Journal of Inclusion Phenomena and Molecular Recognition in Chemistry 7 (1989), 19–26.

The Preparation of Ribose-Free 'Artificial Dinucleotides': A New Approach to Molecular Recognition?

JONATHAN L. SESSLER*, DARREN MAGDA, and JEFF HUGDAHL
Department of Chemistry, University of Texas, Austin, Texas 78712, U.S.A.

(Received: 1 February 1988)

Abstract. The synthesis of the cytosine-guanine linked dimer **1**, and its symmetric guanine-guanine analogue **2** are described. Compound **1** is the first homogeneous ribose-free system to be prepared and characterized in which molecular recognition by complementary guanine-cytosine base-pairing might be possible. It is thus prototypical of a new and potentially large class of 'artificial oligonucleotides'.

Key words. Artificial dinucleotide, molecular recognition, cytosine-guanine dimer.

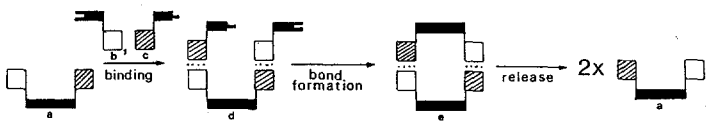
1. Introduction

Recently there has been tremendous interest in reproducing, in simple model compounds, the substrate specificity and catalytic efficiency of enzymatic systems and a number of complicated approaches involving crown ethers [1, 2], cryptands [3], calixarenes [4], spherands [5], clefts [6, 7], cyclophanes [8], and cyclodextrins [9, 10] are currently being pursued. Little effort, however, is being devoted to developing template-type catalysts inspired by DNA replication or ribosomal syntheses [11]. Nonetheless, complementary purine-pyrimidine base-pairing could offer an attractive new approach to molecular recognition [11–12] and might provide a unique means of controlling chemical reactivity. Our interest in this area derives from our long-standing fascination with the conceptual simplicity and chemical efficiency of DNA replication: we remain curious as to just which factors are responsible for self-replication at the molecular level. In particular, we have been intrigued by the question of whether it might be possible to construct artificial self-replicating systems from simple ribose and phosphate-free purine and pyrimidine precursors [13]. This is clearly a highly challenging goal; in this paper we describe the results of initial synthetic studies which might lead to its eventual realization.

A generalized representation of a possible template-based self-replicating system is shown in Scheme 1. Here □ and ▨ could represent complementary base-pairs and —— a rigid spacer group. From the point of strategic planning, several steps can thus be identified as crucial to the success of the overall catalytic sequence. Firstly, it is necessary to prepare covalently linked base-pairs, such as **a**, which are joined in such a way that they could serve as the initial dimeric template. These must be capable of binding by base-pairing the appropriate complementary monomeric substrates (e.g. **b** and **c**), without interference from competitive 'narcissistic' interactions involving just the original template **a**. Secondly, bond formation between the

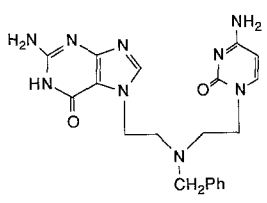
* Author for correspondence.

Scheme 1

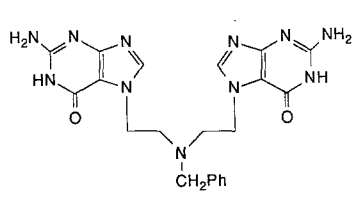


cobound substrates in **d** must be effected (to give **e**). For optimum results, this bond formation step should involve a reaction which is intrinsically exothermic but which would be slow in the absence of the neighboring group effect imposed by the template. Finally, once bond formation is complete, the newly formed dimer must be released from the original template (to give two equivalents of **a**). Clearly, if a *catalytic* self-replicating system is to be constructed, product release (from **e**) must occur at a rate such that a new generation of monomeric substrates may bind to the templates and thus enter into the cycle. For noncatalytic systems, these stringent kinetic and/or thermodynamic restrictions would be relaxed. In this case, changes in solvent or other conditions could be used to 'pry' the newly formed dimer from the original template.

