Nucleotide Base Recognition: Synthesis of Artificial Receptors containing Two Distinct Binding Regions for the Complexation of Bis-thymine Derivatives

Andrew D. Hamilton*[a] and Dorothy Little[b]

[a] *Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260, U.S.A.*
[b] *Department of Chemistry, Princeton University, Princeton, NJ 08544, U.S.A.*

Oxidative coupling of a propargyloxy-substituted, diamidopyridine-based receptor for thymine leads to a new dimeric receptor that shows strong binding to bis-thymine derivatives.

In recent publications we have reported a novel approach to the molecular recognition of nucleotide bases.[1] A family of macrocyclic receptors has been prepared combining both hydrogen bonding to the periphery and aromatic stacking to the plane of a single heterocyclic base [Figure 1(a)].[2] Changing the complementarity of the hydrogen bonding component has led to receptors with different specificities for thymine,[3] guanine,[4] and adenine,[5,6] while the orientation of the π-stacking is influenced by the electronic characteristics of the interacting groups.[7] An important step in this work involves covalently linking individual binding units to provide a multiple receptor for substrates containing several nucleotide base derivatives [*e.g.*, Figure 1(b)]. We herein report the synthesis of a double receptor that contains two *two-site* binding regions for thymine and forms strong complexes with bis-thymine derivatives.[8,9]

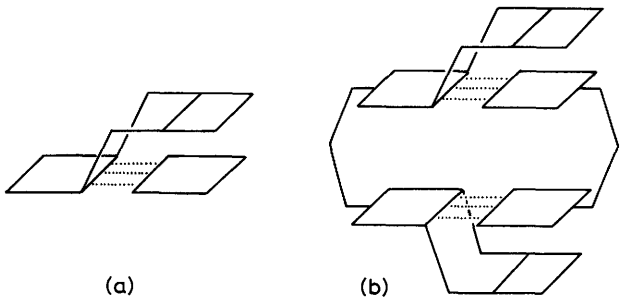
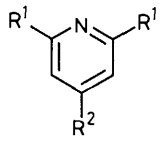


Figure 1

A similar stacking and hydrogen bonding strategy is seen in the DNA-binding antibiotics, triostin[10] and echinomycin,[11] which sandwich a two-base pair sequence between two intercalating quinoxalin rings. Self-association of the quinoxalin rings is inhibited by a bicyclic octapeptide backbone which also forms several hydrogen bonds to the minor groove side of the two-base pairs. In the design of our double receptor, collapse of the cavity will be prevented by positioning a rigid diyne spacer[9b,12] between the two binding regions. This also suggests a facile synthesis of the dimer *via* an acetylene oxidative coupling reaction.[13]

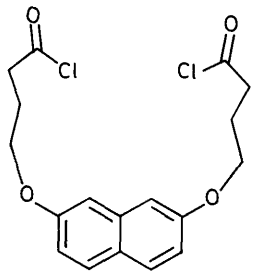
The key 4-propargyloxy-2,6-diaminopyridine (**3**) was prepared by propargylation of dimethyl chelidamate (**1**) (propargyl bromide, Bu^{t}OK) to diester (**2**) followed by conversion



