Pure & Appl. Chem., Vol. 60, No. 4, pp. 533–538, 1988.
Printed in Great Britain.
© 1988 IUPAC

Artificial receptors for biologically active molecules

Andrew D. Hamilton*, Nalin Pant and Alexander Muehldorf

Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA.

Abstract - Artificial receptors for two key classes of biological molecules, nucleotides and peptides, have been prepared. The nucleotide receptors mimick nucleotide binding enzymes by employing a multi-point binding strategy within a flexible cavity. Complementary hydrogen bonding and hydrophobic groups have been linked within a macrocyclic ring to provide a series of thymine receptors. Binding has been studied by NMR and X-ray crystallographic methods which show that both components interact with the bound substrate. A series of potential peptide receptors based on the antibiotic, vancomycin, have also been prepared and their structural and functional similarities to the natural product have been studied.

INTRODUCTION

A key feature in any process of molecular recognition is complementarity between receptor and substrate. A specific receptor must complement its substrate in terms of both shape (providing a cavity of corresponding size and form) and chemical (lining the cavity with groups capable of interacting with regions on the substrate) characteristics. A rigid substrate can then react with a rigid and complementary receptor, with a preorganized binding site, in a fashion resembling a lock and key. In most biological systems, however, the enzyme receptor site is flexible and conformational changes occur to organize the binding site on approach of the substrate. This latter, induced fit, mechanism of recognition has the potential of linking the transmission of a signal to substrate binding (ref. 1). In a program aimed at the development of artificial receptors capable of molecular recognition we are attempting to understand these ideas of complementarity and structural flexibility and to incorporate them into synthetic systems. We have chosen biologically significant molecules, nucleotides and peptides, as our target substrates in the hope that during the course of our investigation of the structural basis of their molecular recognition we may uncover pharmaceutically interesting substances.

1. NUCLEOTIDE RECOGNITION

General strategies

In recent years there has been enormous interest in the design of synthetic molecules that recognize and bind to nucleic acids in a sequence-specific manner (ref. 2). Our initial goal in this area is to develop specific receptors for each of the key nucleotide bases and then to oligomerize them into units for strong and sequence-specific binding to single-stranded nucleic acids. Later work will focus on double-stranded nucleic acid recognition.

In designing receptors for nucleotides much can be learned from the structures of nucleotide binding enzymes. Ribonuclease T_1 cleaves RNA specifically at guanosine. The crystal structure of an enzyme-inhibitor complex (figure 1) shows that the origin of this strong specificity is a three point interaction between protein groups and the nucleotide. Two

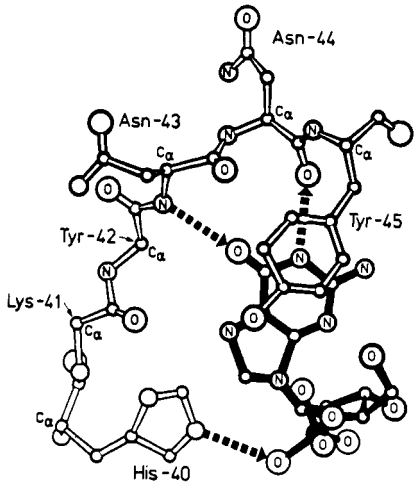


Fig. 1 Guanine (bold lines) binding site of Ribonuclease T_1 (ref. 3).

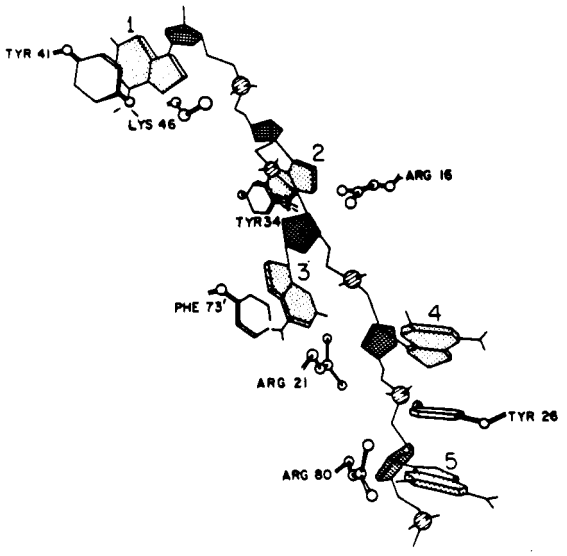


Fig. 2. Single stranded DNA binding site of gene 5 protein of bacteriophage fd (ref. 5)

