*Journal of Inclusion Phenomena and Molecular Recognition in Chemistry* 7 (1989), 27–38.



# Nucleotide Recognition by Macrocyclic Receptors

ANDREW D. HAMILTON,* ALEX MUEHLDORF, SUK-KYU CHANG,**
NALIN PANT, SHYAMAPROSAD GOSWAMI, and DONNA VAN ENGEN
*Department of Chemistry, Princeton University, Princeton, NJ 08544, U.S.A.*

(Received: 1 February 1988)

**Abstract.** It is shown that complementary positioning of recognition sites (particularly hydrogen bonding, stacking and hydrophobic groups) into a macrocyclic structure can lead to very strong and specific complexation of uncharged organic molecules.

**Key words.** Molecular recognition, nucleotide, thymine, guanine, barbiturate.

## 1. Introduction

The recognition and binding of nucleotide substrates by proteins is at the heart of gene expression and metabolic control. The possible future design of 'synthetic repressor' molecules that might artificially activate or control genes depends on an understanding of the key features of nucleotide recognition. The most important recognition features on each nucleotide base are the hydrogen bonding groups at its periphery (Figure 1). These form the basis of the Watson-Crick hydrogen-bonding

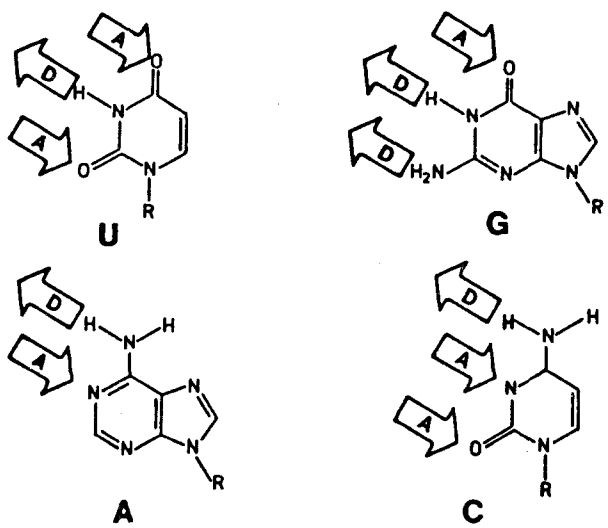


Fig. 1.

* Address correspondence to this author at Department of Chemistry, University of Pittsburgh, Pittsburgh, PA.
** On leave from Chung-Ang University, Seoul, Korea.

scheme and the double helical structure of DNA [1]. However, additional recognition can occur perpendicularly to the hydrogen bonding groups by means of a stacking interaction to the plane of the aromatic base. It is significant that several key nucleotide-binding proteins use both hydrogen bonding and $\pi$–$\pi$ stacking forces to provide strong and selective complexation [1]. For example, ribonuclease $T_1$ binds guanine via two hydrogen bonds between the peptide backbone and the O(6) and NH(1) groups of guanine *plus* a stacking interaction (at 3.4 Å) between a tyrosine residue (Tyr-45) and the purine plane (Figure 2) [2]. The electron density map [2] further suggests that Tyr-45 swings from an 'unbound' conformation into a stacking position on guanine binding, exemplifying an induced fit recognition process [3].

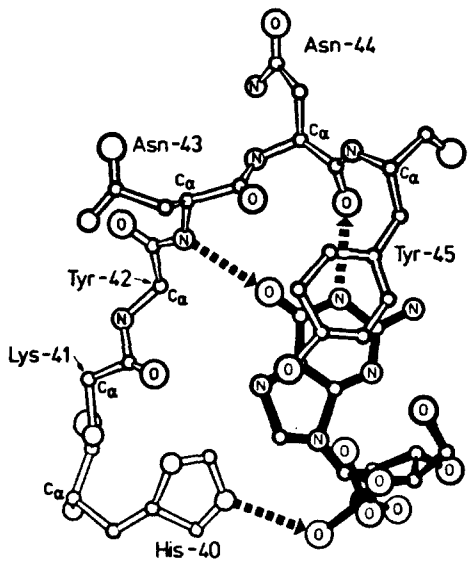


Fig. 2. X-ray structure of ribonuclease-$T_1$ active site.

In designing a new class of artificial receptors for nucleotides we sought to incorporate both recognition features. Our strategy (Figure 3) was to link, within a macrocycle, a group capable of stacking with the nucleotide base to one complementary to its hydrogen bonding periphery. Eventually several of these monomeric hosts might be linked to form a receptor for specific sequences of nucleotide bases.

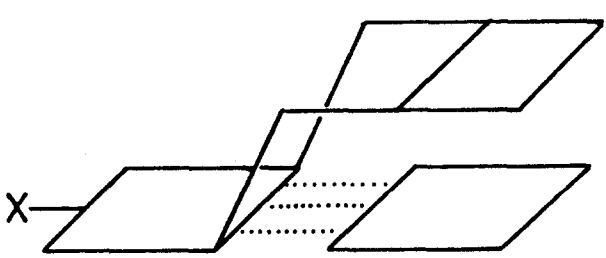


Fig. 3.

