry* **1974**, *13*, 4143.

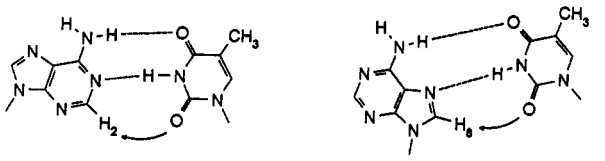


Figure 8. Shielding influence of thymine's carbonyl O2 on the carbon-bound resonances of adenine in both base-pairing modes, as indicated by the arrows.

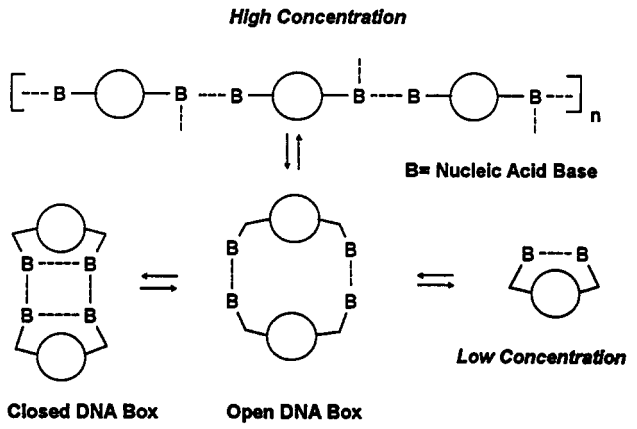


Figure 9. Proposed equilibria among the nucleotide-base monomers.

In nonpolar, poor hydrogen-bonding media (*e.g.* $CDCl_3$), the NMR signals were broadened. We believe this broadening to result from an equilibrium between differently hydrogen-bonded conformations. An examination of CPK molecular models suggested that several equally favorable aggregates might form as shown schematically in Figure 9. These include a linear oligomer that could also be branched, the "DNA box" in which the base pairs are not internally H-bonded, and the "DNA box" in which secondary H-bonding occurs intrapyxially. At low monomer concentration, the individual species could exhibit internal sidearm interactions. The observed signal broadening likely results from a mixture of these species in solution.

As the concentration of monomer in solution increases, so does the likelihood of association (aggregation, oligomerization). The binding forces are generally weak and certainly solvent dependent. Their precise nature defines the extent of aggregation. The solution probably contains few linked monomers, involved in rapid monomer exchanges. If the exchange is rapid, an "average" conformation could be observed by NMR as sharp signals. Dilution would favor monomers or dimers that are not further associated. The coexistence of such species would account for the observed NMR features. Examples of H-bond interactions thought to be possible are illustrated in Figure 10 where → or ← indicates the presence of a donor or acceptor. The possibility that both inter- and intramolecular association may occur is also illustrated (Figure 10).

The combination of inter- and intramolecular hydrogen bonding is appealing in enthalpic terms but may have an unfavorable entropic component. An attempt was made as follows to assess the conformational equilibria. The ^{1}H-NMR spectra ($CDCl_3$) were recorded of solutions containing (i) 4,13-diaza-18-crown-6 (1 mM) and 1-propylthymine[23] (2 mM), (ii) crown and 9-propyladenine[23] (2 mM), and (iii) the crown plus a mixture of the two sidearm analogs (1 mM each). The nucleotide base derivatives exhibited well-resolved signals. Some broadening was observed in signals assigned to the crown's hydrogens, but it was considerably less than that noted for the monomers (*e.g.* A–O–A in Figure 11). In polar media (D_2O, CD_3CN) no signal broadening was observed for the above mixtures.

(23) These compounds were prepared in a way similar to the preparation of 1-(3-bromopropyl)thymine and 9-(3-chloropropyl)adenine but using bromopropane as the starting material.

This model predicts that, at some low concentration, signals should split. As the concentration was lowered, we noted both broadening and new signals when the concentration of A–O–A and T–O–T approached 1 mM and the temperature was lowered (+22 → −60 °C). We are unaware of any precedent for this behavior in any related nucleotide-base-containing systems. Of course, our model systems are unusual in that adenine and thymine are spatially constrained and yet are attached to flexible frameworks. The versatile H-bonding patterns observed in these systems are no doubt due to this combination of properties.

Concentration Dependence Studies. The concentration dependence of the ^{1}H-NMR signals for A–O–A, T–O–T, A–O–T, and the A–O–A + T–O–T mixture was assessed in $CDCl_3$. Monomer concentrations were kept equivalent (*i.e.* [A–O–A] = [T–O–T] = [A–O–T] = [A–O–A + T–O–T] = X). Plots of amino and imido proton resonance positions as a function of total monomer concentration are shown in Figure 12a,b (22 °C).

An intramolecular hydrogen bond should not be affected by concentration (*i.e.*, a plot of δ *vs* concentration should be a straight line with a slope = 0). The shallower the observed slope, the greater the contribution of intramolecular H-bond components is presumed to be. The greatest curvature (amino protons) is apparent (Figure 12a) in the line corresponding to the A–O–A + T–O–T mixture. Similar and even more dramatic results are observed for the imido protons (Figure 12b).

As expected, when mixed with the complementary bases (1:1), both amino and imido resonances are shifted considerably downfield. The increased curvature suggests enhanced intermolecular association, no doubt because the Ade::Thy interactions are more favored. The downfield shifts were largest for A–O–T. These observations are consistent with increased Ade::Thy base pairing; both Ade and Thy are incorporated within the same molecule. In our preliminary designs, A–O–T appeared to be the most convenient monomer for "DNA-box" formation. We did not foresee that intramolecular hydrogen bonding would be enhanced at the expense of intermolecular association.

Intramolecular association is more readily observed in the imido resonance variations. Graphs of chemical shift dependence for the T–O–T and A–O–T monomers showed less curvature than for the mixed A–O–A and T–O–T monomers. Interestingly, we noted that the imido resonances of the T–O–T monomer were considerably more upfield than those exhibited in A–O–T and the monomer mixture, whereas the situation was different in the case of the amino resonances (*i.e.* the chemical shift trend of A–O–A followed very closely those of the other two). Selective irradiation of the most upfield (least H-bonded) N–H resonance of A–O–A (−45 °C, 25 mM in $CDCl_3$) caused a much larger decrease of the C2-H intensity. This result indicated a preference for Hoogsteen, rather than Watson–Crick, interactions.

Temperature Dependence Studies. We suggested above that conformational equilibria among the monomers may account for proton resonance broadening in dilute solutions. We anticipated that the presence of differently hydrogen bonded structures might manifest itself at low temperatures as well. ^{1}H-NMR spectra of A–O–A, T–O–T, A–O–T, and a 1:1 mixture of A–O–A + T–O–T (1 mM total concentration, $CDCl_3$) were recorded at temperatures from +22 to −60 °C. In both the A–O–A and T–O–T cases, signal splitting occurred as the temperature decreased. Signal broadening and splitting were most apparent for the A–O–A monomer. At least two imido resonances could be observed at −50 °C for T–O–T (signals exchanged with CD_3OD). We believe these are due to intra- and intermolecularly hydrogen-bonded species. Some broadening, although considerably less than for the pure monomers, was observed in the A–O–A + T–O–T mixture, but no signal splitting was apparent as observed for the other two monomers. This is probably due to preferential Ade::Thy pairing reducing the number of possible modes of association. Intramolecular association may be more favored in