The postulated self-replicating sequence outlined above is obviously predicated on a number of assumptions. The first, and most fundamental, of these is that base-pairing interactions may be used as controlling elements in simple biomimetic systems. Clearly, until this basic point is established, it makes little sense to discuss subsequent issues such as whether the thermodynamics and kinetics of bond formation and substrate release can be controlled as desired. Unfortunately, this first basic premise has yet to be established: although several generalized biomimetic hydrogen bonding model systems have been reported recently [12, 14, 15], only one involved possible base-pairing interactions [12]. In fact, at present, we are not aware of any covalently linked purine and pyrimidine heterodimers which might be capable of carrying out selective Watson-Crick base-pairing (for various nonbinding dimeric systems see refs. 16–19). Indeed, it is not clear at present that such systems can be made and, if prepared, whether they will be capable of effecting molecular recognition by base-pairing. As our immediate objective, we therefore sought to prepare simple ribose and phosphate-free 'artificial dinucleotides' in which complementary guanine-cytosine base-pairing might be possible. In this paper we report our initial results in this area; we describe the synthesis of the cytosine-guanine dimer **1** and its symmetric guanine-guanine analogue **2**.



1



2

2. Design Considerations

In approaching the problem of developing simple acyclic 'artificial dinucleotides' we were confronted with the choice of which particular base-pair to work with: adenine/thymine or guanine/cytosine. We selected the latter pair, reasoning that the three hydrogen bonds involved in their pairing would lead to stronger interactions in an artificially constructed binding site (Figure 1). The equilibrium constant for guanine/cytosine base-pairing is $3.7\,M^{-1}$ at 32° in DMSO [20]. Under these conditions hydrogen bonding interactions are considered to be dominant. (This should be contrasted to the situation in water where stacking predominates and base-pairing interactions are weak [21].) Since nucleoside bases (especially guanine) are notoriously insoluble in most organic solvents, we could anticipate DMSO to be the solvent of choice of initial enzyme models.

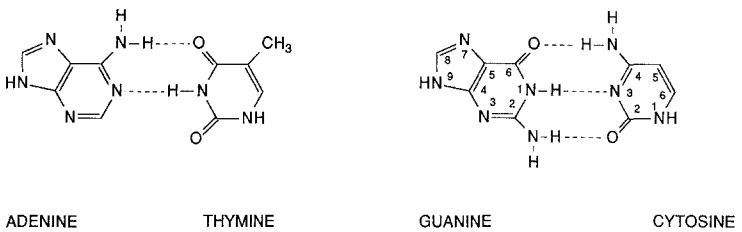


Fig. 1. Watson-Crick base-pairing interactions for common nucleic acid heterocycles.

In a linked dimeric molecule where multiple base-pairs are free to form, the strength of intermolecular interactions should be larger than in a simple cytosine/guanine pair. This assumption is valid provided that *intramolecular* base-pairing does not occur. The system must therefore be designed in such a way as to preclude these interactions. An examination of CPK models suggests that steric constraints would prevent undesirable internal hydrogen bonding in covalently linked guanine-cytosine dimers, such as **1**, in which the two bases are spanned by a five atom tether. Of course, in order to retain the ability to carry out *intermolecular* base-pairing, both the tether itself and its site of attachment to the bases must be chosen so that Watson-Crick hydrogen bonding can still occur. The obvious points of attachment are thus N^1 for cytosine and either N^9, C^8, or N^7 for guanine. Although N^9 substitution is found in DNA, for a variety of synthetic considerations, we chose to prepare N^7-substituted guanine-containing systems. CPK models suggest that no difficulties should be engendered by this choice: N^7-substituted guanines should be able to recognize N^1-substituted acyclic cytosines as complementary substrates.