## 2. Thymine Recognition

The first receptor studied [4] was that for thymine and was based on the triple hydrogen-bonding complementarity between 2,6-diaminopyridines and the imide group of thymine [5] (Figure 4). The stacking unit was derived from 2,7-dihydroxynaphthalene, a component known for its intercalating properties in several DNA-binding drugs (e.g. neocarzinostatin) [6]. The two groups were linked into a macrocycle (**3**) via the synthesis shown in Figure 5. Reaction of 2,7-dihydroxynaphthalene with an ethyl bromoalkanoate (acetone, $K_2CO_3$) gave diester (**1**) which was converted into its diacid chloride **2** (a, HCl-acetone; b, $(COCl)_2$) and reacted, under high dilution conditions with 2,6-diaminopyridine to give **3**. The yield of the final macrocyclization step varied from 20–26% according to ring size.

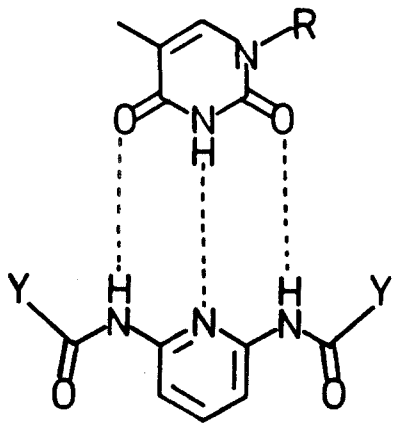

Fig. 4.

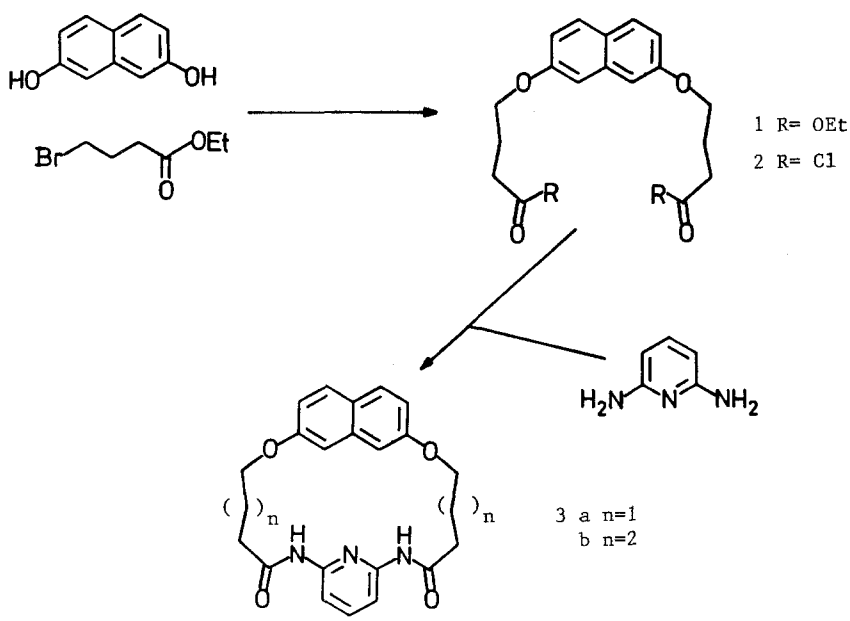


Fig. 5.

The open structure of **3a** (as opposed to an intramolecularly stacked conformation) was confirmed by X-ray crystallography [4] which showed the naphthalene poised at a 127.5° angle to the pyridine ring (Figures 6 and 8a). The thymine recognition properties of **3** were studied using $^1$H NMR and X-ray crystallography. Treatment of a $CDCl_3$ solution of **3a** with one equivalent of *N*-butylthymine (**4**) caused several characteristic changes in the $^1$H NMR of both host and guest. The NH proton resonances on both **3a** and **4** are shifted downfield by 2.25 and 2.6 ppm respectively, confirming the formation of a triple hydrogen bond complex. In addition, upfield shifts (0.19, 0.29 and 0.24 ppm) are seen in the ring-proton, -methyl and -N-methylene resonances of the thymine substrate. These are consistent with the approach of the naphthalene to the thymine plane and the influence of its ring current on the nearby protons on the substrate (the terminal alkyl $CH_3$ on **4** experiences no upfield shift). The X-ray structure of the complex (Figure 7) confirms the ditopic nature of the two recognition sites. Three hydrogen bonds are formed between the pyridine and thymine rings at distances of N—O, 2.87, 2.99 and N—N, 3.06 Å. The naphthalene ring lies directly above the substrate at an angle of 14° and closest inter-plane contact of 3.37 Å. The close contact as well as the ring current induced shifts in the $^1$H NMR confirm that there is a strong correlation between solution and solid state structures. Figures 8a and b show side views of both free and bound forms of **3a** and clearly demonstrate a 'molecular hinge-like motion of the macrocycle. On substrate binding, the naphthalene swings through an arc of 34.1° to within Van der Waals distance (3.4 Å) of the thymine. This induced