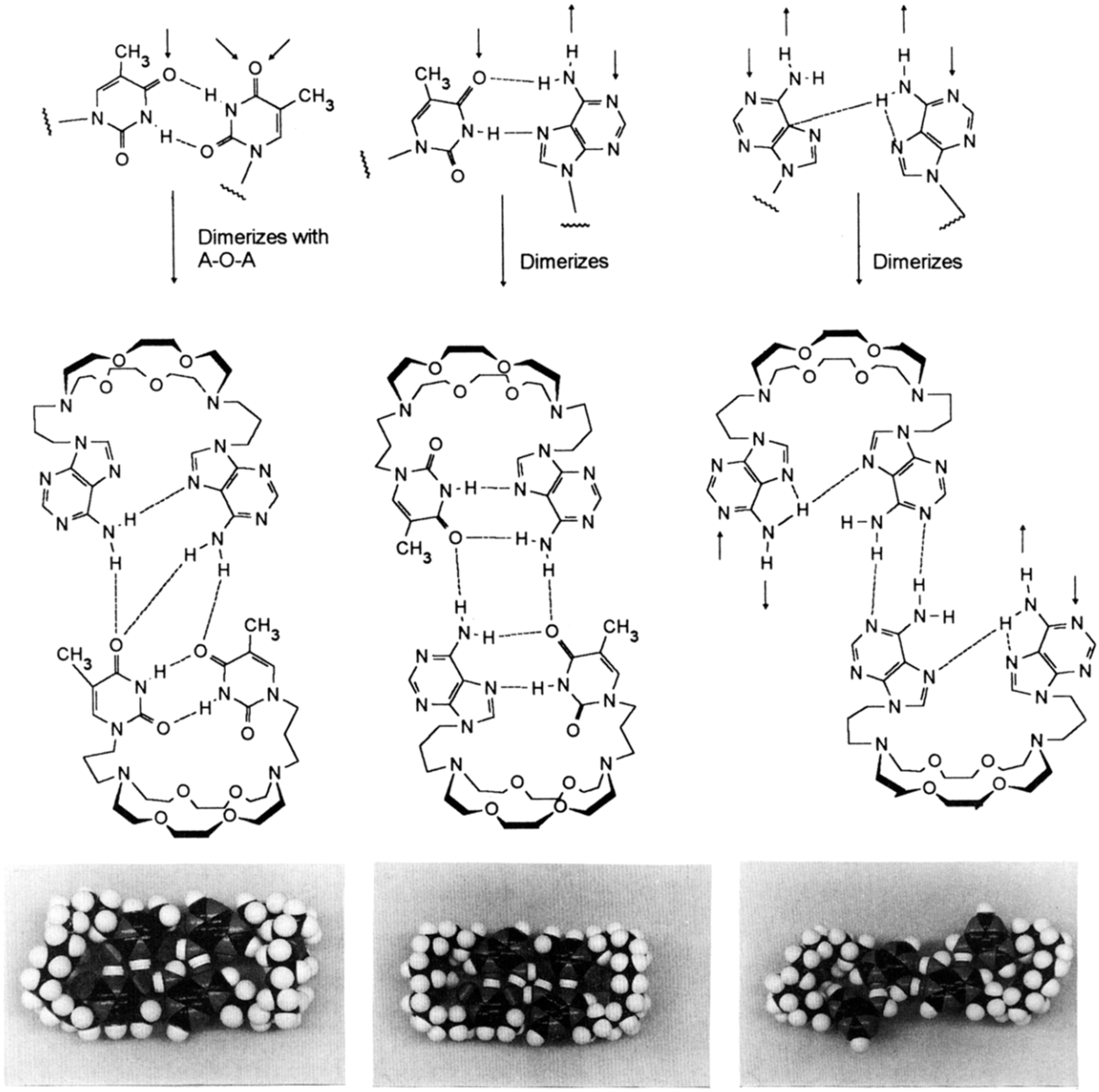


Figure 10. H-bond donors and acceptors present in the intramolecularly associated monomers and dual binding modes for the different monomers. Arrows indicate hydrogen bond donors and acceptors.

A–O–T since both Ade and Thy are present and proximate in the same molecule; no unusual signal splitting was observed. Figure 13a–c shows the proton NMR spectra of the monomers.

At low temperatures, the adenine $-NH_2$ group rotates slowly on the NMR time scale.[22,25] Signals for the individual N–H resonances were observed at –10, –15, and –20 °C for A–O–T, A–O–A + T–O–T, and A–O–A, respectively, at total monomer concentrations of 25 mM (Figure 14). The signal splitting clearly reflects differential H-bonding. In some cases the signal separation is nearly 0.8 ppm. The data also suggest that hydrogen bonding in A–O–T diminishes the rate of amino group rotation more than in the other two cases (individual N–H resonances discernible at a higher temperature). Selective irradiation of thymine's N–H resonance in A–O–T (25 mM, –50 °C) caused a larger decrease of adenine's C8-H resonance, suggesting a preferential Hoogsteen hydrogen-bonding mode.

The N–H resonances of A–O–T are downfield of those assigned either to the A–O–A + T–O–T mixture or to A–O–A itself, at least in the –10 to +25 °C range (25 mM). These data also suggest (as inferred from concentration studies, *vide infra*) that the A–O–T monomer engages in hydrogen bonding more readily than the A–O–A + T–O–T mixture. The N–H signal separation is largest for A–O–T, suggesting that intramolecular H-bonding favors the Hoogsteen over Watson–Crick binding mode (see below).

The behavior of the N–H signals of A–O–A was unexpected. As the temperature decreased, the individual resonances were observed and shifted more downfield than we would have predicted (Figure 14). At –60 °C, the most downfield resonance of A–O–A was 0.7 ppm above any resonance of either A–O–T or the A–O–A + T–O–T mixture. Even the least downfield signal of A–O–A was slightly downfield of the most shifted signals of either A–O–T or A–O–A + T–O–T. These unexpected trends are reminiscent of those previously observed in the concentration dependence studies. We do not fully understand the significance of the small upfield chemical shifts seen in the lower traces of all three cases, just after splitting of the N–H resonances occurs. We speculate that this may be a result of preferential binding modes whose effects, expressed in each individual N–H resonance position, are intensified as the amino groups lose rotational freedom. In

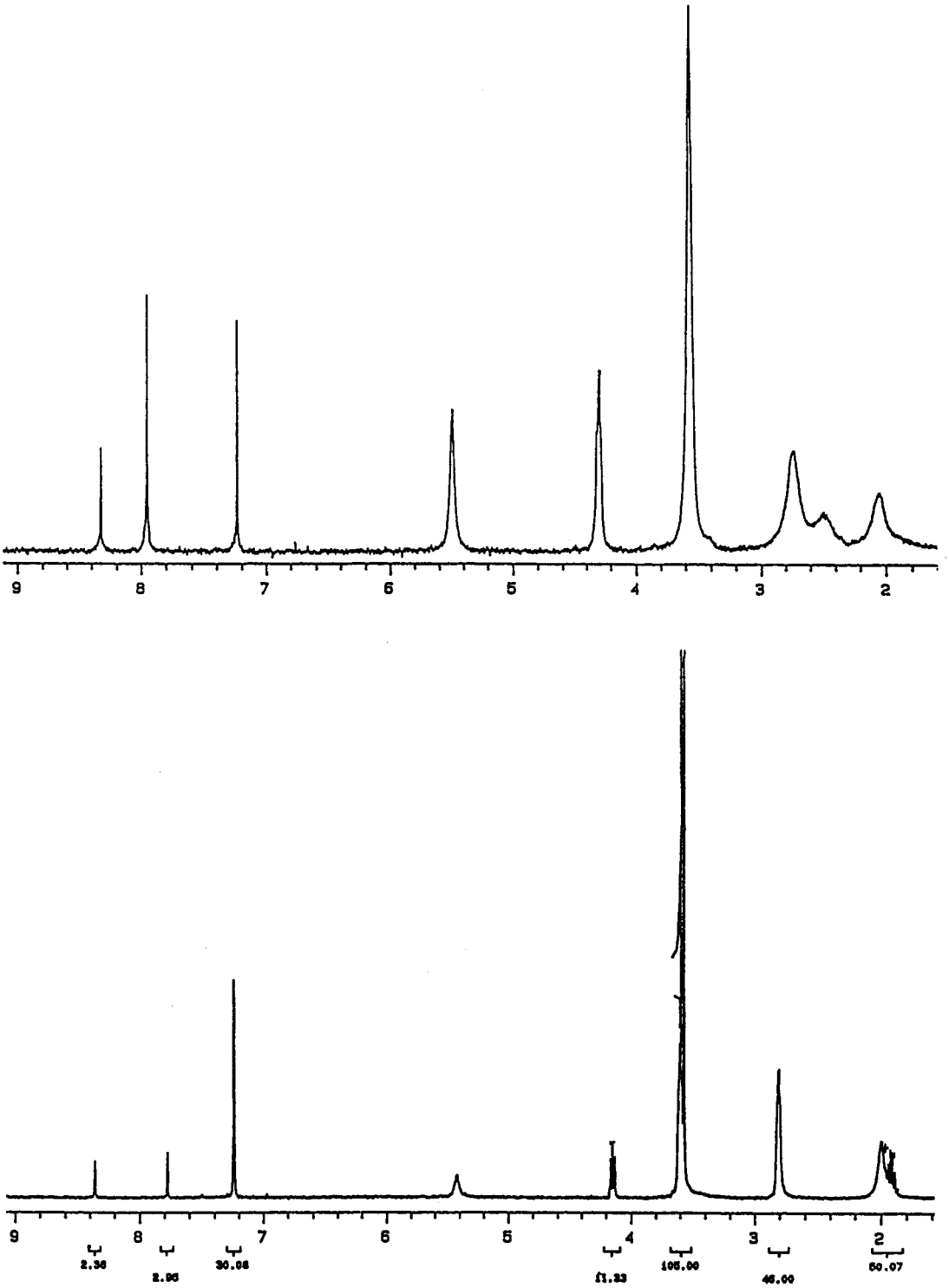


Figure 11. ^{1}H-NMR spectra of A–O–A (1 mM, upper trace) and a mixture of 4,13-diaza-18-crown-6 (1 mM) and 9-propyladenine (2 mM), both at 22 °C.

contrast and as expected, the imido resonances of the T–O–T monomer behaved differently and were upfield of those in A–O–T and A–O–A + T–O–T at all temperatures studied (Figure 15).

