Tetrahedron Letters, Vol. 34, No. 43, pp. 6867-6870, 1993
Printed in Great Britain

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

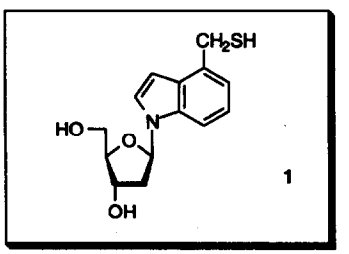
SYNTHESIS OF *N*-(2-DEOXY-β-D-RIBOSYL) 4-THIOMETHYLINDOLE: A PURINE SURROGATE FOR INCORPORATION INTO OLIGONUCLEOTIDES.

Robert S. Coleman,*,[1a] Yong Dong, and John C. Arthur[1b]

Department of Chemistry and Biochemistry
University of South Carolina
Columbia, South Carolina 29208

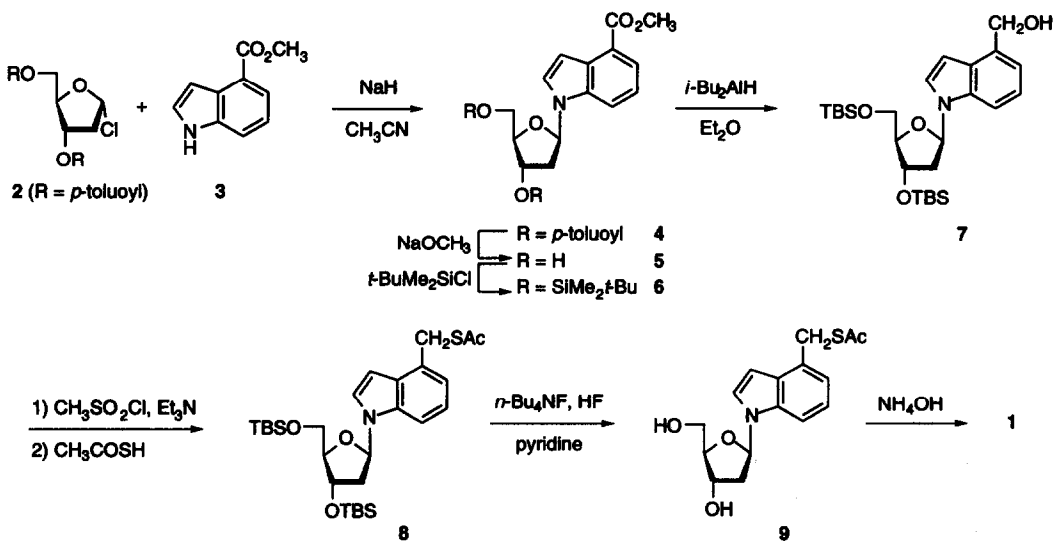
ABSTRACT: *Efficient, high-yielding syntheses of* S-acetyl-N-*(2-deoxy-β-D-ribosyl) 4-thiomethylindole (9) and the corresponding 5′-O-trityl-3′-O-phosphoramidite ester 11 are reported, suitable for direct incorporation into oligodeoxynucleotides by standard solid-phase synthesis techniques.*

In extension of our recent work on the development of modified oligonucleotides containing functionality suitable for site-specific post-synthetic modification,[2,3] we are exploring the use of the indole ring system as a purine surrogate in oligodeoxynucleotides. Recent work on synthetic incorporation of non-natural nucleic acids into oligonucleotides[4] has detailed the introduction of 5-substituted pyrimidines,[5] nucleotide-PAH metabolite adducts,[6] an aminopyrene-guanosine adduct,[7] the photoactive probe 8-bromo-2′-deoxyadenosine,[8] the DNA methyltransferase suicide substrate 5-fluoro-2′-deoxycytidine,[9] and a variety of other substituted nucleic acids including 4-substituted thymidines,[10] 6-*O*-(pentafluorophenyl)-2′-deoxyguanosine,[11] and several thionucleic acids.[12] This recent work includes our introduction of the *S*-cyanoethyl protecting group[3] for incorporation of 2′-deoxy-4-thiouridine into oligonucleotides.[2] Our interest in the development of indoles as purine surrogates arises from the capability to readily synthesize regioisomerically substituted indole systems, thereby allowing chemically reactive functionality or reporter groups to be positioned in different topographical areas of an indole-containing DNA duplex. Furthermore, the indole ring system lacks the functionality that complicates manipulation and incorporation of purines into oligonucleotides. Our prototype system, 4-thiomethylindole, was developed as a scaffolding molecule to position a reactive thiol group in the major groove of duplex DNA. The thiol functionality will be used for the site-specific incorporation of tethered functionality by direct S-alkylation or mixed disulfide formation, in a manner analogous to our published strategy for post-synthetic modification of the thiocarbonyl group of a 2′-deoxy-4-thiouridine-containing oligonucleotide.[2] Herein we report an effective synthesis of *N*-(2-deoxy-β-D-ribosyl) 4-thiomethylindole (**1**), and the corresponding phosphoramidite for incorporation of this heterocycle into oligonucleotides using standard solid-phase synthesis protocols.[13]