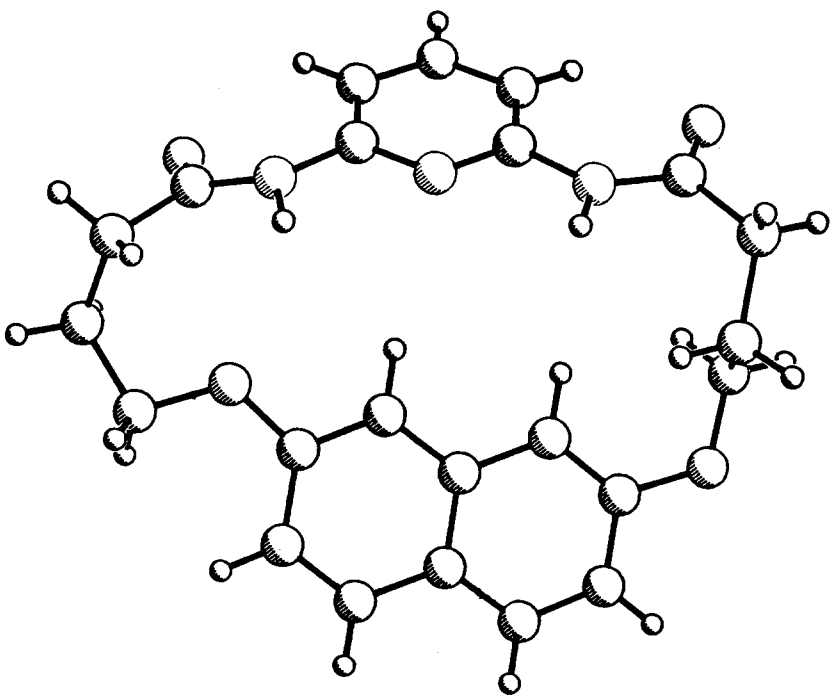

Fig. 6.

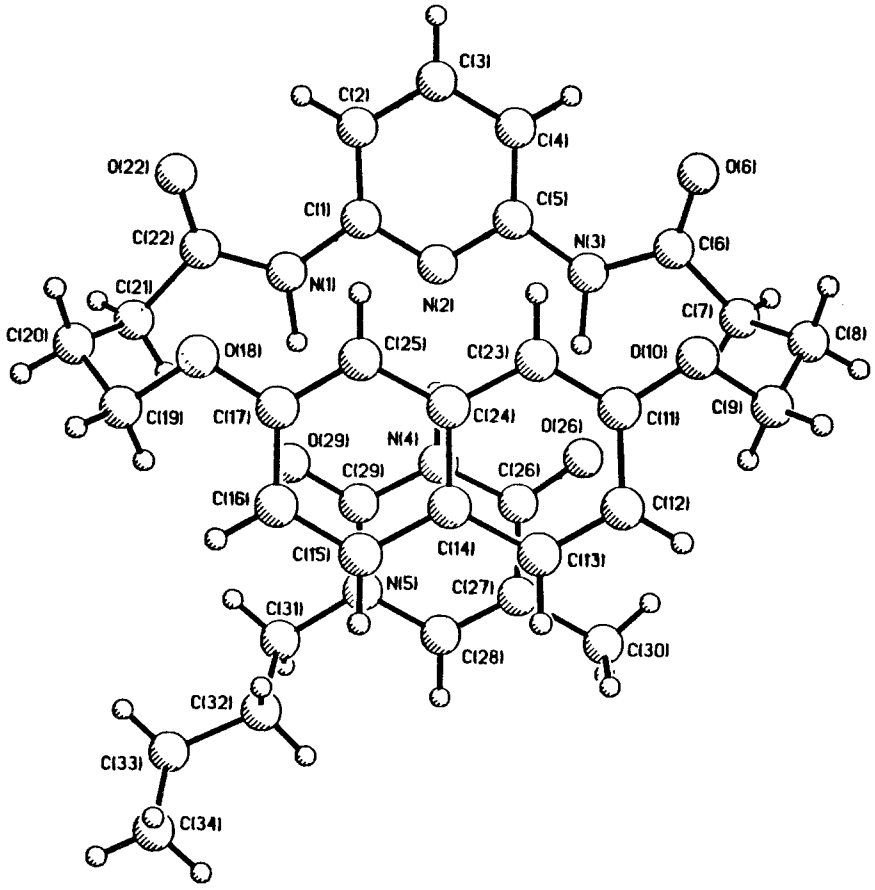
C(3)
C(2)
C(4)
O(22)
O(6)
C(1)
C(5)
C(22)
N(3)
C(6)
C(21)
N(1)
N(2)
C(7)
C(20)
C(8)
O(18)
C(25)
C(23)
O(10)
C(19)
C(17)
C(24)
O(26)
C(11)
C(9)
O(29)
N(4)
C(29)
C(26)
C(16)
C(12)
C(15)
C(14)
C(13)
N(5)
C(27)
C(31)
C(30)
C(28)
C(32)
C(33)
C(34)

Fig. 7.