Concentration dependence experiments were conducted at −30 °C ($CDCl_3$) (Figure 16) because the signal separation was found to be largest at this temperature. A situation similar to those seen before occurred. At low concentrations, the A–O–A + T–O–T resonances are shifted downfield. However, as the concentration increased, those of the A–O–A monomer moved downfield more rapidly. Above about 15 mM, an adenine N–H resonance was the most downfield signal. From the trend, it appears that even the second N–H resonance would surpass the shift of the most downfield signal of A–O–A + T–O–T or A–O–T at concentrations >50 mM. At −30 °C, the adenine resonances in the former two tended to reach saturation faster than those of A–O–A alone. We believe that A–O–A monomers associate while hydrogen bonding intramolecularly (see Figure 10). Such a dimerization would bring four adenine rings approximately into the plane. The organization of these rings and their large aromatic surfaces and corresponding deshielding contributions likely account for the observed downfield shifts.

The top (most downfield) curve for A–O–T (see Figure 16) represents the N–H resonance preferentially involved in hydrogen bonding. The shifts of this resonance do not alter significantly

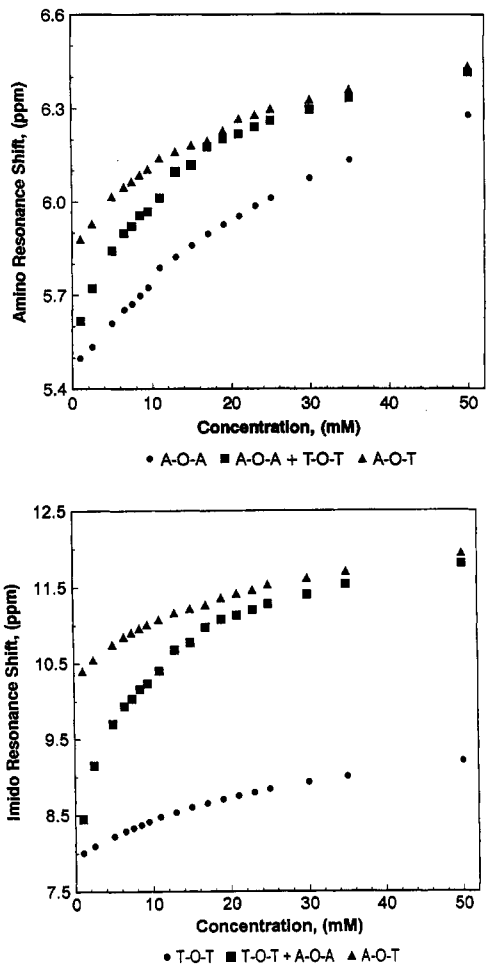


Figure 12. (top) Concentration dependence of the amino groups of the individual monomers and their mixture (22 °C). (bottom) Concentration dependence of the imido groups of the individual monomers and their mixture (22 °C).

over the entire concentration range. In contrast, the bottom (most upfield) line displays greater curvature ($\Delta\delta$ between 2.5 and 50 mM is 3-fold less than that of the lower curve). We interpreted this to mean that the "upper" N–H resonances reflect considerable intramolecular (small concentration variation) Hoogsteen bonding (substantiated by selective irradiation experiments).

Intermolecular interactions, as proposed in the aforementioned equilibrium model, lead to larger observed curvatures in the signal dependence. A similar situation exists for the A–O–A + T–O–T mixture although to a lesser extent ($\Delta\delta$ in the same concentration range is only 1.3-fold larger for the most upfield resonances). Both amino resonances of A–O–A shift to the same extent.

The implications of dual hydrogen-bonding association modes (*i.e.* intra- and intermolecular) may be of significant biological relevance. Base triplets, such as TAT, are known to play a structural rule in determining and stabilizing the tertiary structure of transfer RNA (tRNA) and, along with CGC^+ triplets, are now well-known motifs in DNA recognition.[24] On the other hand, the existence of base quadruplets, such as ATAT (Figure 17), to the best of our knowledge, has only been assessed experimentally and cytosine derivatives.[25] This recognition mode is thought to play an important role in DNA–DNA recognition and DNA recombination and, in the case of GGGG quadruplets, has been suggested to play an important role in the structure of chromosome

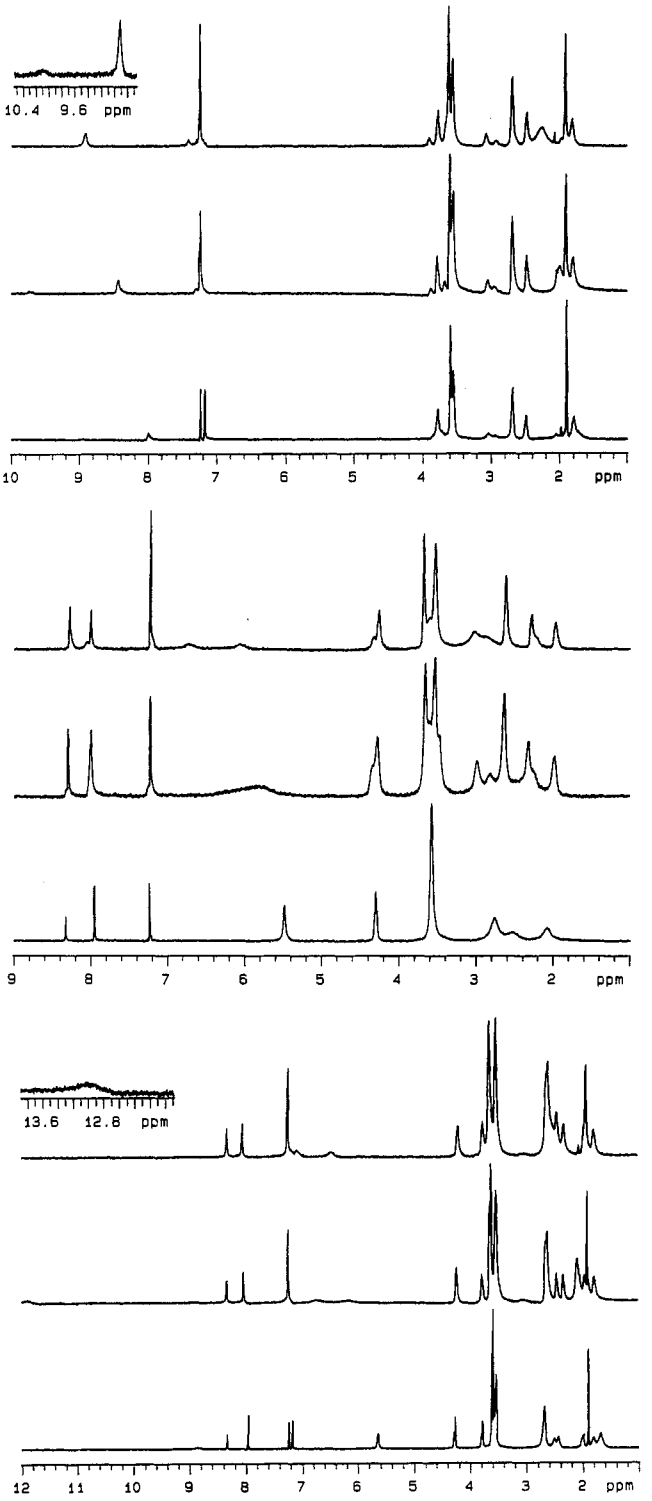


Figure 13. (top) ^{1}H-NMR spectra of T–O–T at 22 °C (lower trace), –30 °C, and –50 °C (upper trace) at 1 mM monomer concentration. (center) ^{1}H-NMR spectra of A–O–A at 22 °C (lower trace), –30 °C, and –50 °C (upper trace) at 1 mM monomer concentration. (bottom) ^{1}H-NMR spectra of the A–O–A + T–O–T mixture at 22 °C (lower trace), –30 °C, and –50 °C (upper trace) at 1 mM monomer concentration.

telomers.[26] Our findings indicate that Hoogsteen rather than Watson–Crick quadruplets predominate, at least with A–O–T.

(24) Griffin, L. C.; Kiessling, L. L.; Beal, P. A.; Gillespie, P.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 7976 and references therein.

(25) (a) Williams, N. G.; Williams, L. D.; Shaw, B. R. *J. Am. Chem. Soc.* **1989**, *111*, 7205. (b) Williams, L. D.; Williams, N. G.; Shaw, B. R. *J. Am. Chem. Soc.* **1990**, *112*, 829.