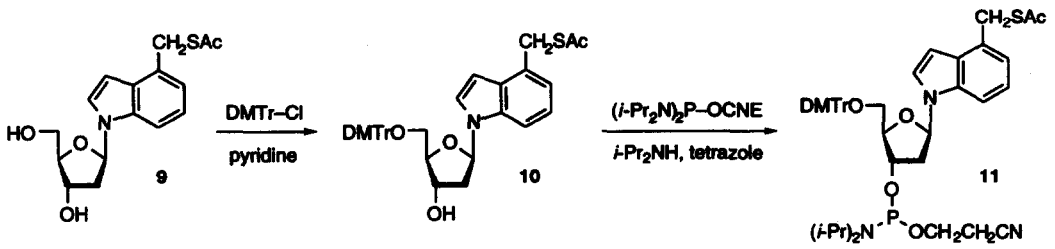


1

N-Ribosylation of methyl indole-4-carboxylate[14] (**3**) with the protected α-chlororiboside **2**[15] using the conditions of Robins[16] (NaH, CH_3CN, 0 → 25 °C) afforded the β-*N*-glycoside **4** as a single diastereomer in 86% yield. Protecting group interchange to a more suitable silyl ether was accomplished by deacylation of **4** with catalytic sodium methoxide (CH_3OH, 25 °C, 8 h) to provide diol **5** (91%) followed by silylation of **5** under standard conditions (*t*-$BuMe_2SiCl$, imidazole, DMF, 25 °C)[17] to afford **6** in quantitative yield. Alternatively, the toluoyl esters of **4** could be selectively removed by reduction with $LiAlH_4$ (1.1 equiv, THF, 0 °C) to afford **5** (70%) without significant reduction of the poorly electrophilic indole-4-carboxylate carbonyl group, which is unreactive towards reduction as a result of resonance with the indole nitrogen lone-pair. Reduction of the methyl carboxylate of **6** to the alcohol **7** with *i*-Bu_2AlH (2.5 equiv, Et_2O, –78 °C, 100%) was followed by one-pot conversion of the alcohol to the corresponding thioacetate by treatment of **7** with methanesulfonyl chloride (1.2 equiv) and Et_3N (2 equiv, –43 °C, 30 min) followed by the addition of thiolacetic acid (1.2 equiv, –43 → 25 °C). This simple procedure afforded the thioacetate **8** in 84% yield. Removal of the silyl ether protecting groups using the conditions described by Gaffney and Jones[18] (*n*-Bu_4NF/HF in pyridine, 25 °C, 40 h) afforded the diol **9**[19] (93%). The *S*-acetyl ester of **9** was hydrolyzed by treatment with conc. NH_4OH (25 °C), which are the reaction conditions used for cleavage of an oligonucleotide from the CPG support at the completion of the synthesis. β-*N*-Ribosylindole **9** proved stable to I_2 in H_2O/pyridine/THF, $P(OCH_3)_3$, and 2% trichloroacetic acid, thereby demonstrating the suitability of this system for direct introduction into an oligonucleotide using standard solid-phase synthesis conditions.[20]



Introduction of a 5′-*O*-protecting group and 3′-*O*-phosphoramidite proceeded without complication. Selective alkylation the 5′-hydroxyl group of **9** with dimethoxytrityl chloride (1.5 equiv, 2.5 equiv *i*-Pr_2NEt, pyridine 24 °C, 86%)[21] afforded the dimethoxytrityl (DMTr) ether **10**. The 3′-*O*-phosphoramidite necessary for oligonucleotide synthesis[13] was introduced by acylation of **10** with *O*-(2-cyanoethyl) *N,N,N′,N′*-tetraisopropylphosphorodiamidite (1.2 equiv, 0.5 equiv *i*-Pr_2NH, 0.5 equiv tetrazole, CH_2Cl_2, 71%)[22] to provide the desired phosphoramidite **11**[23] in good yields.



