Molecular Recognition of Nucleobases via Simultaneous First- and Second-Sphere Coordination

James E. Kickham, Stephen J. Loeb,* and Shannon L. Murphy

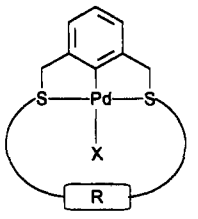
Department of Chemistry and Biochemistry University of Windsor, Windsor, Ontario, Canada N9B 3P4

Received April 5, 1993

The interaction of transition-metal complexes with nucleobases has been studied extensively[1] with emphasis on the development of antitumor agents,[2] reagents for molecular biology,[3] and regulators of gene expression.[4] Concurrently, major advances have been made in designing synthetic receptors for nucleobases by optimizing noncovalent host–guest interactions.[5] Herein, we describe a new approach to designing receptors for the nucleobases. A series of macrocyclic metal complexes has been prepared which bind to a substrate via *simultaneous* first- and second-sphere coordination.[6] This is accomplished by σ-donation to a transition metal (Pd) and noncovalent bonding (hydrogen bonds, $\pi-\pi$ stacking) to peripheral receptor sites on the ligand. Although the combination of metal coordination, hydrogen bonding, and $\pi-\pi$ stacking interactions is common to each nucleobase, the relative placement of these binding sites is different for each and acts as a source of discrimination. Although very few complexes have been characterized which show evidence of simultaneous first- and second-sphere coordination,[6] this type of multiple-point binding has been demonstrated in both the binding of urea and barbiturates to a salicylaldimine-bound UO_2^{2+} moiety[7] and the recognition of amino acids employing Rh and Zn porphyrin complexes.[8]

Chart I shows the metalloreceptors employed in this study. The metal-free macrocycles were prepared in DMF by the Cs^+-mediated ring closure reaction[9] of *m*-xylene-α,α'-dithiol with the appropriate dichloride or ditosylate in 64–86% yield. The palladium and labile CH_3CN group were introduced through direct metalation of the aromatic ring using $[Pd(CH_3CN)_4][BF_4]_2$ in CH_3CN solution,[10] resulting in yields of receptor of 52–77%. Receptors **2** and **3** contain one and three ether oxygens, respectively, for H-bonding to the substrate, and **4** and **5** contain four ether oxygens as well as aromatic units for additional $\pi-\pi$ stacking interactions, while **1** is a model receptor containing none of these secondary binding features.[11]

Chart I



1: R = $(CH_2)_{11}$

2: R = $(CH_2)_2O(CH_2)_2$

3: R = $(CH_2)_2O((CH_2)_2O)_3(CH_2)_2$

4: R = *p*-$CH_2(C_6H_4)O((CH_2)_2O)_2(C_6H_4)CH_2$

5: R = *m*-$CH_2(C_6H_4)O((CH_2)_2O)_3(C_6H_4)CH_2$

X = CH_3CN

Extraction experiments with adenine, cytosine, guanine, and thymine were performed in $(CD_3)_2CO$ and CD_3CN solutions.[12] The results are summarized in graphical form in Figure 1 and provide a good qualitative picture of the relative affinities of these receptors for the four nucleobases.

Cytosine has potential metal-coordination and hydrogen-bonding sites at N3 and N4, two of the three sites employed in Watson–Crick base pairing. Figure 1 shows that, in CD_3CN solution, receptor **3** is selective for cytosine, exhibiting a 56% extraction value compared to less than 5% for the other nucleobases. The ^1H NMR spectrum (CD_3CN) of an isolated sample of [**3**:cytosine][BF_4] showed free *and* H-bonded NH_2 signals at 6.71 and 7.28 ppm for bound cytosine compared to 6.69 ppm for [**1**:cytosine][BF_4], in which no H-bonding can occur.[5] Figure 2 shows the result of an X-ray structural study[13] of [**3**:cytosine][BF_4] which confirms the presence of this multiple-interaction bonding mode. The aromatic N atom of cytosine is bonded to the Pd atom (Pd—N3C, 2.171(6) Å) with simultaneous H-bonding of the NH_2 group to the crown ether portion of the macrocycle (N4C···O1, 3.06(1) Å).

Guanine has potential metal-coordination and H-bonding sites at either N7 and N2 or N3 and N2; however, only one isomer was observed. Extraction experiments show guanine to be the least extracted of adenine, cytosine, and guanine, while thymine was not extracted in any appreciable amount since it lacks the aromatic N-donors for σ-donation to Pd. Significant extraction values for guanine were obtained only in $(CD_3)_2CO$ for **4** and **5**, which contain the aromatic spacing units. ^1H NMR spectra of receptor/purine complexes of **4** and **5** exhibited *upfield* shifts in the range 0.5–0.6 ppm for sets of aromatic protons on the receptors, consistent with significant $\pi-\pi$ interactions[5] accompanying the coordination to Pd. The details of this can been seen in the

(1) Lusty, J. R., Ed. *CRC Handbook of Nucleobase Complexes: Transition Metal Complexes of Naturally Occurring Nucleobases and Their Derivatives*; CRC Press: Boca Raton, FL, 1990; Vol. I, pp 9–99, and references therein.