(26) (a) Chernyi, A. A.; Lysov, Y.; Il'ychova, I. A.; Zibrov, A. S.; Shchyolkina, A. K.; Borisova, O. F.; Mamaeva, O. K.; Florentiev, V. L. *J. Biomol. Struct. Dyn.* **1990**, *8*, 513. (b) Guschlbauer, W.; Chantot, J. F.; Thiele, D. *J. Biomol. Struct. Dyn.* **1990**, *8*, 491. (c) Alberts, B.; Bray, D.; Lewis, L.; Raff, M.; Roberts, K.; Watson, J. D. *Molecular Biology of the Cell*, 2nd ed.; Garland Publishing, Inc.: New York, 1989; pp 481–486.

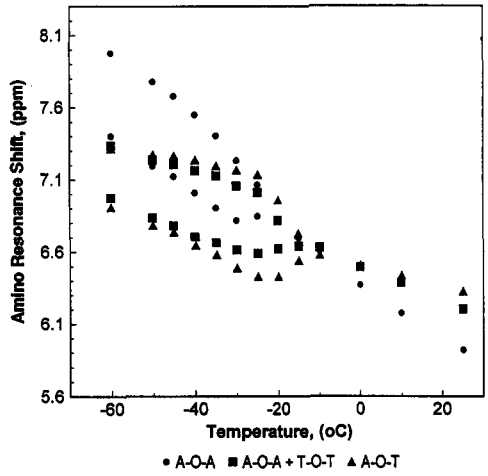


Figure 14. Temperature dependence of the amino groups in the A–O–A (circles) and A–O–T (triangles) monomers and the A–O–A + T–O–T (squares) mixture (25 mM total monomer concentration).

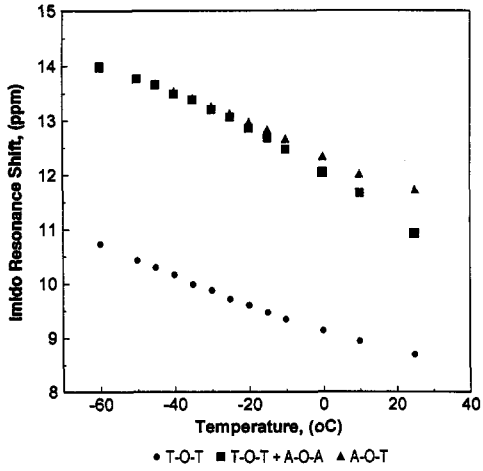


Figure 15. Temperature dependence of the imido groups in the A–O–T (triangles) and T–O–T (circles) monomers and A–O–A + T–O–T (squares) (25 mM total monomer concentration).

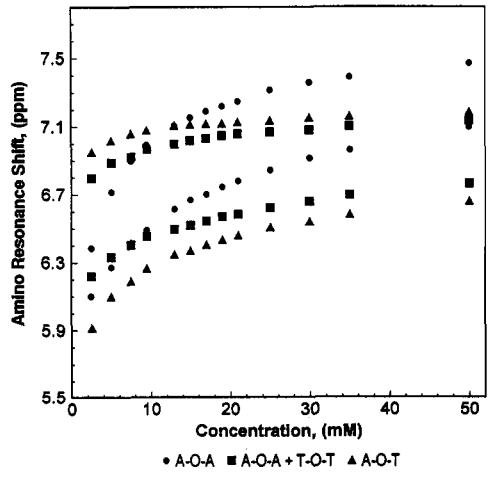


Figure 16. Concentration dependence, at –30 °C, of the different amino resonances in the A–O–A monomer (circles), A–O–T monomer (triangles), and the A–O–A + T–O–T mixture (squares).

Vapor Pressure Osmometry (VPO) Studies of Monomer Aggregation. Although mostly used for molecular weight determinations in aqueous solution, VPO techniques were adapted to give information on the aggregation states of the various nucleotide-base-side-armed monomers described above. The results reflect the extent to which the monomers interact with themselves and/or with each other but do not indicate either the stoichiometry or the association mode. The results of these studies are summarized in Table 1.

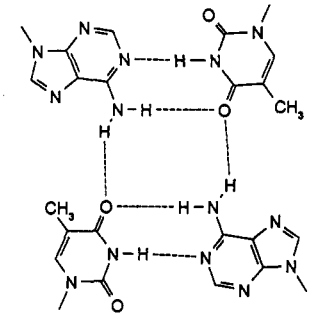


Figure 17. Proposed dual mode of binding for the A–O–A + T–O–T mixture and the A–O–T and A–O–A monomers.

Table 1. Vapor Pressure Osmometry for Nucleotide-Base-Containing Monomers[a]

compd[b]	mol wt (D) theory	mol wt (D) exptl	assocn index[c]
NC(CH$_2$)$_2$⟨N18N⟩(CH$_2$)$_2$CN[d] (**9**)	340.42	339 ● 1	0.99 ± 0.01
A–O–A (**2**)	612.74	859 ± 6	1.39 ± 0.04
T–O–T (**4**)	594.71	852 ± 5	1.43 ± 0.01
A–O–A + T–O–T	603.73[e]	1129 ± 24	1.87 ± 0.06

[a] VPO experiments conducted in $CHCl_3$ at total monomer concentrations of 25 mM (*ca.* 16.7 mmol/kg). [b] Compound studied. See text below for abbreviations. [c] We define "association index" as the ratio of experimentally determined to theoretical molecular weights. [d] ⟨N18N⟩ stands for 4,13-diaza-18-crown-6; dicyano compound **9** was used as the control. [e] The theoretical molecular weight for the 1:1 mixture is (612.74 + 594.71)/2.

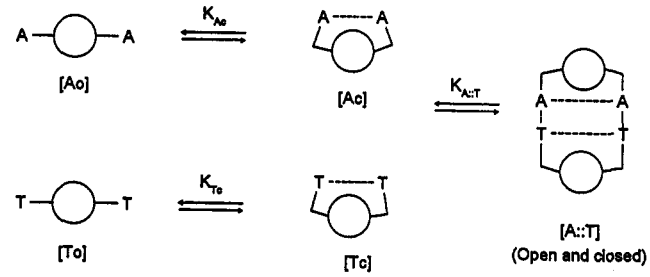


Figure 18. Equilibrium model for determining $K_{A::T}$ for the DNA box.

The molecular weight value obtained for the NC(CH$_2$)$_2$⟨N18N⟩(CH$_2$)$_2$CN control solute established the method's reliability for use in these experiments. Significant association was observed for all monomers. The highest value was obtained for A–O–A + T–O–T as expected for favored Ade::Thy pairing. The individual monomers appeared to associate to about the same extent within experimental error. The values obtained at a monomer concentration of 25 mM were smaller in all three cases compared to those obtained at 37 mM (comparison not shown). Such a concentration dependence was not unexpected (see NMR results above).

Determination of K_{assoc} between the A–O–A and T–O–T Monomers. There are several methodologies and analytical techniques for determining association constants.[27] We chose to monitor by ^{1}H-NMR ($CDCl_3$, 22 °C) the titration of A–O–A with T–O–T. Specifically, changes in the amino group of adenine were followed. The association constant was determined by a nonlinear least-squares fit of the titration curve to the equilibrium model depicted in Figure 18.

The model chosen is relatively simple and considers only intramolecular association of the individual monomers. The

(27) Connors, K. A. *Binding Constants: The Measurement of Molecular Complex Stability*; J. Wiley & Sons; New York, 1987; pp 139–363.

derived association constants should thus be interpreted with due caution. In this milieu, the major driving force for intermolecular association should be Ade::Thy base pairing. In the equilibrium model, K_{Tc} and K_{Ac} represent the intramolecular association constants for the T–O–T and A–O–A monomers, respectively, as isolated species (without contributions from intermolecular interactions). The equilibrium is defined algebraically by eqs 1–6:where $[A]_{total}$ and $[T]_{total}$ represent the total concentrations

$$K_{A_c} = [A_c]/[A_o] \quad (1)$$

$$K_{T_c} = [T_c]/[T_o] \quad (2)$$

$$[A]_{total} = [A_o](1 + K_{A_c}) + [A::T] \quad (3)$$

$$[T]_{total} = [T_o](1 + K_{T_c}) + [A::T] \quad (4)$$

$$K_{A::T} = [A::T]/[A_o][T_o] \quad (5)$$

$$\delta_{obsd} = [A_o]/[A]_{total}(\delta A_o + K_{A_c}\delta A_c) + [A::T]/[A]_{total}(\delta A::T) \quad (6)$$

of the monomers. $[A_o]$ and $[T_o]$ represent the concentrations of "open" monomers, and [A::T], that of the "A::T box". Similarly, δA_c and δT_c stand for the chemical shifts of the intramolecularly associated monomers. δA_o, δT_o, and $\delta A::T$ represent the shifts of the "open" monomers and the Ade::Thy complex, respectively.