highly specific hydrogen bonds are formed between the peptide backbone and the O(6) and NH(1) of guanine. There is a hydrogen bond between His-40 and the phosphate group and a hydrophobic stacking interaction between Tyr-45 and the guanine plane. The electron density map shows two positions for Tyr-45. The first (as shown) in which the tyrosine stacks at van de Waals distance (3·4 Å) from the guanine and the second where it is some distance from the binding site. This suggests that on substrate binding the tyrosine swings into a stacking position, exemplifying an induced fit mechanism of recognition (ref. 3). In contrast the gene 5 protein of bacteriophage fd, a protein involved in unwinding and stabilizing single stranded DNA, shows no sequence specificity in forming a strong complex ($K_S \approx 10^9 M^{-1}$) with a five base fragment of SS DNA (ref. 4). An X-ray structure of the protein-DNA complex (fig. 2) shows two types of intermolecular interactions stabilizing the complex (ref. 5). Four positively charged groups on the protein (Arg 16, 21, 80 and Lys 46) form strong electrostatic interactions with the DNA phosphates while four aromatic residues (Tyr 26, 34, 41 and Phe 73) hydrophobically stack with each of the 5 nucleotide bases (Tyr 26 binding to bases 4 and 5 in figure 2). The lack of any base or sequence selectivity is due to the absence of hydrogen bonds between the protein and the purine or pyrimidine bases.

Thus, a strong, selective receptor for nucleotides should employ a multi-point recognition strategy involving a characteristic H-bonding region, a planar hydrophobic region and, potentially, cationic groups (e.g. fig. 3 for thymine recognition). The schematic design for such a receptor is shown in figure 4 in which H-bonding and hydrophobic groups are incorporated into a macrocyclic ring. Suitable substituents (X in fig. 4) can both modify the basicity of the H-bonding groups and act as a link for oligomerization to future oligonucleotide receptors.

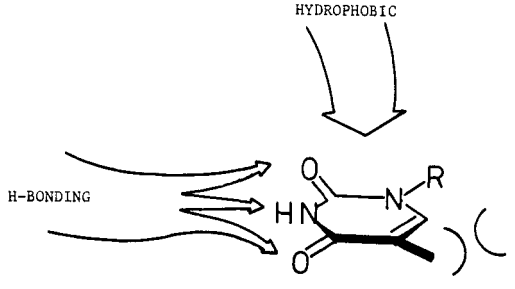


Fig. 3. Main elements for the recognition of nucleotides.

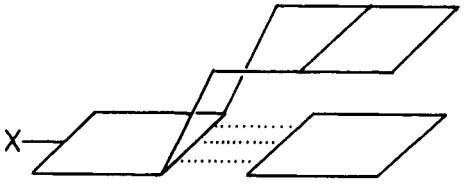


Fig. 4. Schematic design for a multipoint binding nucleotide receptor.

Thymine receptors: Synthesis, structure and binding

The design of a thymine receptor is based on the triple hydrogen bond complementarity between 2,6-diamidopyridines and cyclic imides (ref. 6). The hydrophobic group is derived from naphthalene, a component in several intercalating drugs (ref. 7). The two groups were linked into a macrocycle via the route shown in figure 5. Reaction of 2,7-dihydroxynaphthalene with an ethyl bromoalkanoate gave diester (1) which was readily converted into its diacid chloride and reacted, under high dilution conditions, with 2,6-diaminopyridine to provide the corresponding macrocycle (3). The yield of the final macrocyclization step varied from 20-26% depending on the size of the ring. The structure of (3a) was confirmed by single crystal X-ray analysis which shows an open conformation with the naphthalene poised away from the pyridine ring at an interplane angle of 127.5°.The amide hydrogens project underneath the naphthalene ring providing a partially organized substrate binding region (fig. 6).