(2) Tullius, T. D. In *Metal–DNA Chemistry*; Tulius, T. D., Ed.; ACS Symposium Series 402; American Chemical Society: Washington, DC, 1989; Chapter 1 and references therein.

(3) Barton, J. K. *Comments Inorg. Chem.* **1985**, *3*, 321 and references therein.

(4) Pyle, A. M.; Barton, J. K.; In *Progress in Inorganic Chemistry, Bioinorganic Chemistry*; Lippard, S. J., Ed.; J. Wiley and Sons: New York, 1990; Vol. 38, p 413 and references therein.

(5) (a) Zimmerman, S. C.; Wu, W.; Zeng, Z. *J. Am. Chem. Soc.* **1991**, *113*, 196 and references therein. (b) Tjama, T.; Deslongchamps, G.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1990**, *112*, 8408. (c) Jeong, K. S.; Tjivikua, T.; Muehldorf, A.; Deslongchamps, G.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 201. (d) Deslongchamps, G.; Galan, A.; de Mendoza, J.; Rebek, J., Jr. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 61 and references therein. (e) Osterberg, C. E.; Arif, A. M.; Richmond, T. G. *J. Am. Chem. Soc.* **1988**, *110*, 6903. (f) Smith, D. P.; Baralt, E.; Morales, B.; Olmstead, M. M.; Maestre, M. F.; Fish, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 10647 and references therein.

(6) (a) Colquhoun, H. M.; Stoddart, J. F.; Williams, D. J. *Angew. Chem., Int. Ed. Engl.* **1986**, *25*, 487. (b) Stoddart, J. F.; Zarzycki, R. In *Cation Binding by Macrocyclic Systems*; Inoue, Y., Gokel, G. W., Eds.; Marcel Dekker: New York, 1990; p 631 and references therein.

(7) (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) van Doorn, A. R.; Rushton, D. J.; van Straaten-Nijenhuis, W. F.; Verboom, W.; Reinhoudt, D. N. *Recl. Trav. Chim. Pays-Bas* **1992**, *111*, 421.

(8) (a) Aoyama, Y.; Yamagishi, A.; Asagawa, M.; Toi, H.; Ogoshi, H. *J. Am. Chem. Soc.* **1988**, *110*, 4076. (b) Mizutani, T.; Ema, T.; Yoshida, T.; Kurado, Y.; Ogoshi, H. *Inorg. Chem.* **1993**, *32*, 2072.

(9) (a) Buter, J.; Kellogg, R. M. *J. Org. Chem.* **1981**, 4481–4486. (b) Buter, J.; Kellogg, R. M. *Org. Synth.* **1987**, *65*, 150–152.

(10) Geisbrecht, G. R.; Hanan, G. S.; Kickham, J. E.; Loeb, S. J. *Inorg. Chem.* **1992**, *31*, 3286.

(11) NMR data ($(CD_3)_2CO$) for **1**: δ 6.99–7.04 (m, 3H), 4.46 (dd, 4H), 3.15 (q, 4H), 2.12 (m, 8H), 1.27–1.61, (m, 10H). For **2**: δ 6.95–7.03 (m, 3H), 4.44 (dd, 4H), 4.08 (d, 2H), 3.64 (d, 4H), 3.10 (br s, 2H), 2.07 (s, 3H). For **3**: δ 6.99–7.08 (m, 3H), 4.56 (s, 4H), 4.21 (s, 4H), 3.71 (s, 4H), 3.65 (s, 4H), 3.36 (t, 4H), 2.21 (s, 3H). For **4**: δ 7.54 (d, 4H), 6.98–7.0 (m, 7H), 4.51 (br s, 8H), 4.19 (s, 4H), 3.80 (m, 4H), 3.64 (s, 4H), 2.05 (s, 3H). For **5**: δ 6.98–7.35 (m, 11H), 4.50 (br s, 8H), 4.23 (t, 4H), 3.84 (t, 4H), 3.68 (s, 4H), 2.11 (s, 3H). Satisfactory elemental analyses were obtained for **1–5**.

(12) In a typical extraction experiment, 0.8 mL of a 0.01 M solution of receptor (CD_3CN or $(CD_3)_2CO$) was sonicated for 15 min with a 10-fold excess of solid nucleobase, the mixture filtered, and the ^1H NMR recorded.