These equations were worked into a quadratic expression $ax^2 + bx + c = 0$ (which initially does not consider $[A_c]$ and $[T_c]$), where $x = [A::T]$, $a = 1$, $b = -([A]_{total} + [T]_{total} + 1/K_{A::T})$, and $c = [A]_{total}[T]_{total}$. The data were analyzed using the least-squares-fit software program MINSQ (MicroMath Scientific Software). The results are shown below.[28]

$$K_{T_c} = [T_c]/[T_o] = 0.45\ M^{-1}$$

$$K_{A_c} = [A_c]/[A_o] = 0.34\ M^{-1}$$

$$K_{A::T} = [A::T]/[A_o][T_o] = 855 \pm 94\ M^{-1}$$

$$K_{assoc} = [A::T]/[T]_{total}[A]_{total} = K_{A::T}/(1 + K_{A_c})(1 + K_{T_c}) = 440\ M^{-1}$$

Limiting shifts (δA_o and δT_o) were determined by dilution, at room temperature, of 9-propyladenine and 1-propylthymine until no further shift was detected or until the signal remained discernible from the spectral baseline. This procedure afforded the chemical shifts of the "free" amino and imido resonances. Low-temperature (–60 °C) ^{1}H-NMR of A–O–A and T–O–T (1 mM) afforded approximations for the chemical shifts of the intramolecularly associated species. A typical experimental titration curve is shown in Figure 19.

The values for the intramolecular association constants are lower than we anticipated. Kyogoku *et al.* reported constants of $\approx 3.1\ M^{-1}$ for association of either 9-propyladenine or 1-cyclohexylthymine with itself in chloroform.[29] Our results suggest that intermolecular and intramolecular association compete and decrease each other reciprocally, a possibility discussed above. The association constant ($K_{A:T}$) between A–O–A and T–O–T ($855\ M^{-1}$) is significant if one considers that reported by Kyogoku ($130\ M^{-1}$) for the pairing of adenine and thymine derivatives noted above with each other. This may imply some degree of cooperativity with base pairing. The value of $440\ M^{-1}$ represents a binding constant that takes into account the overall equilibrium.

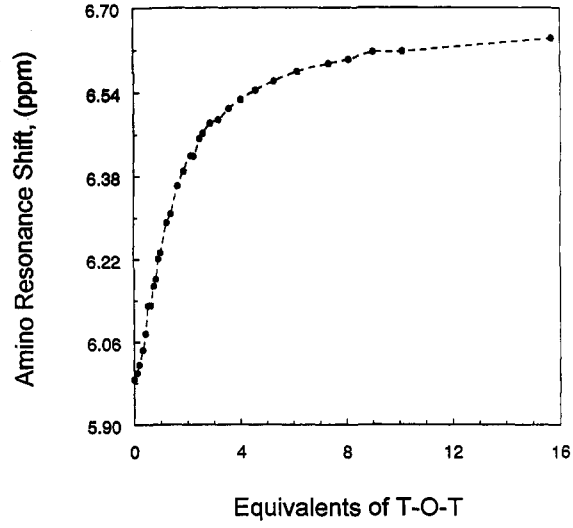


Figure 19. Curve for the titration of the A–O–A monomer. Chemical shift variations of the amino resonance were observed as a function of added complementary monomer T–O–T.

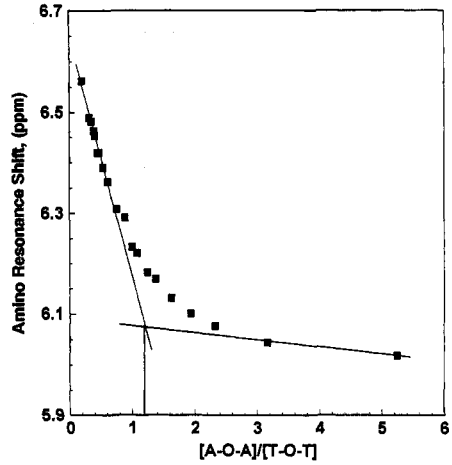


Figure 20. Plot indicating the stoichiometry between the A–O–A and T–O–T monomers. The dotted line suggests a ≈1:1 stoichiometry; lines were not fit mathematically. [A–O–A]/[T–O–T] ratios are displayed only up to a value of 6; however, measurements with ratios of up to ~16:1 were determined.

Some 30 years ago, Ts'o and co-workers determined an association of $192\ M^{-1}$ between thymine and adenine dinucleoside alkyl phosphotriesters in $CDCl_3$ but inferred a non-cooperative process.[30] Sessler *et al.* have determined association constants of 820–1300 M^{-1} for cytosine and guanine derivatives with a designed ditopic receptor in DMSO-d_6.[6c] The contribution to association from H-bonding as compared to Coulombic interactions appeared to be relatively small, however, since triethylamine exhibited about one-half the binding ($K \approx 500\ M^{-1}$) under similar conditions. The groups of Hamilton[31] and Rebek[3a] have observed association constants of 3200 and 90–440 M^{-1}, respectively, in $CDCl_3$ for systems that recognize adenine derivatives using a combination of H-bonding and π-stacking.

From the plot shown in Figure 20 it appears that there may be interactions involving more than two monomers. However, the determined ratio of 1.25 clearly indicates that the monomers interact largely as envisioned (open and closed DNA box formation, see Figure 9; data are shown only up to a ratio of 6).

(28) Relevant statistics are as follows: δA:T = 6.917 ± 0.04 ppm (calculated). Saturation after 16 equiv of T–O–T = 82%. Sum of square deviations = 0.008 09. Standard deviation = 0.019 17.

(29) Iwahashi, H.; Sugeta, H.; Kyogoku, Y. *Biochemistry* **1982**, *21*, 631.

(30) DeBoer, G.; Miller, P. S.; Ts'o, P. O. P. *Biochemistry* **1973**, *12*, 720.

(31) Arnold, K. A.; Viscariello, A. M.; Kim, M.; Gandour, R. D.; Fronczek, F. R.; Gokel, G. W. *Tetrahedron Lett.* **1988**, 3025.

$K_{A:T}$ yields a free energy of association ($-\Delta G°_{298}$) of ≈4 kcal·mol^{-1}. If one considers the presence of four hydrogen bonds in the envisioned DNA box, we may assign a free energy contribution of ≈1 kcal·mol^{-1} per hydrogen bond formed. This energetic contribution is modest but significant considering the feeble nature of the interactions. Values of 1–3 kcal·mol^{-1} per hydrogen bond have been reported for the free energy of binding in related systems.[3,4] The energetic cost of opening a closed hydrogen-bonded conformation and bringing two separate, flexible frameworks together and the concomitant loss of some rotational freedom upon dimerization could well account for this value.

Self-Assembly in the Presence of Complementary Alkanediyl- and Arenediyldiammonium Guests. Diammonium species are excellent complements for crown ethers and may significantly enhance the self-assembly of the DNA box species. Although some alkanediyldiammonium compounds are available commercially, rigid, aromatic diammonium compounds suitable for our experiments required synthesis.

Introduction of aromatic residues in the diammonium cation makes possible stabilization due to π-stacking as well as H-bonding. We recognize the possibility that addition of the diammonium cation to an already complex equilibrium could also afford unexpected complexation modes.

We utilized CPK molecular models for assessing the most appropriate complementary guests (diammonium salts) for these boxes. In a fully extended conformation, all boxes were 17–19 Å long in either a Watson–Crick or Hoogsteen binding mode. Commercially available diaminodecane and diaminododecane are ≈15 and ≈17 Å long, respectively. The diammonium cation resulting from protonation of 4,4′-bis(2-aminoethyl)azobenzene was estimated to be 18 Å long and is colored. Upon the basis of CPK models, considering inherent conformational flexibility, all possible complexation modes seemed plausible.

Synthesis of Diammonium Guests. The general synthetic procedure involved treating a CH_3CN solution of the diamino compounds with $HClO_4$ (room temperature, 1:2 ratio, respectively). Perchloric acid was chosen because its anion is nonnucleophilic. The bis(perchlorates) were isolated by azeotropic removal of water (benzene/absolute EtOH, 5:1) and crystallization from MeCN (pure by ^{1}H-NMR and combustion analysis). 4,4′-Bis(2-aminoethyl)azobenzene was brought to hand and fully characterized by a sequence which involved protection of the amino group of 4-nitrophenethylamine (Aldrich), reductive coupling of the nitro derivatives (Zn/NaOH), and final deprotection (CF_3COOH) to the free amine. The golden orange dye (mp 207–208 °C) was treated as above to form the bis(perchlorate) salt (mp > 300 °C). Unfortunately, complexation studies using this salt could not be conducted due to its poor solubility. Details of the preparation and a solid-state structure of the complex will be reported elsewhere.