The synthesis of **9** proceeds from methyl indole-4-carboxylate **3**[14] in six steps, with an overall yield of 61%. The protected phosphoramidite derivative **11** represents a substrate suitable for direct introduction into an oligodeoxynucleotide using standard solid-phase synthesis protocols. Subsequent to chain elongation, cleavage of the newly synthesized oligonucleotide from the support with NH_4OH will unmask the thiol group, thereby providing functionality for attachment of chemically reactive or reporter groups. With the purine surrogate **1**, this tethered functionality will be positioned within the major groove of duplex B-form DNA. The chemically unique reactivity of thiol group of **1** will likely prove ideal for introduction of a variety of pendant side-chains within an oligonucleotide. Incorporation of this 4-thiomethylindole system into oligodeoxynucleotides and related studies on the use of **1** as a purine surrogate will be reported separately.[20]

ACKNOWLEDGMENTS We thank the American Cancer Society (JFRA-319 and IRG-107P), the Camille and Henry Dreyfus Foundation (NF-89-18), the Carolina Venture Fund, and American Cyanamid, Inc. for their support of this work.

REFERENCES AND NOTES

1. (a) Recipient of a Camille and Henry Dreyfus Foundation Distinguished New Faculty Award (1989-94), an American Cancer Society Junior Faculty Research Award (1990-93), and an American Cyanamid Faculty Award (1993). (b) Recipient of a U. S. Department of Education Graduate Fellowship (1991-92).
2. Coleman, R. S.; Siedlecki, J. M. *J. Am. Chem. Soc.* **1992**, *114*, 9229.
3. Coleman, R. S.; Siedlecki, J. M. *Tetrahedron Lett.* **1991**, *32*, 3033.
4. For reviews, see: Englisch, U.; Gauss, D. H. *Angew. Chem., Int. Ed. Engl.* **1991**, *30*, 613. Matteucci, M. D.; Bischofberger, N. *Ann. Rep. Med. Chem.* **1991**, *26*, 287. "Oligodeoxynucleotides: Antisense Inhibitors of Gene Expression," Cohen, J. S., ed.; CRC Press; 1989.
5. Wagner, R. W.; Matteucci, M. D.; Lewis, J. G.; Gutierrez, A. J.; Moulds, C.; Froehler, B. C. *Science* **1993**, *260*, 1510. Froehler, B. C.; Jones, R. J.; Cao, X.; Terhorst, T. J. *Tetrahedron Lett.* **1993**, *34*, 1003. Froehler, B. C.; Wadwani, S.; Terhorst, T. J.; Gerrard, S. R. *Tetrahedron Lett.* **1992**, *33*, 5307 (C-5 propyne). Sági, J.; Szemzö, A.; Ebinger, K.; Szabolcs, A.; Sági, G.; Ruff, E.; Ötvös, L. *Tetrahedron Lett.* **1993**, *34*, 2191 (C-5 alkyl, alkenyl, and alkynyl). Bradley, D. H.; Hanna, M. M. *Tetrahedron Lett.* **1992**, *33*, 6223 (5-thiocyanato-2′-deoxyuridine). Conte, M. R.; Galeone, A.; Avizonis, D.; Hsu, V. L.; Mayol, L.; Kearns, D. R. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 79 (5-hydroxymethyluracil). Haralambidis, M.; Chai, M.; Tregear, G. W. *Nucl. Acids Res.* **1987**, *15*, 4857 (C-5 appended alkylamines).

6. Steinbrecher, T.; Becker, A.; Stezowski, J. J.; Oesch, F.; Seidel, A. *Tetrahedron Lett.* **1993**, *34*, 1773. Lakshman, M. K.; Sayer, J. M.; Jerina, D. M. *J. Org. Chem.* **1992**, *57*, 3438. Kim, S. J.; Stone, M. P.; Harris, C. M.; Harris, T. M. *J. Am. Chem. Soc.* **1992**, *114*, 5480.

7. Vyas, R. R.; Nolan, S. J.; Basu, A. K. *Tetrahedron Lett.* **1993**, *34*, 2247.

8. Liu, J.; Verdine, G. L. *Tetrahedron Lett.* **1992**, *33*, 4265.

9. MacMillan, A. M.; Chen, L.; Verdine, G. L. *J. Org. Chem.* **1992**, *57*, 2989.

10. Xu, Y.-Z.; Zheng, Q.; Swann, P. F. *J. Org. Chem.* **1992**, *57*, 3839.

11. Gao, H.; Fathi, R.; Gaffney, B. L.; Goswami, B.; Kung, P.-P.; Rhee, Y.; Jin, R.; Jones, R. A. *J. Org. Chem.* **1992**, *57*, 6954.