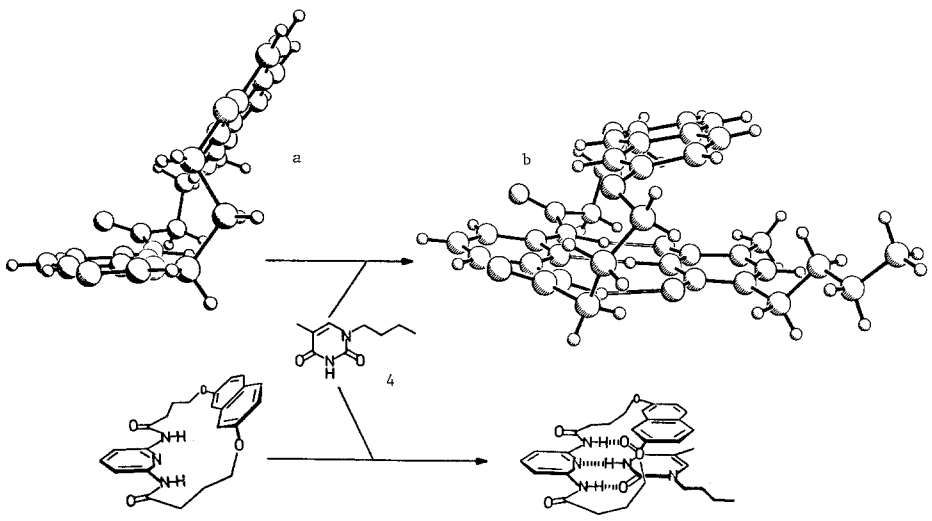
a
b
4

Fig. 8.

fit-like behavior mimics that of the tyrosine residue of ribonuclease $T_1$ which stacks with the guanine substrate.

The ditopic nature of the binding of thymine by **3a** is further reflected in its association constant. These were measured by monitoring the chemical shift of different protons in **3a** as a function of *N*-butylthymine concentration and then analyzing the titration data by means of a Foster-Fife analysis [7]. Both **3a** ($K_s = 290\ M^{-1}$) and the larger **3b** ($K_s = 251\ M^{-1}$) show an approximately threefold increase in association constant for binding to *N*-butylthymine compared to simple 2,6-dibutramidopyridine ($K_s = 90\ M^{-1}$) which lacks the stacking component. This corresponds to a contribution of $\sim 0.75$ kcal/mol from the naphthalene–thymine interaction to the overall binding free energy for **3a : 4** ($\sim 3.5$ kcal/mol). Similar binding enhancements have been seen in related ditopic receptors for adenine [8].

## 3. Guanine Recognition

The hydrogen bonding component of the nucleotide receptor can be readily varied in order to modify its substrate selectivity. For example, we have recently [9] prepared a ditopic (hydrogen bonding and stacking) receptor for guanine based on the triple hydrogen bond complementarity between 7-amino-1,8-naphthyridines and the peripheral hydrogen bonding groups of guanine (Figure 9). The necessary 2-substituted-7-amino-1,8-naphthyridine was synthesized by the route outlined in Figure 10. Condensation of 2,6-diaminopyridine and ethyl acetoacetate in $H_3PO_4$ gave aminonaphthyridone **5** which after acetylation with acetic anhydride and chlorination with $POCl_3$ gave **6** [10]. Reaction of **6** with sodium 2-hydroxyethoxide afforded the aminoalcohol **7** which was then cyclized under high dilution conditions with diacid chloride **2** to form the macrocyclic receptor **8** in 20% yield.

Once again nucleotide base binding was conveniently followed using $^1H$ NMR. Addition of one equivalent of 2′,3′,5′-tri-*O*-pentanoylguanosine **9** to a $CDCl_3$ solution of **8** causes downfield shifts in the NH resonance of **8** and the NH and $NH_2$ resonances of **9** (1.36, 0.31 and 0.25 ppm, respectively), consistent with their forming a triple hydrogen bond complex (as in Figure 9). In addition, all of the naphthalene proton resonances are shifted upfield due to the close approach of the naphthalene to the bound guanine. These shifts are greater for the naphthalene-5, -6

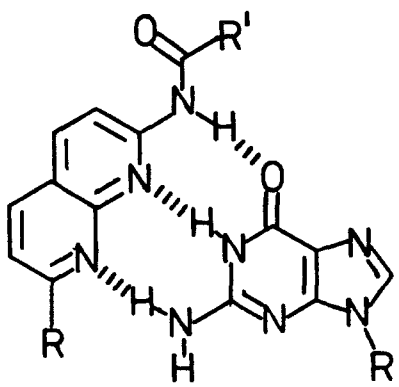


Fig. 9. **9**: R = 2′, 3′, 5′-tri-*O*-pentanoylribose. **10**: R = OEt, R′ = $CH_3$.