3. Synthesis

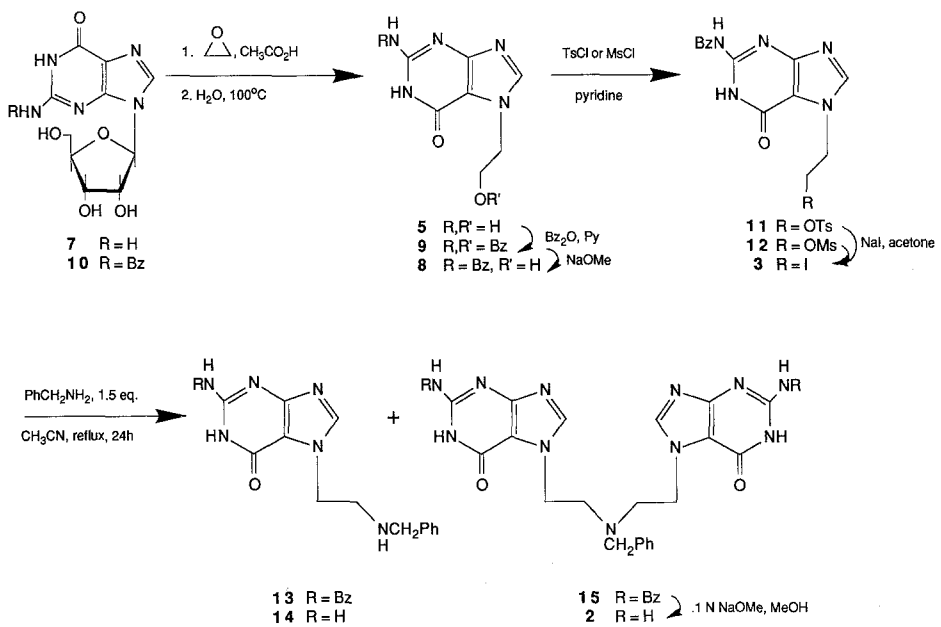
In order to facilitate entry into a wide variety of artificial oligonucleotide systems, electrophilic synthons **3** and **4** were developed. Both compounds have their exocyclic amino group protected as the benzamide. This not only protects this functionality, but also improves the solubility of the heterocycles in organic solvents. Moreover, because both incorporate a reactive electrophilic functionality,

we felt that they could easily be elaborated to give five atom bridging systems by reaction with a variety of monatomic nucleophiles (e.g. sulfide anion, amines, etc.).

The 2-hydroxyethyl substituted forms of both bases (**5** and **6**) were readily obtained following literature procedures [18, 22]. The synthesis of 7-(2-hydroxyethyl)guanine **5** called for the alkylation of guanosine (**7**) with ethylene oxide. We found this reaction especially attractive owing to the absence of the regioisomeric by-products such as those that result from the direct alkylation of guanine [23].

We found that the amino-protected guanine **8** could be synthesized by two methods (Scheme 2) [24]. One involved the selective deprotection of the dibenzoyl derivative **9** [25], obtained from **5** by treatment with benzoic anhydride in pyridine. Alternatively, N^2-benzoylguanosine (**10**) could be alkylated with ethylene oxide and subsequently depurinated by heating in water. The latter procedure was found to be more convenient.