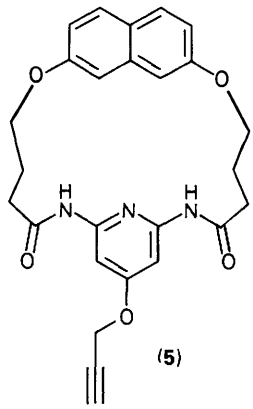
(**1**) $R^1 = CO_2Me$, $R^2 = OH$

(**2**) $R^1 = CO_2Me$, $R^2 = OCH_2CCH$

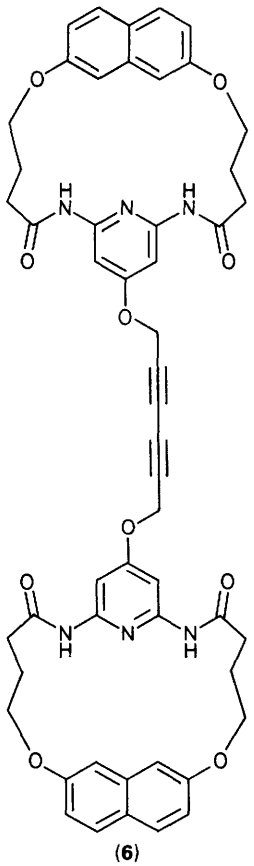
(**3**) $R^1 = NH_2$, $R^2 = OCH_2CCH$



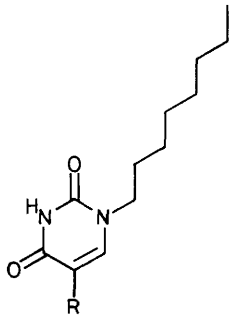
(**4**)



(**5**)



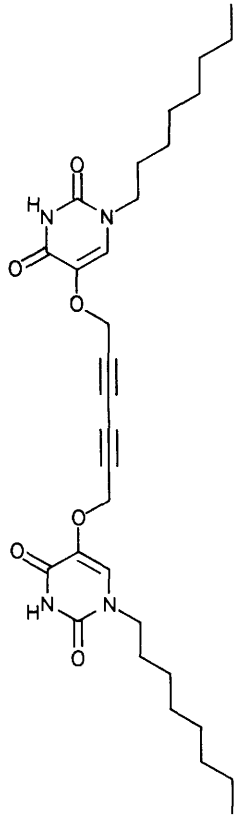
(**6**)



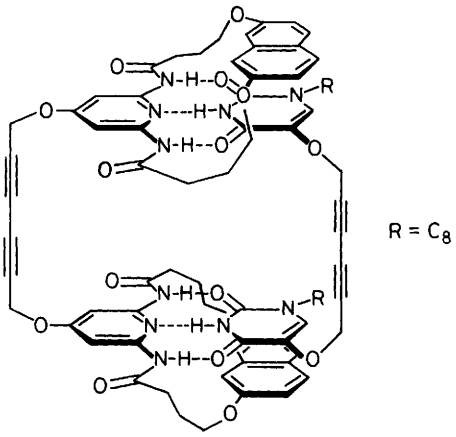
(**7**) R = H

(**8**) R = OH

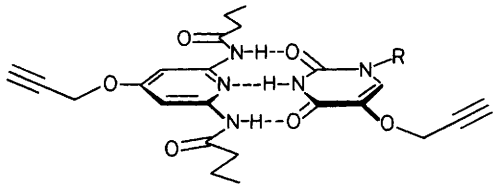
(**9**) R = OCH_2CCH



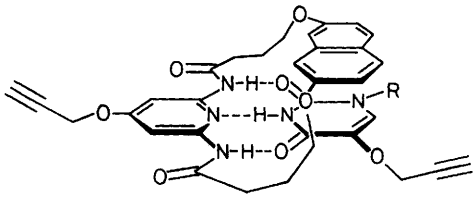
(**10**)



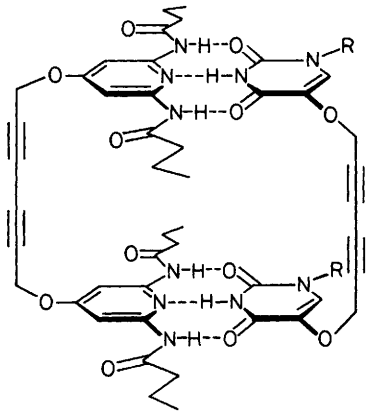
(11)



(12)



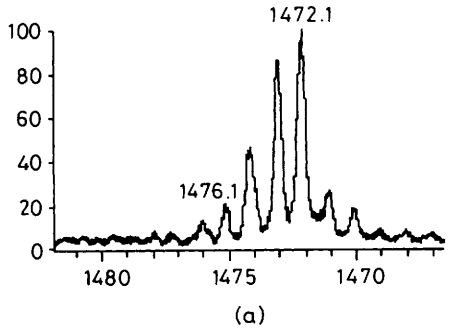
(13)