(13) (a) Crystal data: [**3**:cytosine][BF_4], $C_{22}H_{31}BF_4N_4O_4PdS_2$, monoclinic, $P2_1/n$, a = 12.877(4), b = 16.567(3), and c = 13.741(5) Å, β = 101.53(3)°, V = 2866(3), Z = 4, d(calcd) = 1.56 g cm^{-3}, μ(Mo Kα) = 8.365 cm^{-1}, Rigaku AFC6S diffractometer, 2335 unique reflections with $F_o^2 > 3\sigma F_o^2$, 242 variables, R = 4.57, R_w = 3.99%. (b) Crystal data: [**5**:guanine(BF_3)][BF_4], $C_{34}H_{37}B_2Cl_3F_7N_5O_5PdS_2$, monoclinic, $P2_1/c$, a = 14.589(6), b = 20.444(6), and c = 14.849(4) Å, β = 110.78(2)°, V = 4140(4), Z = 4, d(calcd) = 1.65 g cm^{-3}, μ(Mo Kα) = 8.102 cm^{-1}, Rigaku AFC5R diffractometer, 3112 unique reflections with $F_o^2 > 3\sigma F_o^2$, 363 variables, R = 5.97, R_w = 5.38%. (c) TEXSAN-TEXRAY Structure Analysis Package, Molecular Structure Corporation (1985). (d) Cromer, D. T.; Waber, J. T. *International Tables for X-ray Crystallography, Vol. IV*; The Kynoch Press: Birmingham, United Kingdom, 1974; Table 2.2A. (e) Ibers, J. A.; Hamilton, W. C. *Acta Crystallogr.* **1974**, *17*, 781. (f) Cromer, D. T. *International Tables for X-ray Crystallography, Vol. IV*; The Kynoch Press: Birmingham, United Kingdom, (1974); Table 2.3.1.

0002-7863/93/1515-7031$04.00/0

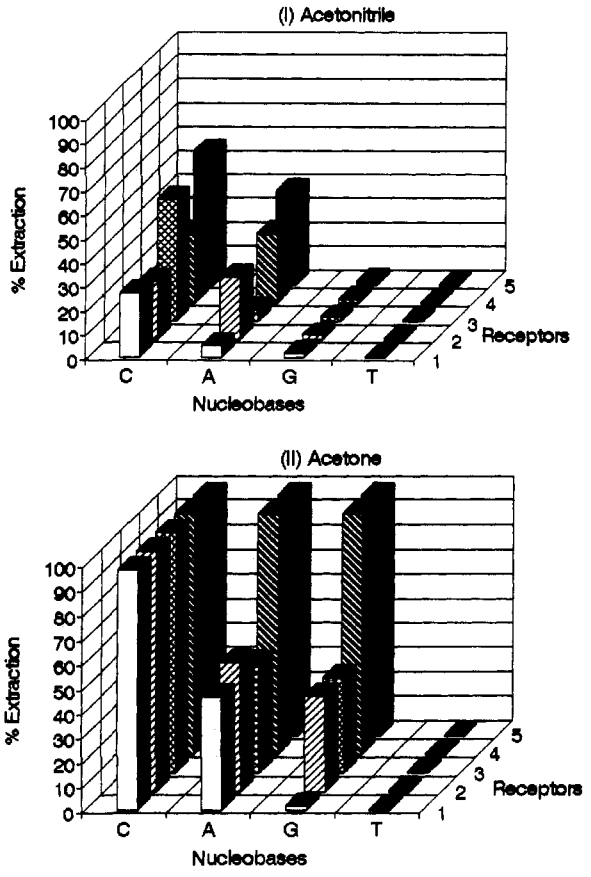


Figure 1. Plots of percent extraction of solid nucleobase (cytosine (C), adenine (A), guanine (G), and thymine (T)) into solution of receptors **1–5**. (I) CD_3CN and (II) $(CD_3)_2CO$.

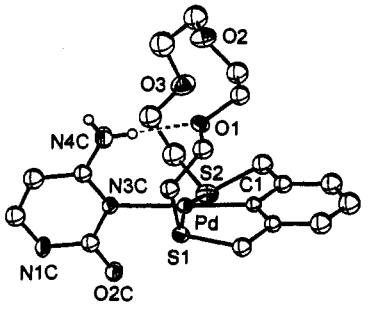


Figure 2. Perspective ORTEP view of the [**3**:cytosine]$^+$ cation. C atoms are numbered sequentially beginning at C1. Pd–S1 2.308(3), Pd-S2 2.300(3), Pd–C1 1.993(7), and Pd–N3C 2.171(6) Å; S1–Pd–S2 161.80-(8), S1–Pd–C1 85.5(2), S2–Pd–C1 81.7(2), S1–Pd–N3C 92.0(2), S2–Pd–N3C 101.2(2), and C1–Pd–N3C 176.7(3)°; N4C···O1 3.06(1) Å.