Scheme 2

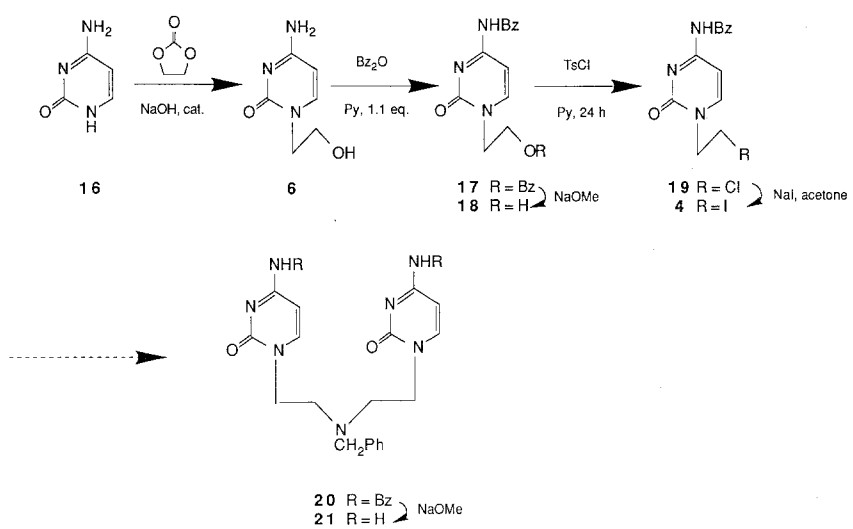


Once the protected species **8** was in hand, the next step involved activating the hydroxyl substituent. Our first attempt resulted in its conversion to the tosylate **11** by reaction with *p*-toluenesulfonyl chloride in pyridine. Compound **11** proved to be inadequately reactive for our purposes. The methanesulfonyl derivative **12** was therefore prepared from **8** using similar conditions. The mesylate was found to be more reactive than the tosylate, giving the desired reaction with benzylamine in acetonitrile, although somewhat slowly. Compounds **11** and **12** were therefore converted to the iodide analogue **3** using sodium iodide in refluxing acetone. Reaction of **3** with benzylamine produced three compounds **13–15**; the relative yields depended on the stoichiometry of the reaction. If a large excess of benzylamine is employed, both the secondary amine **13** and its deprotected analogue **14**

were formed. The latter presumably resulted from the nucleophilic attack by benzylamine at the benzamide carbonyl. Interestingly, the monomeric precursor is much less susceptible to debenzoylation under these conditions. When 1.5 eq. of benzylamine was used, the protected secondary amine **13** was obtained in ca. 50% yield as the major product. The protected dimer **15** was also isolated in 10% yield as a minor byproduct. The dimer was deprotected by treating with sodium methoxide/methanol under reflux, to produce **2** in nearly quantitative yield.

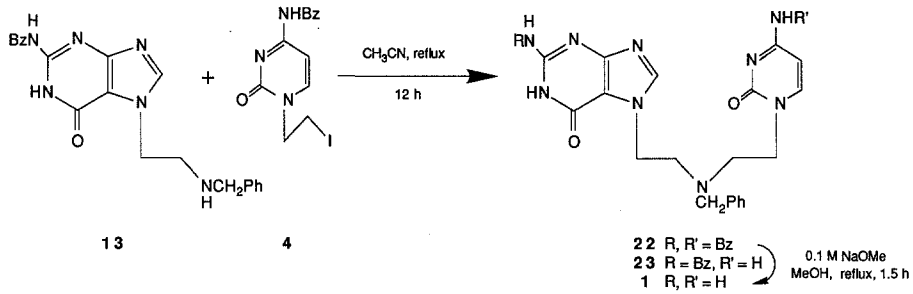
The cytosine analogues were produced in a manner similar to the guanine derivatives, with a few minor differences (Scheme 3). For instance, 1-(2-hydroxyethyl)cytosine **6** was produced by reaction of cytosine (**16**) with ethylene carbonate [22]. Also, the selective deprotection of dibenzoylated 1-(2-hydroxyethyl)cytosine **17** failed under a variety of conditions, contrary to our expectations [26]. Fortunately, the desired monoprotected compound **18** was easily separated from **17** after the reaction of benzoic anhydride and **6**. Finally, reaction of **18** with *p*-toluenesulfonyl chloride/pyridine at room temperature led to the chloride **19**, presumably via displacement by chloride of the intermediate tosylate. This suggests that the cytosine electrophiles are in general more reactive than the guanine analogues. Nevertheless, the chloride was converted to the iodide **4** after initial attempts at substitution failed. (It is conceivable that **4** could be reacted with benzylamine to produce a protected dimer **20**, which could be further elaborated to afford **21**, a *bis*-cytosine dimer analogous to **2**; this chemistry is being explored at the moment.)

Scheme 3



Compound **4** was found to react readily with the secondary amine **13** in refluxing acetonitrile to produce the fully protected, 'mixed dimer' **22** (Scheme 4). The monobenzoyl species **23** could be generated by deprotection of the cytosine moiety with sodium methoxide/methanol at room temperature, while one hour at reflux was required to achieve the fully deprotected 'dinucleotide' **1**.

