Tetrahedron Letters, Vol. 34, No. 8, pp. 1275-1278, 1993
Printed in Great Britain

0040-4039/93 $6.00 + .00
Pergamon Press Ltd

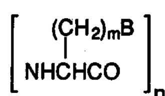
Peptide-Based Nucleic Acid Surrogates Incorporating Ser[CH2B]-Gly Subunits

Philip Garner* and Ji Uk Yoo

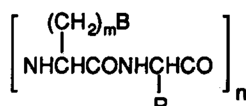
Department of Chemistry
Case Western Reserve University
Cleveland, Ohio 44106-7078

Abstract: *The synthesis of building blocks corresponding to four natural nucleobases (A, C, G, & T) is presented and it is demonstrated that such units may be linked together using standard peptide coupling techniques without racemization or β-elimination.*

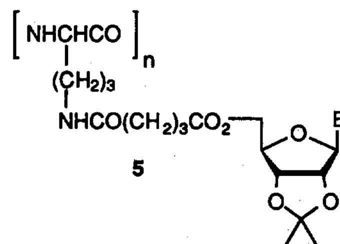
The quest to develop new drug therapies based on sequence-specific interactions between complementary nucleic acids is an exciting and rapidly growing field of chemical research.[1] Consequently, there is considerable interest in developing oligonucleotide surrogates that are capable of maintaining Watson-Crick (Hoogsteen) base-pairing to native RNA (DNA) targets but do not incorporate the usual phosphodiester linkages which are susceptible to nucleases and incompatable with passive membrane transport. An interesting approach to this problem involves the use of backbones made up of *peptide linkages* which connect the base-containing subunits. Besides their obvious resistance to nucleases, such peptide nucleic acid (PNA) surrogates would also be amenable to block or solid phase peptide synthesis techniques.



1 (m = 1)
3 (m = 4)

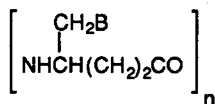


2 (m = 1)
4 (m = 4)

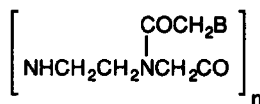


5

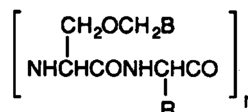
B = nucleobase



6



7



8

The concept of peptide-based nucleic acid surrogates is itself not new. Jones and coworkers prepared polymers incorporating repeat unit 1 (B = T, n ≈ 9-20) but found that they did not interact appreciably with polyadenylic acid (poly-A).[2] Shvachkin's group made a variety of well-defined homo- and heteronucleopeptides corresponding to 1 as well as "mixed" nucleopeptides incorporating subunit 2 and noted that the latter formed stable complexes with complementary duplexes.[3] De Konig and Pandit prepared nucleopeptides corresponding to 3 and 4 but found no interaction between poly-3 and poly-A.[4] Polymers made up of the more complex unit 5, on the other hand, did show binding to their complementary nucleic acid sequences.[5] More recently, Weller and coworkers reported the synthesis of nylon-based surrogates 6[6] while Egholm et al demonstrated that oligomers made up of the achiral unit 7 form stable 2:1 complexes with complementary DNA.[7]