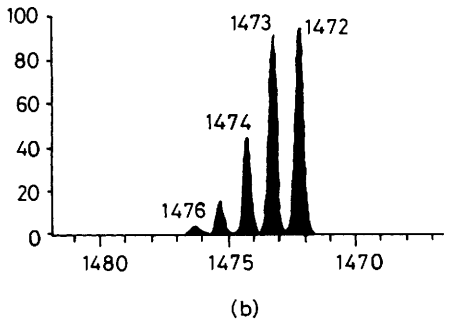
(14)

of the carboxymethyl groups to amines under Curtius conditions.[14] High dilution coupling of (**3**) and diacid chloride (**4**)[3] (CH_2Cl_2, Et_3N) gave macrocycle (**5**)[15] (11% yield) which was oxidatively dimerized [tetramethylethylenediamine (TMEDA), CuCl, O_2] to double receptor (**6**)† in 70% yield. In order to study the double binding properties of (**6**) we have constructed a $CDCl_3$-soluble, bis-thymine derivative containing an analogous diyne spacer. 1-Octyluracil (**7**)[15] was converted to its 5-hydroxy derivative (**8**) (i, Br_2, H_2O; ii, pyridine)[16] which was then propargylated (propargyl bromide, NaOH)[17] to (**9**) and oxidatively dimerized (TMEDA, CuCl, O_2) to bis-thymine (**10**)† in 59% yield from (**8**).

In spacer length as well as hydrogen bonding and aromatic stacking characteristics, double receptor (**6**) provides a



(a)



(b)

Figure 2. (a) $M + H^+$ peak from FAB MS of (**11**). (b) Computer simulation of isotope pattern for $C_{82}H_{91}N_{10}O_{16}$.

† *Selected spectroscopic data* for (**6**): ^1H NMR ($CDCl_3$): δ 7.83 (4H, br. s, NH), 7.69 (4H, d, *J* 9 Hz, naphthalene H-4,5), 7.65 (4H, s, pyridine H), 7.05 (4H, d, *J* 2 Hz, naphthalene H-1,8), 7.01 (4H, dd, *J* 2, 9 Hz, naphthalene H-3,6), 4.88 (4H, s, pyridine OCH_2), 4.24 (8H, t, *J* 6 Hz, OCH_2CH_2), 2.50 (8H, m, CH_2CO), 2.23 (8H, m, OCH_2CH_2). (**10**): ^1H NMR ($CDCl_3$): δ 8.98 (2H, br. s, NH), 7.07 (2H, s, thymine H), 4.80 (4H, s, CH_2O), 3.71 (4H, t, *J* 7.5 Hz, CH_2N), 1.68 (4H, m, CH_2CH_2N), 1.30 [20H, m, $Me(CH_2)_5$], 0.88 (6H, t, *J* 6 Hz, CH_3). Complex (**10**:**6**): ^1H NMR ($CDCl_3$): δ 10.64 (2H, br. s, imide NH), 9.58 (4H, br. s, amide NH), 7.72 (4H, s, pyridine H), 7.33 (4H, d, *J* 9 Hz, naphthalene H-4,5), 6.74 (4H, dd, *J* 9, 2 Hz, naphthalene H-3,6), 6.69 (2H, s, thymine H), 6.58 (4H, d, *J* 2 Hz, naphthalene H-1,8), 4.94 (4H, s, pyridine OCH_2), 4.79 (4H, s, thymine OCH_2), 4.12 (8H, m, CH_2CH_2O), 3.39 (4H, t, *J* 6 Hz, CH_2N), 2.52 (8H, m, CH_2CO), 2.20 (8H, m, CH_2CH_2O), 1.70 (4H, m, CH_2CH_2N), 1.30 [20H, m, $Me(CH_2)_5$], 0.90 (6H, m, Me).