Association Studies. NMR studies of complexation using alkyldiyldiammonium cations (as their ClO_4^- salts) in $CDCl_3$ were hampered by insolubility of the salts. No salt could be detected (^{1}H-NMR) in solution after filtering a suspension (10 mM A–O–A and either 5 mM 1,10-diaminodecane·$HClO_4$ or its 12-methylene analog). In contrast, unsubstituted 4,13-diaza-18-crown-6 dissolved the diammonium salts under identical conditions.

The diammonium perchlorate salt derivatives were quite soluble in CD_3CN. However, A–O–A, T–O–T, and their mixture displayed very poor solubility, although T–O–T appeared to be more soluble than A–O–A. Neither warming nor sonicating afforded clear solutions. No assembly could be observed in DMSO-d_6, probably because of its high polarity. We were able, however, to obtain solutions of both host and guest using a combination of solvents. Ultimately, we used a 10% CD_3OH:$CDCl_3$ mixture which is intermediate in polarity and H-bonding ability between $CDCl_3$ and DMSO-d_6. Trideuteriomethanol, rather than CD_3OD, was used because H/D exchange would prevent observation of the amino and imido resonances. In this solvent combination at the concentrations used (10 mM total monomer and 5 mM guest), hydrogen bonding is compromised by the 250- to 500-fold excess of CD_3OH.

Table 2. Chemical Shift Variations Observed on Complexation of Nucleotide-Base Monomers and Diammonium Salts in 10% CD_3OH:$CDCl_3$ at 22 °C

guest[a]	resonance[b]	A–O–A[c]	T–O–T	A–O–A + T–O–T
1,10-decanediyl-diammonium salt	Ade C2-H	0		0
	Ade C8-H	0.084		0.080
	Thy C6-H		0.146	0.142
	N–H[d]	0.101		0.138
	$CH_2CH_2NH_2$	–0.031	–0.016	–0.042
	$-(CH_2)_6-$	–0.046	–0.013	–0.045
1,12-dodecanediyl-diammonium salt	Ade C2-H	0		0
	Ade C8-H	0.089		0.096
	Thy C6-H		0.135	0.160
	N–H	0.089	*e*	0.134
	$CH_2CH_2NH_2$	–0.017	0	–0.021
	$-(CH_2)_8-$	–0.033	0	–0.027

[a] 1,10-Diaminodecane and 1,12-diaminododecane were added as their perchlorate salts. [b] Ade and Thy stand for adenine and thymine; subscripts refer to proton positions within the residues. [c] Shifts are given in ppm (δ) downfield from residual $CHCl_3$. The total concentration of monomers is 10 mM, and that of the guest is 5 mM. The numbers in the table represent Δδ values. For the guests' resonances, the differences relate to experiments using 4,13-diaza-18-crown-6 as the control. For the rest of the signals, these differences relate to experiments with the same monomers but in the absence of guests. Negative and positive signs indicate upfield and downfield shifts, respectively. [d] We could not observe the $-NH_3^+$ resonances. [e] Indicates that the imido resonance was not observed.

We first recorded "reference" spectra of the individual monomers, and their mixture, in the absence of any diammonium salt. Both complexation by the crown ring and shielding from the base residues were expected to shift the alkyldiammonium resonances to higher fields. Reference spectra of diammonium salt compounds were recorded in the presence of 4,13-diaza-18-crown-6 to control for simple H-bonding effects. The results obtained are recorded in Table 2.

From these results, it appears that the binding mode is as envisioned. The magnitudes of the Δδ values are modest, a fact we anticipated because of the CD_3OH concentration (see above). Upon addition of either dialkylammonium salt, the C–H resonances of the bases shifted downfield. As predicted, the N–H resonances of adenine also shifted downfield. This strongly suggests enhanced hydrogen bonding due to box formation. The only potential H-bond acceptors except for the solvent (which remains constant) are adenine's N1 and N7 atoms. Experiments using dodecylammonium perchlorate [$CH_3(CH_2)_{11}NH_2$·$HClO_4$] indicate that the downfield shifts are not due simply to protonation of adenine, a fact that further supports our belief that it arises from increased base pairing. The upfield Δδ values for the internal $-(CH_2)_n-$ resonances indicate a shielding effect. The central methylene resonances are most shifted, probably due to the proximity of the nucleic acid bases as envisioned in our model of the DNA box (Figure 1).

Since the 4,13-diaza-18-crown-6 control is a better binder of both diammonium salts (it dissolves them readily in $CDCl_3$, see above), the differential upfield shifts (Δδ) indicative of shielding due to box formation may actually be underestimated (*i.e.* their intrinsic values should actually be more negative).

The chemical shift (Δδ) values are minimal for the TT box. This is presumably because its cavity size is the smallest or least complementary to the diammonium ions (poor inclusion of the guest) and because thymine is inherently less aromatic than adenine. These is a substantial downfield shift of the C6–H resonance of thymine, perhaps due to increased Thy::Thy or Ade:

:Thy pairing. The $\Delta\delta$ values for the AA and AT boxes are about the same and further assess the major role played by diammonium binding on box formation. On the basis of the amino resonance $\Delta\delta$ values, it appears that Ade::Thy base pairing is favored, if only slightly.

Conclusion

Several novel lariat ether compounds having alkyl sidearms terminated in nucleotide bases were prepared. Three of these, A–O–A, A–O–T, and T–O–T, were examined in detail, and C–O–C was prepared but not evaluated. NMR and VPO solution studies showed that both intramolecular sidearm interactions and aggregation occur. In both intra- and intermolecular interactions, the Hoogsteen hydrogen bonding plays a significant role. The monomers aggregate to form a dimeric box. The association constant ($CDCl_3$) between A–O–A and T–O–T is found by NMR studies to be 855 M^{-1}. Formation of an induced-fit receptor between A–O–A and T–O–T is also reported in which the guest compound is a diammonium salt. The A–O–A·T–O–T box presumably assembles about the salt, binding the latter as a ditopic receptor system. These studies demonstrate that small molecules may be used effectively to probe biologically relevant interactions and also to mimic the fashion in which far larger systems operate.

Experimental Section

^{1}H- and ^{13}C-NMR were recorded on a Varian VXR-400S NMR spectrometer (see below for details) or on a Hitachi Perkin-Elmer R-600 high-resolution NMR spectrometer. $CDCl_3$ was used as the solvent, unless otherwise specified. Chemical shifts are given in ppm (δ) downfield from internal Me_4Si (TMS), from DSS (for D_2O solutions), or from residual, non-deuterated, solvent ($CHCl_3$, (δ 7.240), CD_3–SO–CD_2H (δ 2.495), or HCD_2CN (δ 1.930)) as a reference and are reported in the following order: chemical shift, peak multiplicity (broad, s = singlet, d = doublet, t = triplet, m = multiplet), integration, and assignment. Infrared spectra were recorded on a Perkin-Elmer 599 infrared spectrophotometer and were calibrated against the 1601-cm^{-1} band of polystyrene. DCI and FAB mass spectra were determined on a VG Trio-2 spectrometer, using methane or ammonia as the reagent gas (DCI), or 3-nitrobenzyl alcohol as the matrix (FAB), and are reported as follows: m/z (%FS). Vapor pressure osmometry experiments were carried out using a Wescor model 5500 instrument. Melting points were determined on a Thomas Hoover 6406-K capillary melting point apparatus and are uncorrected. Combustion analyses were performed by Atlantic Microlab, Inc., Atlanta, GA, and are reported as percents. TLC analyses were performed on aluminum oxide 60 F-254, neutral (type E), or on silica gel 60 F-254. Preparative chromatographic columns were packed with aluminum oxide (50 g/g of sample, activated, neutral Brockmann 150 mesh), standard grade. Centrifugally accelerated PTLC was performed on a Harrison Research Model 7924 chromatotron, using 2- and 4-mm-layer-thickness circular plates, coated with either aluminum oxide GF or silica gel GF. All reactions were conducted under dry N_2. Acetonitrile, benzene, and DMF were dried and distilled over CaH_2. THF and Et_2O were dried over and distilled from sodium benzophenone ketyl. Dry solvents were stored over activated molecular sieves (3 Å). All other solvents and reagents were of the best grade available commercially and were used without further purification, unless otherwise stated.