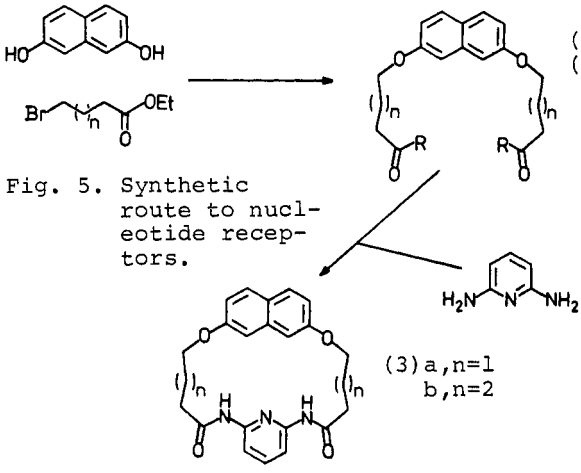


Fig. 5. Synthetic route to nucleotide receptors.

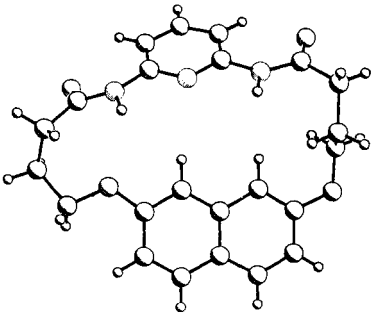


Fig. 6. X-ray structure of (3a).

Treatment of a $CDCl_3$ solution of (3a) with one equivalent of 1-butylthymine (4) results in several characteristic changes in the ^1H NMR spectrum. The NH protons on both (3a) and (4) are shifted downfield by 2.25 and 2.6 ppm, respectively, reflecting the formation of a triple hydrogen bonded complex (ref. 6). However upfield shifts are seen in the thymine-6-proton, -ring methyl and -N-methylene resonances while no significant shift is found for the alkyl methyl group (fig. 7). The

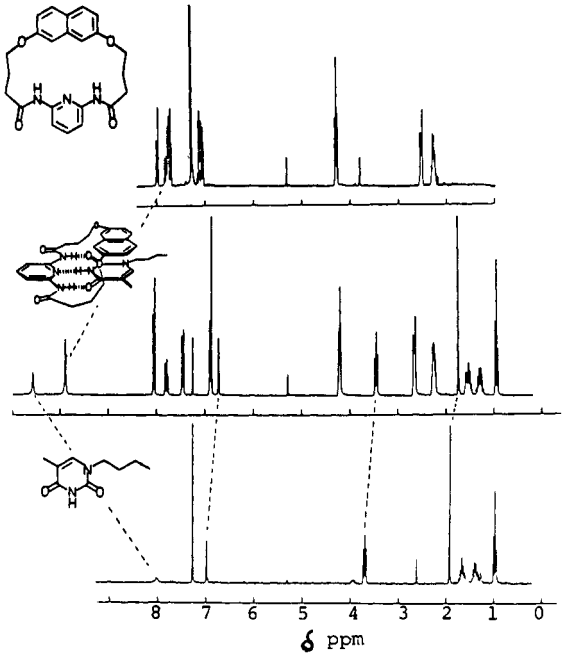


Fig. 7. ^1H NMR Spectra of (3a), (3a):(4), (4) in $CDCl_3$.

selective upfield shifts of certain substrate protons are consistent with the close approach of the naphthalene to the substrate and its participation in binding (ref. 8). These results are in contrast to those with 2,6-diacetamidopyridine which shows downfield shifts of its NH protons on hydrogen bonding to (4) but exhibits no upfield shifts in the substrate ring protons (ref. 6).