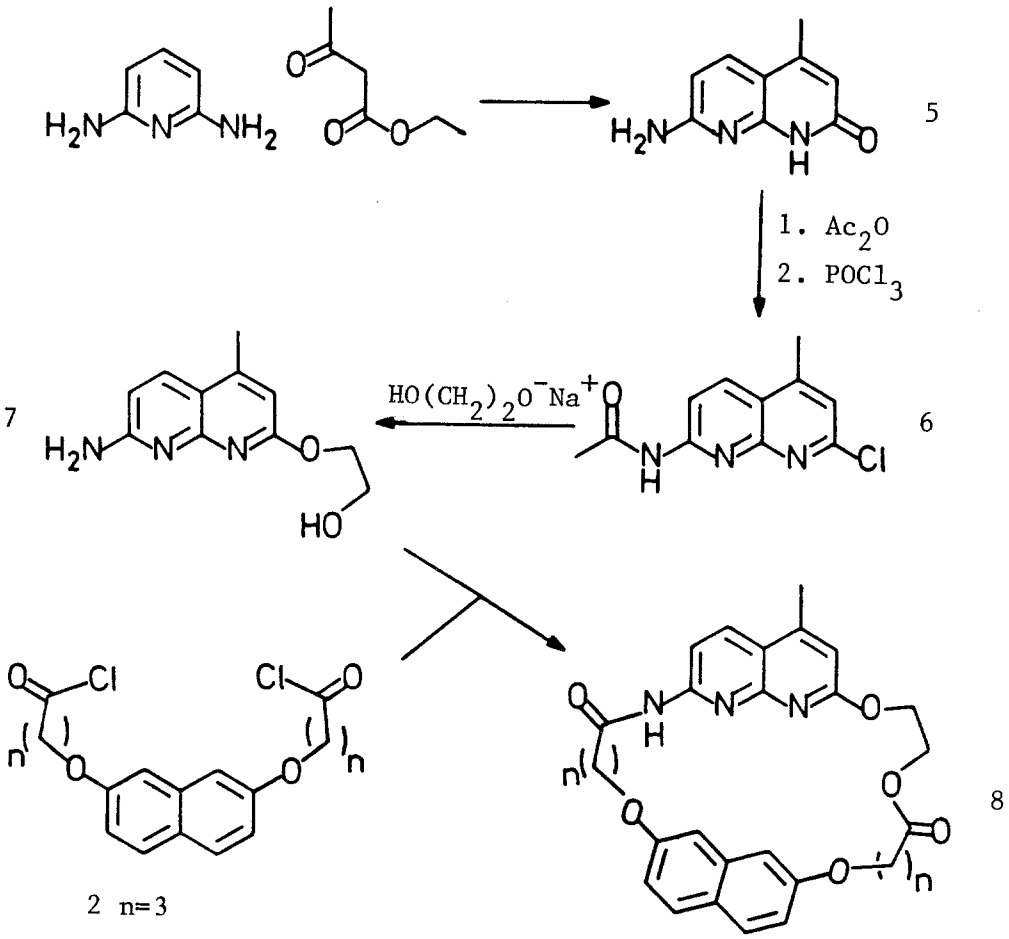
5
1. $Ac_2O$
2. $POCl_3$
6
$HO(CH_2)_2O^-Na^+$
7
2 n=3
8

Fig. 10.

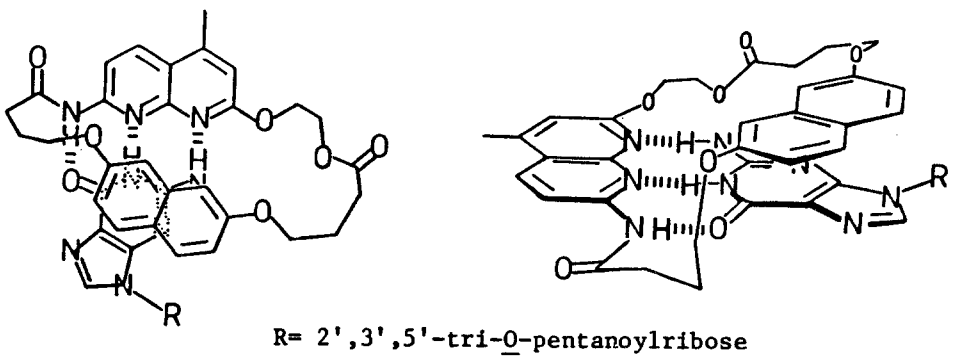
R= 2',3',5'-tri-O-pentanoylribose

Fig. 11.

and -8 protons (0.15, 0.14 and 0.18 ppm) than for those at the -4,3 and -1 positions (0.1, 0.09, and 0.11 ppm), suggesting a skewed orientation of the naphthalene relative to the guanine (Figure 11). In this structure for the complex **8 : 9** (Figure 11) the -5, -6 and -8 protons on the naphthalene would lie closer to the purine and naphthyridine ring currents and would, thus, experience greater upfield shifts. As in the thymine case, the stacking interaction causes a substantial increase in the association constant between **9** and naphthyridine macrocycle **8** ($K_s = 530\ M^{-1}$) compared to acyclic naphthyridine **10** ($K_s = 126\ M^{-1}$). The contribution of guanine–naphthalene stacking to the overall binding energy between **8 : 9** ($\sim 4$ kcal/mol) is approximately 1 kcal $mol^{-1}$, in good agreement with the thymine case.