Monomer Association Studies. NMR Sample Preparation. All glassware was dried in an electric oven at 110 °C overnight. All glassware, equipment, and volumetric flasks containing powdered samples were introduced in a dessicator and left *in vacuo* (≤0.1 Torr), over phosphorus pentoxide (P_2O_5), for at least 12 h prior to sample preparation. All solutions were prepared and transferred in a "glovebag" (Aldrich) under dry argon (P_2O_5 dessicant present inside). All deuterated solvents were dispensed from sealed ampules and were of the highest purity available commercially (ISOTEC, Inc., Aldrich Chemical Co.). Water content in samples prepared under such conditions were typically ≈2.5×10^{-4} M, as judged from ^{1}H-NMR integrals.

A–O–A Monomer Titration Experiments. The total monomer concentration, [A–O–A + T–O–T], was kept constant (to minimize concentration-dependent shifts) at 25 mM in all samples. A total of 30 individual samples (to minimize trace water) were prepared and placed in 5-mm (diam) × 20-mm NMR tubes. The ratios of A–O–A:T–O–T were varied from 1:0.11 to 1:16. One-third each of the samples had ratios 1:0.11–1:1, 1:1–1:3, or 1:3–1:16. The amino resonance was monitored as a function of [T–O–T]. A value of 82% saturation was obtained for the amino resonance after addition of 16 equiv of T–O–T. The association constant ($K_{A:T}$) was then determined by nonlinear least-squares fit (software program MINSQ) of the titration curve to the equilibrium model (see text).

Variable-Temperature Experiments. These were carried out using the VT controller of the Varian VXR-400S, which is stable to ±0.2 °C.

Diammonium Salt Addition Experiments. Samples were prepared as previously described. All samples were dissolved in 10% v/v CD_3OH:$CDCl_3$ to minimize any solvent-induced shifts. [Total monomer] = 10 mM; [diammonium cation] = 5 mM.

Deuterium Exchange at Adenine's C8-H. H–D exchange was accomplished by the method of Chan *et al.*[32] D_2O solutions (5 mM 99.96 atom % D) of the appropriate compounds were placed in 5-mm (diam) × 20-mm tubes and purged with argon. The tubes were kept at 95–98 °C for *ca.* 60 min and cooled, and the spectra were recorded. A nonheated sample served as the control.

Vapor Pressure Osmometric Determinations. Preliminary Calibration with aqueous standards was achieved by processing three standards of known osmolality: 100, 290, and 1000 mmol/kg. A bovine serum solution of known osmolality (284–290 mmol/kg) gave a reading of 289 mmol/kg.

Direct Calibration—Organic Solvent. A calibration curve was obtained by using four solutions of 4,13-diaza-18-crown-6 in $CHCl_3$ (HPLC grade, 30.81, 53.03, 72.12, and 91.26 mmol of solute/kg of solvent). The instrument was zeroed by using pure solvent, until a stable reading (±2 mmol/kg) was obtained during three consecutive trials.

Molecular Weight Measurements. The instrument was calibrated using *N,N'*-bis(cyanomethyl)-4,13-diaza-18-crown-6 in $CHCl_3$ (see the Experimental Section for preparation; cation binding was previously reported[31]).

9-(3-Chloropropyl)adenine (1). To adenine (27 g, 0.2 mol) and NaH (60% oil dispersion, 10 g, 0.25 mol) in dry DMF (500 mL) was added, in a stream, 22 mL (0.22 mol) of 1-bromo-3-chloropropane, and the mixture was stirred for 18 h. The suspension was filtered off (Celite) and evaporated *in vacuo*, and the resulting slurry was diluted with water (300 mL). The mixture was extracted ($CHCl_3$, 3 × 800 mL), the extracts were dried ($MgSO_4$), and the solvent was removed. Crystallization from EtOH (180 mL) yielded **1** (16.3 g, 39%) as white crystals (mp 187–189 °C, resolidify, then mp >300 °C). ^{1}H-NMR (60 MHz, DMSO-d_6): 2.18 (m, 2H, CH_2–CH_2–CH_2), 3.58 (t, 2H, –CH_2Cl), 4.23 (t, 2H, Ade–CH_2), 7.10 (broad s, 2H, –NH_2, exchanged in D_2O), 8.05 (s, 2H, C8-H) + C2-H). Anal. Calcd for $C_8H_{10}N_5Cl$: C, 45.40; H, 4.80; N, 33.10%. Found: C, 45.20; H, 4.80; N, 33.00%.

***N,N'*-Bis[3-(9-adeninyl)propyl]-4,13-diaza-18-crown-6 (2).** N^9-(3-Chloropropyl)adenine (**1**) (1.63 g, 7.7 mmol) and 4,13-diaza-18-crown-6 (2.03 g, 7.6 mmol) were ground in a mortar and placed in a two-neck flask. The mixture was heated gradually under nitrogen in an oil bath until melted (110–115 °C). The temperature was maintained at ≈150 °C for 12 h, while the mixture was stirred mechanically with a high-torque stirrer. The flask was cooled, the solid digested in boiling EtOH (50 mL), the suspension filtered off, and the solvent removed *in vacuo*. The resulting thick, slightly green oil was dissolved in $CHCl_3$ and chromatographed over Al_2O_3 (0–6% MeOH:$CHCl_3$). Some monosubstituted product **7** (see below) was recovered from the early fractions. The oily product was crystallized from MeCN (≈100 mL/g) to give **2** (25–35%) as a white amorphous solid (mp 162–164 °C). ^{1}H-NMR: 2.002 (m, 4H, CH_2–CH_2–CH_2), 2.421 (t, 4H N[crown]–CH_2), 2.679 (t, 8H, CH_2–N–CH_2 within crown), 3.554 (t, 8H, O–CH_2–CH_2–N), 3.642 (s, 8H, O–CH_2–CH_2–O), 4.289 (t, 4H, Ade–CH_2), 6.244 (broad s, 4H, –NH_2, exchanged in D_2O), 7.952 (s, 2H, C8-H), 8.325 (s, 2H, C2-H). IR: 3400 (br), 3180, 2970, 2900, 2840, 1690, 1620, 1580, 1490, 1425, 1360, 1340, 1315, 1255, 1215, 1165, 1080 cm^{-1}. MS (DCI): 613 (100), 286 (38), 258 (46), 184 (26), 172 (45), 100 (52), 85 (76), 49 (15). Anal. Calcd for $C_{28}H_{44}N_{12}O_4$: C, 54.89; H, 7.24; N, 27.43%. Found: C, 54.62, H, 7.25; N, 27.35%.

1-(3-Bromopropyl)thymine. To a solution of thymine (12.6 g, 0.1 mol) in excess hexamethyldisilazane (HMDS, 60 mL, 0.28 mol) was added a catalytic amount of (TMS)Cl (6 mL, 0.05 mol), and the mixture was refluxed for 21 h. Excess HMDS was removed *in vacuo* to give crude bis(O-silylated)thymine. The crude material was stirred at room

(32) Chan, S. I.; Schweizer, M. P.; Ts'o, P. O. P.; Helmkamp, G. K. *J. Am. Chem. Soc.* **1964**, *86*, 4182.

temperature with excess 1-bromo-3-chloropropane (180 mL, 1.82 mol) for 10 d, water (400 mL) was added, and the two-phase system was stirred for 30 min. The mixture was then extracted with CH_2Cl_2 (3 × 300 mL), dried over Na_2SO_4, and concentrated. Addition of hexanes (700 mL) precipitated the crude product. Crystallization from absolute EtOH (50 mL) afforded N^1-(3-bromopropyl)thymine (15.1 g, 62%) as white crystals (mp 137–139 °C). ^{1}H-NMR (60 MHz, DMSO-d_6): 1.70 (s, 3H, $-CH_3$), 2.08 (m, 2H, CH_2–*CH_2*–CH_2), 3.53 (m, 4H, $-CH_2Br$ + Thy–CH_2), 7.35 (s, 1H, C6-H), 8.90 (broad s, 1H, imide-H, exchanged in D_2O). Anal. Calcd for $C_8H_{11}N_2O_2Br$: C, 38.89; H, 4.49; N, 11.34%. Found: C, 38.78; H, 4.45; N, 11.15%.