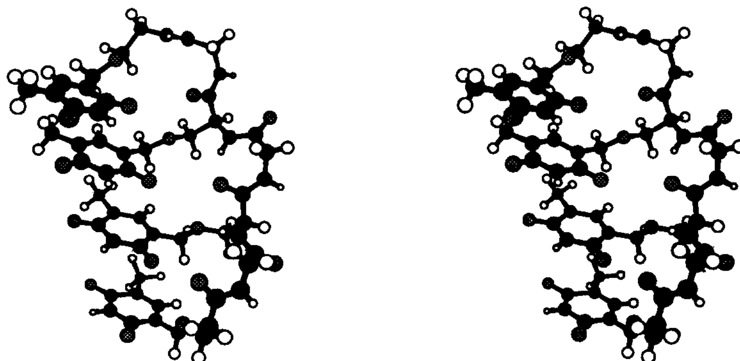


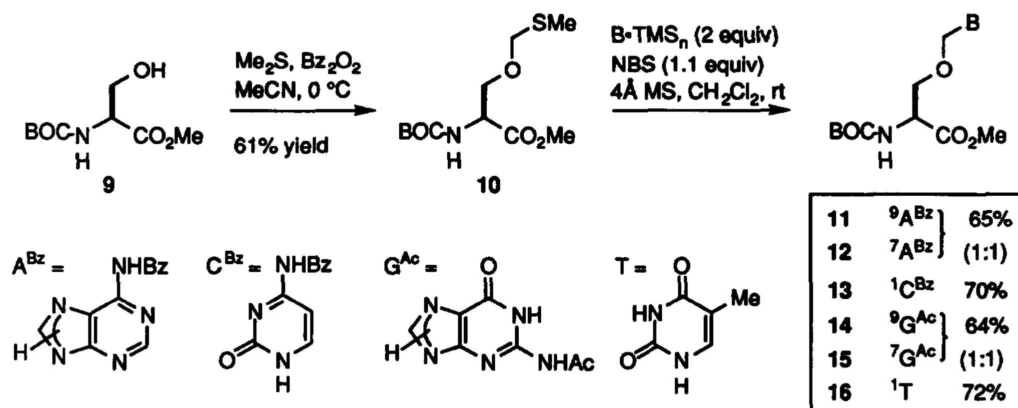
Figure 1. Stereoview of an energy-minimized molecular model of (Ser[CH_2T]-Gly)$_3$-Ser[CH_2T] with nucleobases in the B-DNA helix conformation. The terminal CO_2H and NH_2 groups (corresponding to the 5' & 3' ends respectively) have been deleted.

One of the drawbacks associated with systems 1 through 4 (as well as 6) is the need to synthesize the corresponding nucleoaminoacid building blocks in enantiomerically pure form. We felt that nucleic acid surrogate 8 might offer some advantages in this respect since it incorporates readily available α-amino acids. Here, the nucleobase is attached to a serine residue via a hemiaminal linkage which preserves the natural N-glycoside (O-C-N) substructure. This connector provides an additional H-bond acceptor and lateral flexibility which may lead to better binding to certain nucleic acid structures. There is also the potential for attaching chemical probes, etc. onto the spacer amino acid or N-methylation to prevent degradation by proteases. Molecular modeling studies by Weller suggest that peptides 2 incorporating glycine spacers appear to be well-suited for binding to complementary nucleic acids within the B-helix motif.[8] Our own preliminary modeling studies seem to indicate that repeat unit 8 may also accomodate the B-DNA helix conformation without undue steric strain (cf. Figure 1).[9]

Our synthesis of monomeric building blocks BOC-Ser[CH_2B]-OMe corresponding to all four natural nucleobases (B = A, C, G, & T) is shown in Scheme 1.[10] First, the known L-serine derivative 9 is converted to

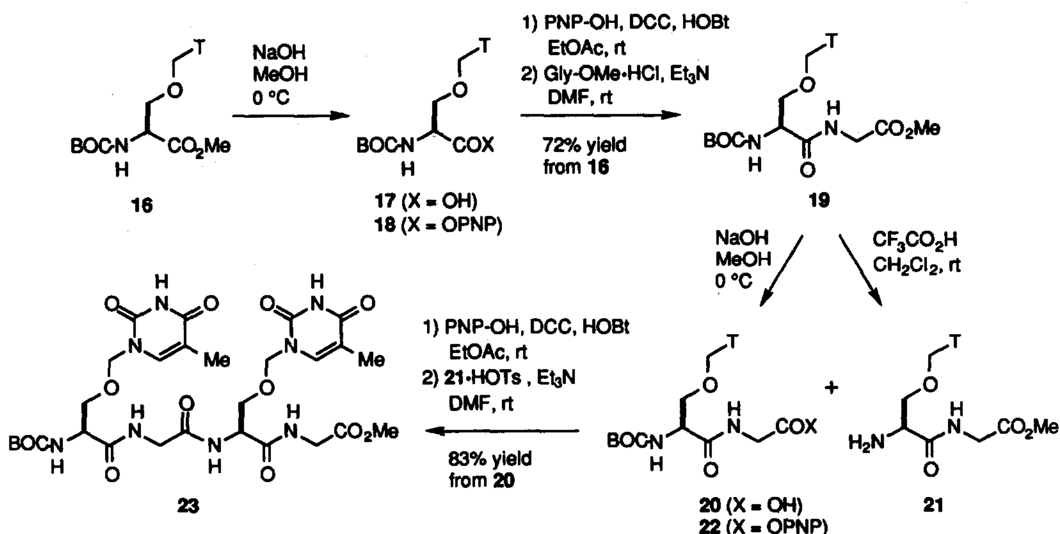
its methylthiomethyl (MTM) ether **10** in good yield using the method of Kyler.[11] This compound was coupled to silylated N^6-benzoyladenine (A^{Bz}•2TMS),[12] N^4-benzoylcytosine (C^{Bz}•2TMS),[8] N^2-acetylguanine (G^{Ac}•3TMS),[13] and thymine (T•2TMS)[14] after activation with NBS[15] to give the corresponding acyclic nucleosides **11-16**. The yields of purified nucleosides are in the range of 64-72% and, in the case of the purine bases, the N^7 & N^9 regioisomers were readily separated from each other by simple flash chromatography. For these N-acylated purines, regiochemical assignments were made (^1H & ^{13}C NMR) by comparing the relative chemical shifts of H-8, C-4, C-5, C-8, & C-1' (cf. reference 12). It is noteworthy that these conditions for both MTM ether formation and nucleosidation are mild enough to tolerate the presence of an acid-sensitive BOC group. Alternative methods for both MTM ether formation (DMSO + Ac_2O, rt)[16] and thioglycoside nucleosidation (B•TMS_n, NIS, TfOH, rt)[17] were found to be decidedly inferior with our substrates.