## 4. Barbiturate Recognition

Our success in incorporating hydrogen-bonding groups into macrocyclic receptors prompted us to extend the approach to other biologically significant substrates. This and other work [8, 11, 12] suggested that incorporation of several inwardly-facing hydrogen bonding groups into a cleft or cavity would lead to strong binding to substrates with complementary shape and hydrogen bonding characteristics. We focused our attention on the barbiturate family of drugs (Figure 12) which are widely used as sedatives [13] and as anticonvulsants [14]. Molecular modelling studies suggested that incorporation of two 2,6-diaminopyridine units into a macrocyle would allow the complexation of all six accessible hydrogen-bonding groups (four CO lone pairs and two NHs) in 5,5-disubstituted barbiturates (Figure 13). Appropriate choice of spacer Y may further allow a secondary recognition of the substituents in the 5,5-positions. For example, a diphenylmethane unit (as Y) should accommodate the 5,5-ethyl groups of barbital **12**.

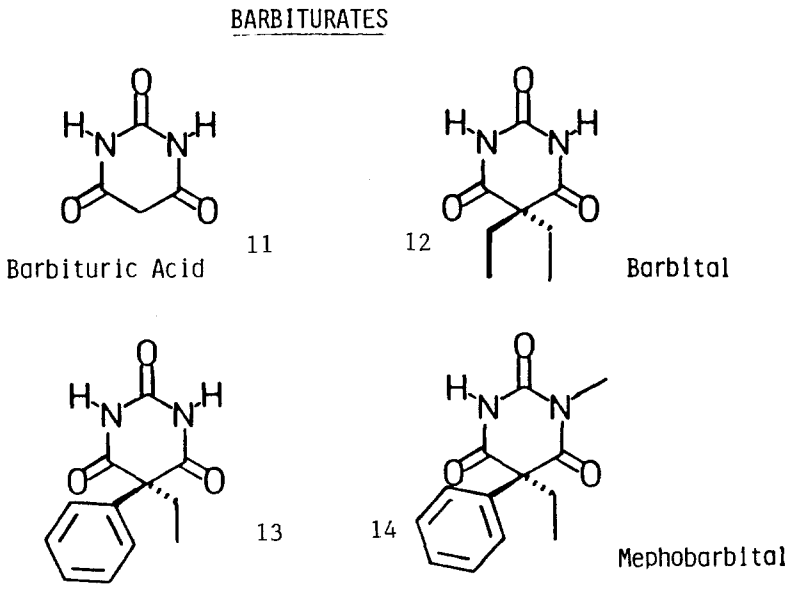


Fig. 12.

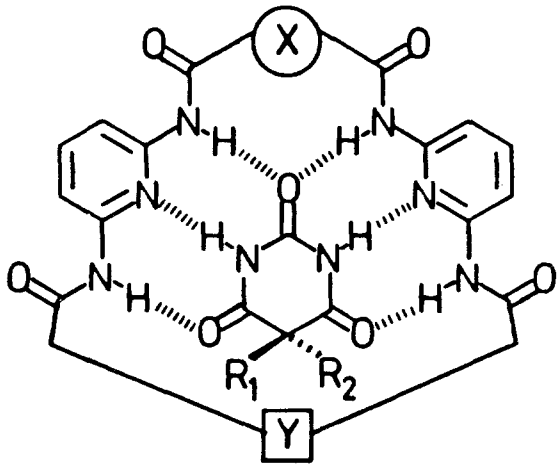


Fig. 13.

The first barbiturate receptors **18, 19, 20,** were prepared by standard high dilution methods from diamine **15** and acid chlorides **16** and **17** [15] (Figure 14). Their open conformation was confirmed by an X-ray structure of **19** which shows a tetrahydrofuran molecule occupying the central cavity (Figure 15). There is a high degree of preorganization in this structure which requires very small conformational changes to achieve the proposed hexahydrogen bonded complex. Complex formation was followed by $^1$H NMR. Addition of one equivalent of barbital **12** to a $CDCl_3$ solution of **18** caused large downfield shifts of the amide resonances of **18** (1.65 and 1.63 ppm) and the imide resonance of **12** (4.38 ppm) indicating the formation of a