Table 1. Association constant (K_s) values.

Complex	K_s/mol^{-1} dm^3	Complex	K_s/mol^{-1} dm^3
(**12**)	5.21×10^2	(**14**)	4.50×10^3
(**13**)	1.60×10^3	(**11**)	2.03×10^4

matched fit to bis-thymine (**10**) and can readily form a complex of type (**11**).† Addition of one equivalent of (**10**) to a $CDCl_3$ solution of (**6**) gave ^{1}H NMR changes consistent with this formulation.† In particular, large downfield shifts of the amide and imide resonances (δ 1.66 and 1.75) and upfield shifts of the thymine ring–H and N–CH_2 (δ 0.38 and 0.32) and naphthalene–(H-1, -8), and –(H-4, -5) (δ 0.47, 0.27, and 0.36) resonances confirm the presence of both hydrogen bonding and aromatic stacking interactions in (**11**).[3] Further addition of (**10**) to (**6**) gave a saturation curve that established 1 : 1 stoicheiometry and after Foster–Fife[18] analysis yielded an association constant of 2.03×10^4 mol^{-1} dm^3. Evidence that (**11**) is a discrete 1 : 1 complex was obtained from mass spectrometry. Fast atom bombardment (FAB) ionization of a 1 : 1 mixture of (**6**) and (**10**) in *p*-nitrobenzyl alcohol gave an $M + H^+$ peak at $m/z = 1472$ (Figure 2) corresponding to (**11**), with no detectable peaks at higher mass.

The importance of dimerization of both hydrogen bonding and aromatic stacking components was assessed by studying a series of complexes based on monomeric or partial dimeric receptors (**12**—**14**). The association constants for these complexes are collected in Table 1 and show an expected increase in binding strength as the number of binding regions increases from 1 to 4. The K_s values for (**12**) and (**13**) are larger (~5 fold) than those for analogous, unsubstituted amidopyridine receptors[3] due to the increased basicity of the pyridine from the 4-propargyloxy substituent. However, the contribution from π-stacking [3—4 fold increase in K_s from (**12**—**13**)] remains the same.[3,4] Dimerization of the binding regions [(**12**—**14**), (**13**—**11**)] results in a ~10 fold increase in K_s. That the effect of dimerization is not larger may be due to some strain in forming the macrocyclic hydrogen bonded complex.[19]‡

In summary, we have shown that individual receptors for thymine can be dimerized to double receptors that show increased binding to bis-thymine derivatives. These complexes have important implications in construction of supramolecular assemblies held together by hydrogen bonds.[8a] The central cavity of (**11**) and (**14**) resembles that of a two base pair unit of duplex nucleic acids and its intercalation properties are presently being investigated.

We thank the National Institutes of Health (GM35208) for partial support of this work.