Scheme 4



The ^{1}H NMR spectra of dimers **1** and **2** are clearly different (c.f. Table I). Not only are the expected changes due to differences in chemical structure observed, but also slight perturbations exist for the acidic protons of guanine. (The cytosine amino protons resonate in the same region of the spectrum as the phenyl protons (at ca. δ 7.00), and hence are not well resolved.) For instance, as compared to the homodimer **2**, the lactam and amino protons of **1** are shifted downfield by 36 and 18 Hz, respectively. Shifts of similar magnitude were reported for the purine N^1-H and C^2-NH_2 protons in a dimeric Watson-Crick type complex formed from cytidine and guanosine in DMSO [20]. This suggests that the heterodimer **1** may be undergoing intermolecular hydrogen bonding to form, by biomimetic molecular recognition, a supramolecular complex (Figure 2). We are currently exploring this intriguing possibility.

Table I. Selected ^{1}H NMR data for 'artificial dinucleotides' and synthetic precursors[a]

Compound	Purines				Pyrimidines		
	C8-H	C2-NH_2	C2-NHBz	N1-H	C5-H	C6-H	C4-NHBz
18					7.27	8.02	11.13
19					7.32	8.18	11.23
4					8.15	8.89	12.38
8	8.09		12.33	11.85			
11	8.14		12.21	11.98			
12	8.24		12.41	11.89			
3	8.23		12.40	11.88			
13	8.14		3.38	3.38			
14	7.89	6.13		b			
15	7.96		12.22	11.81			
2	7.73	6.05		10.62			
22	8.11		12.10	11.18	7.22	7.53	12.10
23	8.03		12.05	b	5.62	7.28	
1	7.78	6.10		10.72	5.57	7.20	

[a] DMSO-d_6; 360 MHz

[b] Peaks too broad to be assigned chemical shift values.

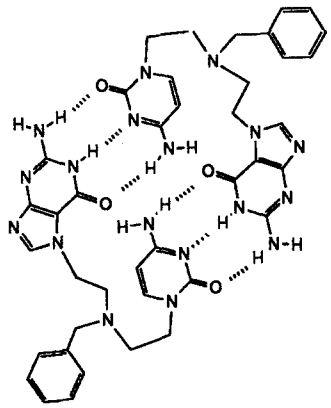


Fig. 2. Schematic representation of possible supramolecular complex which might be formed from the 'artificial dinucleotide' **1**.

4. Conclusions

We have described here the synthesis of the first rationally designed 'artificial dinucleotide', **1**, in which self-recognition by Watson-Crick hydrogen bonding might be possible. The synthetic strategy employed involved the preparation of several activated 'tailed' guanine and cytosine precursors and their subsequent reaction with benzylamine. It appears likely that this procedure will prove general and that it will be possible to couple these reactive synthons with a variety of other nucleophiles. In this way it should prove possible to prepare a variety of novel structures including other dimeric systems and unprecedented higher 'artificial oligonucleotides'. Studies with these should allow us to probe further important factors, such as hydrogen bonding and stacking, which influence base-pairing in both simple systems and naturally occurring oligonucleotides.

Acknowledgements

This work was supported by the National Science Foundation (PYI Award 1986), The Camille and Henry Dreyfus Foundation (New Faculty Award 1984), and the Procter and Gamble Co.

References and Notes

1. R. M. Izatt, J. S. Bradshaw, S. A. Nielsen, J. D. Lamb, J. J. Christensen, and D. Sen: *Chem. Rev.* **85**, 271 (1985).
2. F. Vögtle and H. Sieger: *Complexation of Uncharged Molecules and Anions by Crown-Type Host Molecules* (Host Guest Complex Chemistry-Macrocycles, Eds. F. Vögtle, E. Weber), pp. 319–373. Springer Verlag (1985).
3. (a) P. G. Potvin and J. M. Lehn: *Design of Cation and Anion Receptors, Catalysts and Carriers* (Progress in Macrocyclic Chemistry, v. 3, Eds. R. M. Izatt, J. J. Christensen), pp. 167–239. John Wiley and Sons (1987); (b) J. M. Lehn: *Science (Washington D.C.)* **227**, 849 (1985).
4. (a) C. D. Gutsche: *Calixarenes and the Art of Molecular Basket Making* (Progress in Macrocyclic Chemistry, v. 3, Eds. R. M. Izatt, J. J. Christensen), pp. 93–165. John Wiley and Sons (1987); (b) C. D. Gutsche: *Top. Curr. Chem.* **123**, 1 (1984).