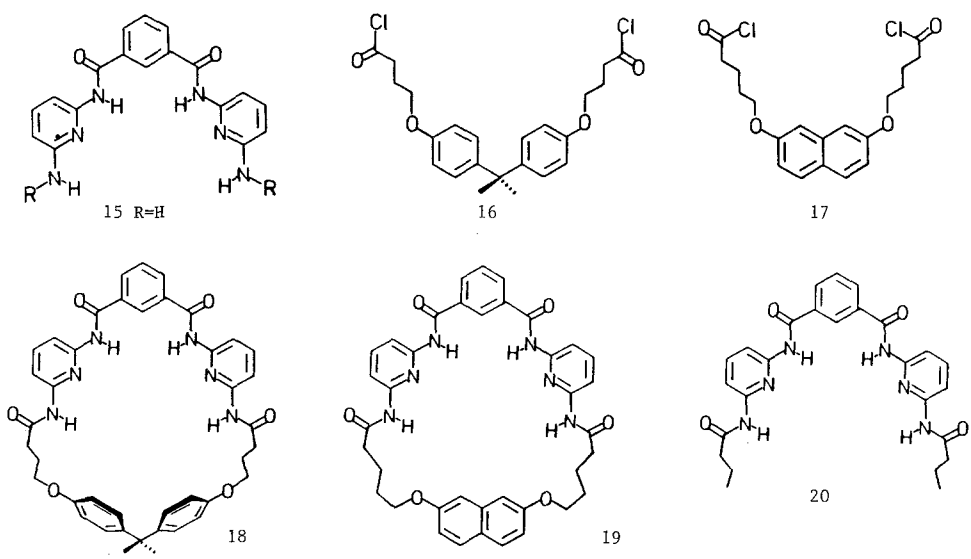


Fig. 14.

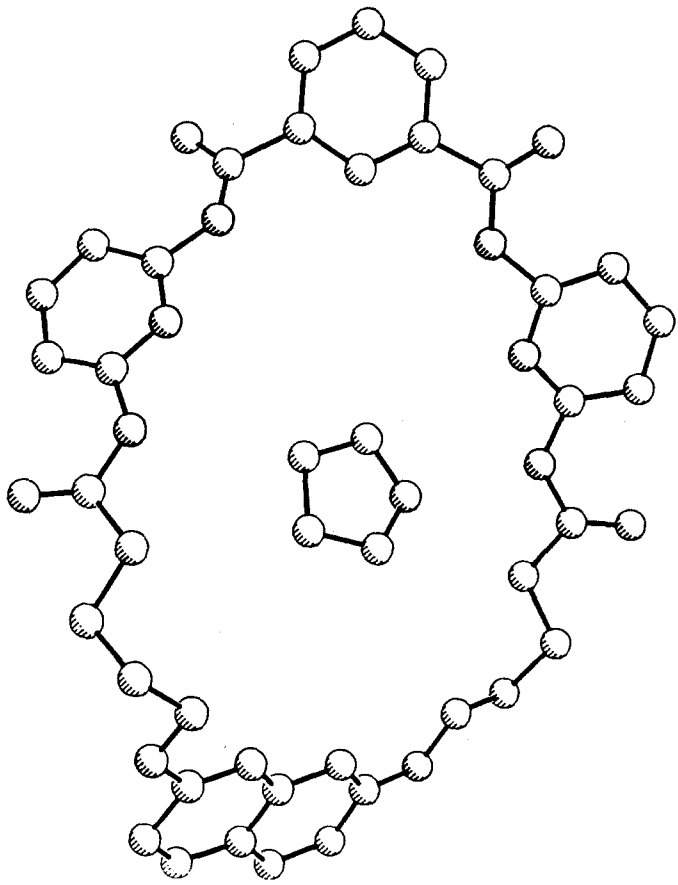

Fig. 15.

hexahydrogen bonded complex as shown in Figure 16. Also, the $CH_2$ and $CH_3$ resonances of the barbital ethyl groups were shifted upfield (0.25 and 0.23 ppm) confirming their proximity to the diphenylmethane cleft in Figure 16. A large downfield shift (0.4 ppm) is seen in the isophthaloyl-2-proton resonance and CPK models suggest that in the complex this proton lies in the deshielding region of the barbital-2-carbonyl group. In addition the isophthaloyl resonances in uncomplexed **18** are broad due to the conformational mobility of the macrocycle. In the complex (Figure 16) the motion of the isophthaloyl group is restricted and its $^1H$ resonances sharpen. Essentially similar NMR changes are seen with the other receptors, **19** and **20**, in their interactions with non-*N*-alkylated barbiturates.

Association constants were determined from $^1H$ NMR titration data using either Foster-Fife [7] or nonlinear least squares analysis and are collected in Table I. The strongest complex ($K_s = 1.37 \times 10^6\ M^{-1}$) is formed between barbital **12** and diphenylmethane receptor **18**. This result is expected due to the strong complementarity in both shape and hydrogen bonding specificity that exists between **12** and **18** (Figure 16). Alkylation of one barbiturate N-atom (as in mephobarbital **14**) essentially removes three H-bonding groups from participating in complexation and