structure of [**5**:guanine(BF_3)][BF_4], Figure 3. The guanine derivative is bound to Pd at N7 (Pd—N7G, 2.155(7) Å) while being simultaneously π-stacked between the aromatic rings of the receptor. The three aromatic units are essentially parallel, with interplanar distances between receptor and substrate of *ca.* 3.35 Å. Most importantly, the fused-ring carbon atom C5G is positioned exactly between the centers of the parallel aromatic rings of the receptor, an "offset" situation that maximizes the

(14) (a) Desiraju, G. R. In *Crystal Engineering: The Design of Organic Solids*; Materials Science Monographs 54; Elsevier Publishers: New York, NY, 1989; pp 92–108. (b) Hunter, C. A.; Saunders, J. K. M. *J. Am. Chem. Soc.* **1990**, *112*, 5525. (c) Burley, S. K.; Petsko, G. A. *J. Am. Chem. Soc.* **1986**, *108*, 7995.

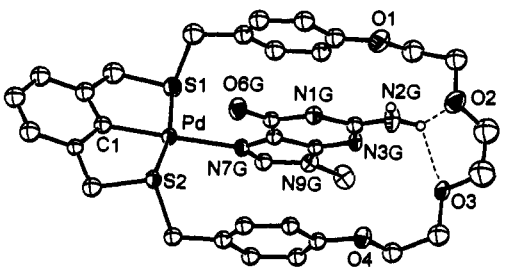


Figure 3. Perspective ORTEP view of the [**5**:guanine(BF_3)]$^+$ cation. C atoms are numbered sequentially beginning at C1. F atoms have been omitted for clarity. Pd–S1 2.296(3), Pd–S2 2.294(3), Pd–C1 2.00(1), and Pd–N3C 2.155(7) Å; S1–Pd–S2 168.6(1), S1–Pd–C1 84.2(3), S2–Pd–C1 85.1(3), S1–Pd–N3C 98.6(2), S2–Pd–N3C 92.1(2), and C1–Pd–N3C 177.1(4)°; N2G···O2 3.16(1) and N2G···O3 2.87(1) Å.

degree of π–π interaction[14] and can only occur with coordination to the Pd atom at N7. In addition, the amino group is hydrogen-bonded in a bifurcated mode to the two aliphatic ether oxygen atoms (N2G···O2, 3.16(1); N2G···O3, 2.87(1) Å).

Adenine has potential metal-coordination and hydrogen-bonding sites at N1 and N6 (employed in Watson–Crick base pairing), N7 and N6 (employed in Hoogsteen base pairing), or N3 and N6 (Type III). In extraction experiments with **1**, **2**, and **3**, two isomers for an adenine-bound substrate were observed, and ^{1}H NMR spectra in $(CD_3)_2CO$, employing 1:1 solutions of receptor and 9-ethyladenine, showed isomer ratios of 1.03:1, 1.96:1, and 1.64:1, respectively. These ratios represent a competition between Watson–Crick and Hoogsteen-type binding modes since CPK models show that Type III binding is not feasible with these receptors. Watson–Crick–type binding allows for optimization of the H-bonding interaction, but the higher basicity of N7 favors the Hoogsteen-type binding mode. In the case of **1**, in which no H-bonding occurs, the WC:H ratio is close to 1:1, but when the possibility of H-bonding is introduced for **2** and **3**, the Watson–Crick-type interaction is favored. This is more pronounced for **2**, and CPK models show that, although both **2** and **3** can undergo H-bonding, the short polyether chain for **2** is locked in a rigid, preorganized conformation with a smaller spacing between the Pd and O binding sites which favors Watson–Crick-type binding. The presence of π–π stacking allows for higher percent extraction values for **4** and **5**. Although CPK models show that placements of the Pd and ether oxygen atoms might actually favor a Type III binding mode, the predominate isomer is likely the Hoogsteen binding type. This orientation of the substrate would be favored since it involves the more basic N7 donor and optimizes the π–π stacking interactions, as observed in the solid-state structure of [**5**:guanine(BF_3)][BF_4].

The ability of this series of metalloreceptors to exhibit simultaneous first- and second-sphere coordination with nucleobase substrates has been verified. The molecular recognition properties of these receptors has also been demonstrated in a qualitative manner, and the potential of these complexes for application to biological systems is the subject of a current study.

Acknowledgment. We thank NSERC of Canada for financial support of this research, and S.L.M. is grateful for an NSERC of Canada undergraduate research scholarship.

Supplementary Material Available: Tables of X-ray experimental details, atom positional parameters, bond distances and angles, thermal parameters, and hydrogen atom parameters (22 pages); listings of observed and calculated structure factors for [**3**:cytosine][BF_4] and [**5**:guanine(BF_3)][BF_4] (38 pages). Ordering information is given on any current masthead page.