The structure of the complex (3a):(4) was confirmed by X-ray crystallography which shows (fig. 8) three hydrogen bonds between the pyridine and thymine rings at distances of (N...N) 3.06, (N...O) 2.87, 2.99 Å. The naphthalene lies directly above the substrate with an angle of 14° between the rings and a closest inter-plane contact of 3.37 Å. The position of the naphthalene accounts for the upfield shifts of the ring protons on the substrate (fig. 7) and indicates a strong similarity between solution and solid state structures. The structures of free and complexed receptor (3a) are seen from the side view, in figure 9. This clearly shows the 'molecular hinge' motion of the naphthalene ring which, on substrate complexation, swings through an arc of 34.1° to within van de Waals distance of the thymine ring. This 'induced fit'-like behavior mimics the recognition of nucleotides by ribonuclease T_1 in which a tyrosine moves into place above the bound guanine (ref. 3).

In order to increase the potential interaction between naphthalene and thymine rings we have recently prepared the larger macrocycle (3b). This would be expected to position the naphthalene further forward, more parallel and closer to the substrate. Preliminary NMR binding data indicate that this is the case. The upfield shifts of the thymine ring protons are consistently larger for (3b) than those of the smaller macrocycle (3a), reflecting more significant ring current effects in (3b)(fig. 10).

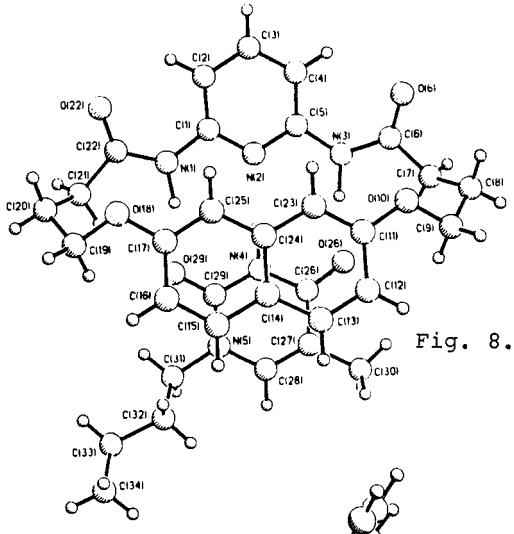
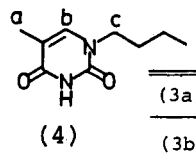


Fig. 8. X-Ray structure of (3a):(4).



	Δδ ppm		
	a	b	c
(3a)	0.19	0.29	0.24
(3b)	0.23	0.37	0.35

Fig. 10. Upfield shifts of (4) ring protons.

Fig. 9. X-ray structure of (3a) and (3a):(4) viewed from the side.

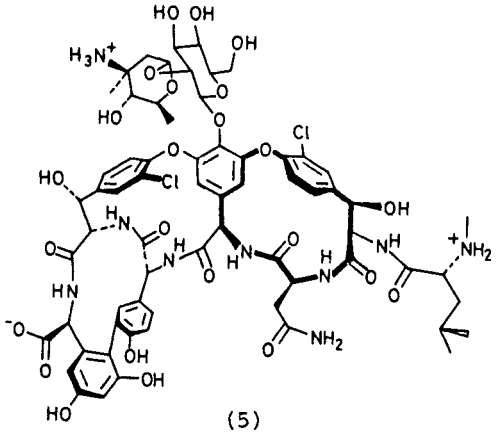
This approach to the design of nucleotide receptors holds considerable potential. Many variations on the basic theme can be made to provide receptors with different binding strengths and specifities. Changing the nature of the hydrogen bonding group (e.g. to 2-amido-1,8-naphthyridines for guanine binding) should lead to different recognition characteristics. Addition of cationic groups onto the hydrophobic region should enhance water solubility and also increase binding strength via potential electrostatic interaction with the phosphodiester backbone of polynucleotides. Oligomerization (e.g. through the 4-position of the pyridine in (3)) should lead to oligonucleotide receptors. Finally minor modifications to the design of the macrocycles may provide synthetic receptors that recognize the major or minor groove of double-stranded nucleic acids via a multi-point binding strategy.