5. (a) D. J. Cram: *CHEMTECH* **17**, 120 (1987); (b) D. J. Cram: *Acc. Chem. Res.* **98**, 1041 (1984).
6. (a) J. Rebek, Jr.: *Science (Washington D.C.)* **235**, 1478 (1987); (b) J. Rebek, Jr.: *Acc. Chem. Res.* **98**, 258 (1986).
7. C. S. Wilcox, L. M. Greer, and V. Lynch: *J. Am. Chem. Soc.* **109**, 1865 (1987).
8. Y. Murakami: *Top. Curr. Chem.* **115**, 107 (1983).
9. I. Tabushi: *Pure Appl. Chem.* **58**, 1529 (1986).
10. R. Breslow: *Acc. Chem Res.* **13**, 170 (1980).
11. Catalytic systems based on polymerized purines and pyrimidines have been reported: (a) T. Inoue and L. E. Orgel: *Science (Washington D.C.)* **219**, 859 (1983); R. Naylor and P. T. Gilham: *Biochemistry* **5**, 2722 (1966).
12. M. S. Kim and G. Gokel: *J. Chem. Soc., Chem. Commun.*, 1686 (1987).
13. Self-replicating systems based on hexanucleotides have, however, been reported: G. v. Kiedrowski: *Angew. Chem., Int. Ed. Engl.* **25**, 932 (1986).
14. (a) J. Rebek, Jr. and D. Nemeth: *J. Am. Chem. Soc.* **108**, 5637 (1986); (b) J. Rebek, Jr., B. Askew, P. Ballester, C. Buhr, S. Jones, D. Nemeth, and K. Williams: *J. Am. Chem. Soc.* **109**, 5033 (1987); (c) J. Rebek, Jr., B. Askew, P. Ballester, C. Buhr, A. Costero, S. Jones, D. Nemeth, and K. Williams: *J. Am. Chem. Soc.* **109**, 6866 (1987).
15. A. D. Hamilton and D. Van Engen: *J. Am. Chem. Soc.* **109**, 5035 (1987).
16. A. F. Maggio, M. Lucas, J. L. Barascut, A. Pompon, and J. L. Imbach: *Nouv. J. Chim.* **10**, 643 (1986).
17. P. Brookes and P. D. Lawley: *Biochem. J.* **77**, 478 (1960).
18. P. Brookes and P. D. Lawley: *J. Chem. Soc.* **1961**, 3923 (1961).
19. C.-W. Chen and H. W. Whitlock, Jr.: *J. Am. Chem. Soc.* **100**, 4921 (1978).
20. R. A. Newmark and C. R. Canter: *J. Am. Chem. Soc.* **90**, 5010 (1968).
21. (a) M. P. Schweizer, S. I. Chan, and P. O. P. Ts'o: *J. Am. Chem. Soc.* **87**, 5241 (1965); (b) A. D. Broom, M. P. Schweizer, and P. O. P. Ts'o: *J. Am. Chem. Soc.* **89**, 3612 (1967); (c) M. P. Schweizer, A. D. Broom, P. O. P. Ts'o, and D. P. Hollis: *J. Am. Chem. Soc.* **90**, 1042 (1968); (d) S. R. Jadkunas, C. R. Cantor, and I. Tinoco Jr.: *Biochemistry* **7**, 3164 (1968).
22. N. Ueda, K. Kondo, M. Kono, K. Takemoto, and M. Imoto: *Die Makromolekulare Chemie* **120**, 13 (1968).
23. K. K. Ogilvie, S. L. Beaucage, and M. F. Gillen: *Tetrahedron Lett.* **35**, 3203 (1978).
24. All new compounds gave spectral and C. I. mass spectrometric data in accord with the assigned structures.
25. J. L. Sessler, D. J. Magda, V. Lynch, D. I. Bernstein, and G. M. Schiff: *Nucleosides and Nucleotides*, in press.
26. H. Khorana, A. F. Turner, and J. P. Vizsolyi: *J. Am. Chem. Soc.* **83**, 686 (1961).