***N,N'*-Bis[3-(1-thyminyl)propyl]-4,13-diaza-18-crown-6 (4). Procedure 1.** 4,13-Diaza-18-crown-6 (1.52 g, 5.80 mmol) in MeCN (20 mL) was added to a solution of 1-(3-bromopropyl)thymine (4.24 g, 17.20 mmol) in MeCN (40 mL). Na_2CO_3 (6.0 g, 56.6 mmol) and NaI (0.03 g, 0.2 mmol) were added to the clear solution. The resulting suspension was stirred at room temperature for 8 d [some monosubstituted crown remained (TLC)]. The solution was warmed (≈45 °C) and stirred for an additional 6 d. The white suspension was cooled, filtered, and evaporated *in vacuo* to afford a thick, pale brown oil. Dichloromethane (30 mL) and C_6H_6 (2 × 40 mL) were added successively and removed in turn under reduced pressure. This procedure afforded a foamy solid that was dissolved in $CHCl_3$ and chromatographed over Al_2O_3 (deactivated by adding 4% v/wt H_2O, 0–3% MeOH:CH_2Cl_2). Those fractions containing impure products were further purified by centrifugally accelerated PTLC (Al_2O_3 rotor, 3.5% MeOH:CH_2Cl_2). Early fractions, upon crystallization from EtOAc, afforded 30–40% of monosubstituted product **5** as slightly yellow crystals. The later fractions afforded a foamy white solid which crystallized from EtOAc/MeOH (15:1, 25 mL) to give **4** (10–15%) as small white crystals (mp 129–130 °C). ^{1}H–NMR: 1.760 (m, 4H, CH_2–*CH_2*–CH_2), 1.877 (s, 6H, $-CH_3$), 2.475 (t, 4H, N[crown]–CH_2), 2.675 (t, 8H, CH_2–N–CH_2 within crown), 3.538 (t, 8H, O–*CH_2*–CH_2–N), 3.583 (s, 8H, O–CH_2–CH_2–O), 3.764 (t, 4H, Thy–CH_2), 7.169 (s, 2H, C6-H), 9.152 (broad s, 2H, imide-H, exchanged in D_2O). IR: 3450 (br), 3230, 3110, 3010, 2915, 2860, 1650, 1445, 1350, 1275, 1245, 1220, 1175, 1100, 1065 cm^{-1}. MS (DCI): 595 (100), 286 (12), 258 (38), 172 (30), 127 (49), 100 (37), 95 (27), 49 (68), 45 (16). Anal. Calcd for $C_{28}H_{46}N_6O_8$: C, 56.55; H, 7.80; N, 14.13%. Found: C, 56.36; H, 7.84; N, 13.95%.

Procedure 2. A suspension of 1-(3-bromopropyl)thymine (5.03 g, 20.36 mmol) in MeCN (60 mL) was stirred until a clear solution was obtained (≈1 h). 4,13-Diaza-18-crown-6 (2.13 g, 8.13 mmol) and Et_3N (2.14 g, 21.16 mmol) were then added, and the solution was stirred for 8 d. After 6 d, a white precipitate appeared. The suspension was filtered, and the white solid was washed with cold MeCN (10 mL) and then crystallized (MeCN, 25 mL/g) to give **4** (2.65 g, 55%) as colorless crystals (mp 130–132 °C).

***N*-[3-(1-Thyminyl)propyl]-4,13-diaza-18-crown-6 (5).** 4,13-Diaza-18-crown-6 (2.0 g, 7.6 mmol), **6** (1.88 g, 7.6 mmol), and Na_2CO_3 (1.62 g, 15.2 mmol) were stirred at room temperature for 8 d. The reaction was filtered and the solvent removed *in vacuo*. Column chromatography (Al_2O_3, 0–3% MeOH:$CHCl_3$) gave an oil. Crystallization from EtOAc afforded the monosubstituted product **5** as slightly yellow crystals (40%, mp 111–112 °C). ^{1}H-NMR: 1.813 (m, 2H, CH_2–*CH_2*–CH_2), 1.913 (s, 3H, $-CH_3$), 2.528 (t, 2H, N[crown]–CH_2), 2.721 (t, 4H, CH_2–N–CH_2, unsubstituted side), 2.773 (t, 4H, CH_2–N–CH_2, substituted side), 3.616 (m, 17H, CH_2–OCH_2CH_2–OCH_2 + NH), 3.830 (t, 2H, Thy–CH_2), 7.280 (s, 1H, C6-H), 9.05 (broad s, 1H, imide-H, exchanged in D_2O). IR: 3650–3200 (br), 2860, 1930, 1660, 1435, 1350, 1275, 1220, 1175, 1100, 1065, 1040, 1000, 945, 885, 805, 765, 735, 680, 645 cm^{-1}. Anal. Calcd for $C_{20}H_{36}N_4O_6$: C, 56.06; H, 8.47; N, 13.07%. Found: C, 55.88; H, 8.40; N, 13.00%.

***N*-[3-(9-Adeninyl)propyl]-4,13-diaza-18-crown-6 (7).** The procedure was identical to that used for **2** except that the reaction time was 6 h. Early fractions collected during the chromatographic separation (Al_2O_3, 0–6% MeOH:$CHCl_3$) yielded **7** (35 ± 5%) as small white needles (mp 72–74 °C), after two crystallizations from MeCN. ^{1}H-NMR: 2.001 (m, 2H, CH_2–*CH_2*–CH_2), 2.435 (t, 2H, N[crown]–CH_2), 2.688 (t, 4H, CH_2–N–CH_2, unsubstituted side), 2.716 (t, 4H, CH_2–N–CH_2, substituted side), 3.574 (m, 17H, CH_2–OCH_2CH_2–OCH_2 + NH), 4.314 (t, 2H, Ade–CH_2), 5.760 (broad s, 2H, $-NH_2$, exchanged in D_2O), 8.016 (s, 1H, C8-H), 8.322 (s, 1H, C2-H). Anal. Calcd for $C_{20}H_{35}N_7O_4$: C, 54.90; H, 8.06; N, 22.41. Found: C, 54.55; H, 7.96; N, 22.10%.

***N*-[3-(1-Thyminyl)propyl]-*N'*-[3-(9-adeninyl)propyl]-4,13-diaza-18-crown-6 (8).** A solution of **7** (0.64 g, 1.46 mmol) in MeCN (30 mL), 1-(3-bromopropyl)thymine (0.73 g, 2.93 mmol), and Et_3N (0.42 mL, 3 mmol) was stirred at room temperature. After 24 h, a thick, pale brown oil separated and 1 mL of absolute EtOH was added. Afer 11 d, the solvent was removed under reduced pressure to afford a brownish foam. Column chromatography (Al_2O_3, 0–3% MeOH:$CHCl_3$) gave a white foam in 55% yield that was homogeneous by TLC. ^{1}H-NMR: 1.680 (m, 2H, Thy–CH_2–*CH_2*), 1.887 (s, 3H, Thy–CH_3), 2.020 (m, 2H, Ade–CH_2–*CH_2*), 2.429 (t, 2H, Thy–$(CH_2)_2$–*CH_2*), 2.485 (t, 2H, Ade–$(CH_2)_2$–*CH_2*), 2.659 (m, 8H, CH_2–N–CH_2), 3.505 (5, 4H, O–CH_2–*CH_2*–N, in thymine's side), 3.549 (t, 4H, O–CH_2–*CH_2*–N, in adenine's side), 3.589 (s, 8H, O–CH_2–CH_2–O), 3.769 (t, 2H, Thy–CH_2), 4.282 (t, 2H, Ade–CH_2), 6.319 (broad s, 2H, $-NH_2$, exchanged in D_2O), 7.142 (s, 1H, C6-H), 8.036 (s, 1H, C8-H), 8.348 (s, 1H, C2-H), 11.551 (broad s, 1H, imido-H, exchanged in D_2O). MS (FAB): 604 (10), 329 (5), 309 (100), 275 (40), 207 (25).

***N,N'*-Bis(cyanomethyl)-4,13-diaza-18-crown-6 (9).** 4,13-Diaza-18-crown-6 (2.0 g, 7.63 mmol) was added to dry acetone (100 mL) followed by addition of Na_2CO_3 (2.12 g, 20 mmol). To the stirred and hot suspension was added, dropwise, chloroacetonitrile (1.73 g, 22.9 mmol) dissolved in 30 mL of the same solvent. The mixture was refluxed for 18 h. After cooling, the suspension was filtered and the solvent was removed *in vacuo* to yield a yellow solid. Column chromatography (Al_2O_3, 0–3% v/v MeOH:CH_2Cl_2) followed by crystallization from THF (50 mL) afforded the product (1.78 g, 68%) as white crystals (mp 107–108 °C). ^{1}H-NMR: 2.77 (t, 8H, CH_2–N–CH_2 within the crown), 3.59 (s, 8H, O–CH_2–CH_2–O), 3.62 (t, 8H, O–*CH_2*–CH_2–N), 3.98 (s, 4H, $-CH_2$-CN). IR: 3420 (br), 2880, 2210, 1970, 1650, 1500, 1490, 1455, 1430, 1400, 1320, 1300, 1280, 1260, 1240, 1190, 1170, 1115 cm^{-1}. Anal. Calcd for $C_{16}H_{28}N_4O_4$: C, 56.45; H, 8.29; N, 16.46%. Found: C, 56.55; H, 8.34; N, 16.52%.

Acknowledgment. We thank the NIH for a grant (GM-36262) that supported this work.