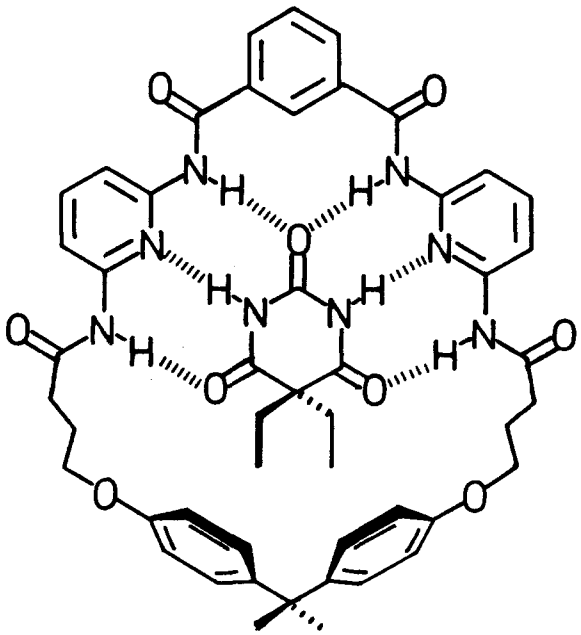

Fig. 16.

results in a 1000 fold decrease in binding to **18**. When the inwardly-pointing binding site is no longer enforced by a macrocyclic superstructure (as in acyclic **20**) association to barbital **12** diminishes by almost 100-fold. Finally, incorporation into the barbiturate-5 position of a bulky substituent that cannot fit neatly into the receptor cavity (as with phenobarbital **13** and **18**) causes a 10-fold reduction in the binding constant. A degree of secondary recognition of the 5, 5-substituents in the barbiturate can also be seen. While **18** binds barbital **12** more strongly than phenobarbital **13**, receptor **19**, which contains a flat naphthalene moiety, binds **13** nearly twice as strongly as **12**.

In summary, we have shown that complementary positioning of recognition sites (particularly hydrogen bonding, stacking and hydrophobic groups) into a macrocyclic structure can lead to very strong and specific complexation of uncharged organic molecules.

Table I. Association constants for the receptor–barbiturate interaction.

| Receptor | Barbiturate | $K_s$, $M^{-1}$ (25°C, $CDCl_3$) |
|---|---|---|
| **20** | Barbital (**12**) | $2.08 \times 10^4$ |
| **18** | Mephobarbital (**14**) | $6.80 \times 10^2$ |
| **19** | Phenobarbital (**13**) | $2.80 \times 10^5$ |
| **18** | Phenobarbital (**13**) | $1.97 \times 10^5$ |
| **19** | Barbital (**12**) | $1.35 \times 10^5$ |
| **18** | Barbital (**12**) | $1.37 \times 10^6$ |

## Acknowledgements

We thank KOSEF, Korea for a fellowship to S.K.C. and the National Institutes of Health for financial support of this work.

## References

1. W. Saenger: in *Principles of Nucleic Acid Structure*, Springer-Verlag, New York, 1984, p. 105.
2. U. Heinemann and W. Saenger: *Nature* **299**, 27 (1982).
3. A. Fersht: in *Enzyme Structure and Function*, Freeman, Reading, 1977, p. 262.
4. A. D. Hamilton and D. Van Engen: *J. Am. Chem. Soc.* 1987, **109**, 5035.
5. B. Feibush, A. Figueroa, R. Charles, K. D. Onan, P. Feibush, and B. Kargar: *J. Am. Chem. Soc.* **108**, 3310 (1986).
6. K. Edo, M. Mizugaki, Y. Koide, H. Seto, K. Furihata, N. Otake, and N. Ishida: *Tetrahedron Lett.* **26**, 331 (1985).
7. R. Foster and C. A. Fife: *Prog. Nucl. Magn. Reson. Spectrosc.* **4**, 1 (1969).
8. J. Rebek Jr., B. Askew, P. Ballester, C. Buhr, S. Jones, D. Nemeth, and K. Williams: *J. Am. Chem. Soc.* **109**, 5033 (1987).
9. A. D. Hamilton and N. Pant: submitted for publication.
10. E. V. Brown: *J. Org. Chem.* **30**, 1607 (1965).
11. J. Rebek Jr.: *Science* **235**, 1478 (1987).
12. M. Newcomb, S. S. Moore, and D. J. Cram: *J. Am. Chem. Soc.* **99**, 6405 (1977); V. M. L. J. Aarts, C. J. V. Staveren, P. D. J. Grootenhuis, J. V. Eerden, L. Kruise, S. Harkema, and D. N. Reinhoudt: *J. Am. Chem. Soc.* **108**, 5035 (1986); R. E. Sheridan and H. W. Whitlock Jr.: *J. Am. Chem. Soc.* **108**, 7120 (1986).
13. J. A. Vida: in *Burger's Medicinal Chemistry*, Wolff, M. E., Ed., Wiley-Interscience, New York, 787 (1981).
14. E. I. Isaacson and J. N. Delgado: in *Burger's Medicinal Chemistry*, Wolff, M. E. Ed., Wiley-Interscience, New York, 829 (1981).
15. S. K. Chang and A. D. Hamilton: *J. Am. Chem. Soc.*, in press.