2. PEPTIDE RECOGNITION

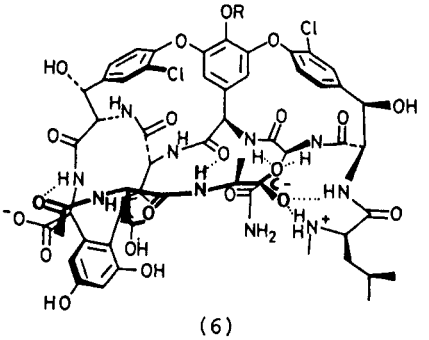
Peptide recognition presents a more complex problem than in the nucleotide case, due to the greater chemical variation and structural flexibility of the substrates. We are concentrating our efforts on the design of receptors for a single dipeptide unit, acylated-D-alanine-D-alanine. This dipeptide forms the carboxylate terminus of key mucopeptide precursors in bacterial cell wall formation.

General strategy

The vancomycin family of antibiotics, e.g. vancomycin (5), functions by recognizing and binding to this D-Ala-D-Ala carboxylate terminus and, thus, preventing the final step in bacterial cell wall biosynthesis (ref. 9). The structure of the active complex has been proposed to be (6) in which six hydrogen bonds are formed between antibiotic and dipeptide (ref. 10). The two peptide methyl groups occupy hydrophobic regions formed by the triphenyl diether and biphenyl groups leading to a strong and highly substrate- and stereospecific complex. Our approach is to reduce the complexity of (5) in order to determine the minimum structural unit required for peptide recognition (ref. 11). Interestingly, five of the six hydrogen bonds in (6) occur on or near the right hand ring (5, bold lines) which forms a carboxylate-binding pocket for the substrate (ref. 12). Synthetic analogs of this right hand ring might be expected to show similar properties to vancomycin and, so, become our first targets.



(5)



(6)

Synthesis and properties of vancomycin analogs

A number of simplifying modifications were made in the target structure. For example, dimethylglycine was chosen as the N-terminal residue as using a tertiary amine (vs secondary in (5)) simplifies the manipulation of protecting groups but does not appear (from CPK models) to hinder substrate binding. The synthetic route is outlined in figure 11. Dinitrotyrosine ethyl ester, protected as its tertiary butoxycarbonyl derivative, was tosylated and reacted with phenol (7) to provide diphenyl ether (8) in 86% yield. Acid cleavage of the BOC group followed by reaction of the resultant amine with dimethylglycine acid chloride formed

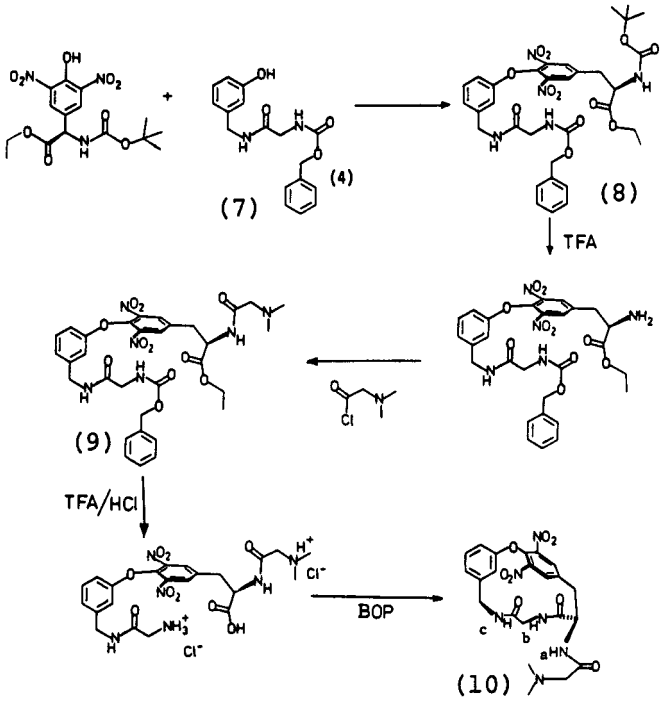


Fig. 11. Synthesis of vancomycin analog (10)