Received, 3rd March 1989;§ Com. 9/00942F

References

1 A. D. Hamilton, N. Pant, and A. V. Muehldorf, *Pure Appl. Chem.*, 1988, **60**, 533; A. D. Hamilton, A. Muehldorf, S. K. Chang, N. Pant, S. Goswami, and D. Van Engen, *J. Incl. Phenom.*, 1989, **7**, 27.
2 For non-macrocyclic receptors with similar recognition characteristics, see J. Rebek, Jr., *Science (Washington, DC)*, 1987, **235**, 1478; J. Rebek, Jr., *J. Molec. Recog.*, 1988, **1**, 1.
3 A. D. Hamilton and D. Van Engen, *J. Am. Chem. Soc.*, 1987, **109**, 5035.
4 A. D. Hamilton and N. Pant, *J. Chem. Soc., Chem. Commun.*, 1988, 765.
5 S. Goswami, D. Van Engen, and A. D. Hamilton, *J. Am. Chem. Soc.*, 1989, **111**, 3425.
6 For a different approach to adenine recognition see J. Rebek, Jr., B. Askew, P. Ballester, C. Buhr, S. Jones, D. Nemeth, and K. Williams, *J. Am. Chem. Soc.*, 1987, **109**, 5033. For cytosine binding see K. S. Jeong and J. Rebek, Jr., *J. Am. Chem. Soc.*, 1988, **110**, 3327, and references cited therein.
7 A. V. Muehldorf, D. Van Engen, J. C. Warner, and A. D. Hamilton, *J. Am. Chem. Soc.*, 1988, **110**, 6561.
8 Synthetic bis-hydrogen bonding receptors: see (a) M. Kim and G. W. Gokel, *J. Chem. Soc., Chem. Commun.*, 1987, 1687; (b) D. T. Browne, J. Eisinger, and N. J. Leonard, *J. Am. Chem. Soc.*, 1968, **90**, 7302; (c) F. Kotzyba-Hibert, J. M. Lehn, and P. Vierling, *Tetrahedron Lett.*, 1980, 941; (d) M. R. Johnson, I. O. Sutherland, and R. F. Newton, *J. Chem. Soc., Perkin Trans. 1*, 1980, 587.
9 Synthetic bis-stacking (intercalating) receptors: (a) S. C. Zimmerman and C. M. Van Zyl, *J. Am. Chem. Soc.*, 1988, **109**, 7894; (b) C.-W. Chen, and H. W. Whitlock, *ibid.*, 1978, **100**, 4921; (c) For a review on polyfunctional DNA Intercalators, L. P. G. Wakelin, *Med. Res. Rev.*, 1986, **6**, 275.
10 A. H. J. Wang, G. Ughetto, G. J. Quigley, and A. Rich, *J. Biomolec. Struct. Dyn.*, 1986, **4**, 319.
11 G. Ugetto, A. H. J. Wang, G. J. Quigley, G. A. van der Marel, J. H. Von Boom, and A. Rich, *Nucleic Acids Res.*, 1985, **13**, 2305.
12 E. T. Jarvi and H. W. Whitlock, Jr., *J. Am. Chem. Soc.*, 1980, **102**, 657.
13 A. S. Hay, *J. Org. Chem.*, 1962, **27**, 3320.
14 D. G. Markees, V. C. Dewey, and G. W. Kidder, *J. Med. Chem.*, 1968, **11**, 126; B. Feibush, A. Figueroa, R. Charles, K. D. Onan, P. Feibush, and B. L. Kargar, *J. Am. Chem. Soc.*, 1986, **108**, 3310.
15 Prepared by a modification of C. C. Cheng and L. R. Lewis, in 'Synthetic Procedures in Nucleic Acid Chemistry,' eds. W. W. Zorbach and R. S. Timpson, Interscience, New York, 1968, p. 109.
16 See; D. W. Visser, in 'Synthetic Procedures in Nucleic Acid Chemistry,' eds. W. W. Zorbach and R. S. Timpson, Interscience, New York, 1968, p. 428.
17 P. F. Torrence, J. W. Spencer, and A. M. Bobst, *J. Med. Chem.*, 1978, **21**, 228.
18 R. Foster and C. A. Fife, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1969, **4**, 1. Titrations were carried out at 25 °C. [(**6**)] = 6×10^{-4} M; [(**10**)] = 0.6—6×10^{-3} M. At high guest: host ratios (> 6 : 1) there is a deviation from linearity due to the competitive formation of higher-order aggregates.
19 W. P. Jencks, *Proc. Natl. Acad. Sci. USA*, 1981, **78**, 4046.

‡ An idealized geometry for the complex is shown by (**11**). In a large molecular aggregate of this type a number of conformations are possible. In particular, the relatively minor contribution of π-stacking to the overall binding free energy in (**13**) and (**11**) indicates that, in $CDCl_3$ solution, non-stacked conformations with the naphthalene away from the thymine ring will also be present. In addition, rotation of one of the thymine rings by 180° in (**11**) will lead to a diastereoisomeric complex. These different conformational forms are in fast exchange at room temperature.

§ Received in revised form, 20th September 1989.