Scheme 1



That these novel aminoacid nucleosides could be incorporated into an oligopeptide structure of type **8** without racemization or β-elimination was demonstrated for the thymine series (Scheme 2). Peptide bond formation was best achieved using the p-nitrophenyl (PNP) active ester method.[18] Thus, saponification of **16** produced the carboxylic acid **17** which was converted to its PNP ester **18** and coupled directly with glycine methyl ester to give the dipeptide **19** in 72% overall yield after chromatographic purification. This compound served as the common building block for further peptide elongation as follows: Repetition of the saponification sequence with **19** led to the dipeptide carboxylic acid **20** whereas treatment of **19** with trifluoroacetic acid (TFA) produced the complementary dipeptide amine **21** which was isolated as its p-TsOH salt. Free acid **20** was then activated as the PNP ester **22** and coupled with the amine **21** to give the tetrapeptide **23** in 83% isolated yield after flash chromatography. The 400 MHz ^1H NMR spectrum of this compound showed only traces of a possible diastereomer suggesting that racemization had been minimal. The stage is now set for elongation of **23** into oligopeptide structures and biophysical evaluation of their interaction with nucleic acid targets.[19]

Scheme 2



REFERENCES AND NOTES

1. Select reviews: Nielsen, P. E. *Bioconjugate Chem.* **1991**, *2*, 1 ; Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 544; Hélène, C.; Toulme, J.-J. *Biochim. Biophys. Acta* **1990**, *1049*, 99; Dervan, P. B. *Science*, **1986**, *232*, 464.
2. Jones, A. S. *Int. J. Biol. Macromol.* **1979**, *1*, 194; Buttrey, J. D.; Jones A. S.; Walker, R. T. *Tetrahedron* **1975**, *31*, 73.
3. Cf. Tyaglov, B. V.; Permogorov, V. I.; Chernykh, N. A.; Semiletov, Yu. A.; Konde, K.; Shvachkin, Yu. P. *Zh. Obshch. Khim.* **1987**, *57*, 2124, and references cited therein.
4. De Koning, H.; Pandit, U. K. *Rec. Trav. Chim.* **1971**, *91*, 1069, and references cited therein.
5. Takemoto, K.; Inaki, Y. *Polym. Mat. Sci. Eng.* **1988**, *58*, 250.
6. Huang, S.-B.; Nelson, J. S.; Weller, D. D. *J. Org. Chem.* **1991**, *56*, 6007.
7. Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895.
8. Weller, D. D.; Daly, D. T.; Olsen, W. K.; Summerton, J. E. *J. Org. Chem.* **1991**, *56*, 6000.
9. Molecular modeling was done using the Biograf 3.1 software package. Conformational sampling was achieved via quenched dynamics followed by all-atom E-minimization. Relative strain was deduced by comparing $E_{helix} - E_{free}$ for both B-PNA and B-DNA models.
10. All new compounds were characterized by IR, ^{1}H & ^{13}C NMR, and HR FAB MS.
11. Medina, J. C.; Salomon, M.; Kyler, K. S. *Tetrahedron Lett.* **1988**, *29*, 3773.
12. Huang, Z.; Schneider, K. C..; Benner, S. A. *J. Org. Chem.* **1991**, *56*, 3869.
13. Azuma, T.; Isono, K. *Chem. Pharm. Bull.* **1977**, *25*, 3347.
14. Nishimura, T.; Iwai, I. *Chem. Pharm. Bull.* **1964**, *12*, 352.
15. Sugimura, H.; Osumi, K.; Yamazaki, T.; Yamaya, T. *Tetrahedron Lett.* **1991**, *32*, 1813.
16. Ogilvie, K. K.; Ngen-ba, N.; Hamilton, R. G. *Can. J. Chem.* **1984**, *62*, 1622.
17. Knapp, S.; Shieh, W.-C. *Tetrahedron Lett.* **1991**, *32*, 3627.
18. Bodansky, M.; Bodansky, A. *The Practice of Peptide Synthesis*; Springer-Verlag: Berlin, 1984; pp 114-115 & 131-132.

19. This work was supported in part by Public Health Service Grant GM 35557 administered by the NIH. The authors wish to thank Satish Pillai for his help with the molecular modeling studies and Dr. Ramakanth Sarabu (Hoffmann-La Roche) for helpful comments and advise.

(Received in USA 3 November 1992; accepted 9 December 1992)