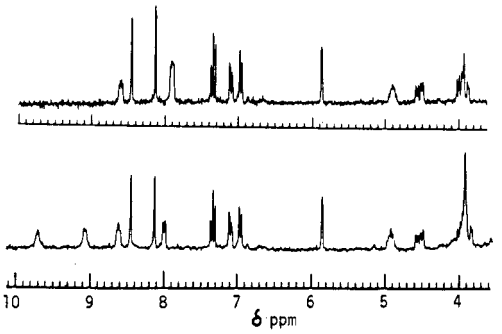


Fig. 12. ^{1}H NMR spectra of a) (10), b) (10)HCl in d_6 DMSO.

tripeptide (9) in 72% yield. Finally, cleavage of the protecting groups by acid hydrolysis followed by cyclization of the amino acid (using benzotriazol-1-yloxytris(dimethylamino)phosphonium hexafluorophosphate) gave cyclic diphenyl ether peptide (10) in 12% yield.

The ^{1}H NMR spectrum of (10) in d_6DMSO (fig. 12, low field region) shows an upfield shifted singlet at 5·80 ppm. This is due to the 2-H of the benzylamine ring which is constrained by the cyclic peptide to lie under the dinitrophenyl group. Similar upfield shifts of the equivalent proton are seen both in simpler model systems (ref. 12) and in vancomycin (ref. 10). The amide -H resonances have been assigned by decoupling and NOE experiments, amides a and b coming at 7.91 and c at 8.60 ppm. Formation of the hydrochloride salt of (10) results, in addition to other changes, in a downfield shift of amide a (figure 12) to 9.08 and an ammonium -H resonance at 9.71 ppm. This shift of 1·17 ppm is presumably due to H-bonding between the carbonyl oxygen of amide a and the ammonium proton. The interaction of (10).HCl with carboxylate substrates is presently under investigation.

REFERENCES

1. For general discussion see A.R. Fersht, Enzyme Structure and Mechanism, W.H. Freeman, New York (1985).
2. P.D. Dervan, Science, 232, 464 (1986).
3. U. Heinemann and W. Saenger, Nature, 299, 27 (1982).
4. J.E. Coleman and J.T. Oakley, C.R.C. Crit. Rev. Biochem., 7, 247 (1980).
5. G.D. Brayer and A. McPherson, Biochemistry, 23, 340 (1984).
6. B. Feibush, A. Figueroa, R. Charles, K.D. Onan, P. Feibush and B. Kargar, J. Am. Chem. Soc., 108, 3310 (1986).
7. J.W. Cown and C.C. Hanstock, J. Am. Chem. Soc. 104, 3213 (1982).
8. A.D. Hamilton and D. van Engen, submitted for publication.
9. D.H. Williams, V. Ragananda, M.P. Williamson, and G. Gojesen, Topics in Antibiotic Chemistry, 5B p 123, ed. P. Sammes, Ellis Harwood, Chichester (1980).
10. M.P. Williamson and D.H. Williams, J. Am. Chem. Soc., 103, 6580 (1981). C.M. Harris, H. Kopeka, and T.M. Harris, J. Am. Chem. Soc. 105, 6915 (1983).
11. M.J. Mann, N. Pant, and A.D. Hamilton, J. Chem. Soc. Chem. Commun., 158 (1986).
12. M.P. Williamson, D.H. Williams, and S.J. Hammond, Tetrahedron, 40, 569 (1984).