12. Nikiforov, T. T.; Connolly, B. A. *Tetrahedron Lett.* **1992**, *33*, 2379. Rajur, S. B.; McLaughlin, L. W. *Tetrahedron Lett.* **1992**, *33*, 6081. Rao, T. S.; Jayaraman, K.; Durland, R. H.; Revankar, G. R. *Tetrahedron Lett.* **1992**, *33*, 7651. Xu, Y.-Z.; Zheng, Q.; Swann, P. F. *Tetrahedron Lett.* **1992**, *33*, 5837. See also reference 2.

13. Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett.* **1981**, *22*, 1859. Matteucci, M. D.; Caruthers, M. H. *J. Am. Chem. Soc.* **1981**, *103*, 3185. "Oligonucleotide Synthesis: a Practical Approach," Gait, M. J., ed.; IRL Press; Washington; 1984. For recent reviews, see: Engels, J. W.; Uhlmann, E. *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 716. Caruthers, M. H.; Barone, A. D.; Beaucage, S. L.; Dodds, D. R.; Fisher, E. F.; McBride, L. J.; Matteucci, M.; Stabinsky, Z.; Tang, J.-Y. *Methods Enzym.* **1987**, *154*, 287.

14. Kozikowski, A. P.; Ishida, H.; Chen, Y.-Y. *J. Org. Chem.* **1980**, *45*, 3350.

15. Bhat, C. in "Synthetic Procedures in Nucleic Acid Chemistry," vol. 1, p. 521; Zorbach, W. W.; Tipson, R. S., eds.; Interscience Publishers; 1967.

16. Kawakami, H., Matsushita, H.; Naoi, Y.; Itoh, K.; Yoshikoshi, H. *Chem. Lett.* **1989**, 235. Ramasamy, K.; Robins, R. K.; Revankar, G. R. *Nucleosides Nucleotides* **1988**, *7*, 385. Sanghvi, Y. S.; Hanna, N. B.; Larson, S. B.; Robins, R. K.; Revankar, G. R. *Nucleosides Nucleotides* **1987**, *6*, 761.

17. Corey, E. J.; Venkateswarlu, A. *J. Am. Chem. Soc.* **1972**, *94*, 6190.

18. Gaffney, B. L.; Jones, R. A. *Tetrahedron Lett.* **1982**, *23*, 2257.

19. Compound **9** was characterized: ^{1}H NMR (500 MHz, d_6-acetone) δ 7.56 (d, J = 3.4 Hz, 1 H, C3-H), 7.50 (d, J = 8.3 Hz, 1 H, C7-H), 7.11 (dd, J = 8.3, 3.4 Hz, 1 H, C6-H), 7.04 (d, J = 7.2 Hz, 1 H, C5-H), 6.56 (d, J = 3.4 Hz, 1 H, C2-H), 6.44 (dd, J = 7.8, 5.9 Hz, 1 H, C1′-H), 4.57 (ddd, J = 9.5, 5.7, 3.1 Hz, 1 H, C3′-H), 4.41 (s, 2 H, CH_2S), 3.97 (ddd, J = 9.5, 1.9, 1.4 Hz, 1 H, C4′-H), 3.70 (ABX, J = 4.3, 1.9, 1.4, Δν = 13.6 Hz, 2 H, C5′-H), 2.60 (ddd, J = 13.6, 7.8, 5.7 Hz, 1 H, C2′-H), 2.35 (ddd, J = 13.3, 5.9, 3.1 Hz, 1 H, C2′-H), 2.33 (s, 3 H, CH_3CO); ^{13}C NMR (125 MHz, d_6-acetone) δ 194.7, 136.7, 129.4, 128.5, 125.3, 122.0, 120.7, 109.9, 101.0, 87.6, 85.3, 71.9, 62.9, 40.6, 31.2, 29.8; IR (neat) ν_{max} 3387, 2924, 1686, 1439, 1287, 1053, 752 cm^{-1}; EIMS, *m/z* 321 (M^+, 50), 205 (40), 156 (35), 142 (90), 130 (base), 117 (30).

20. Coleman, R. S.; Arthur, J. C., unpublished studies.

21. Seela, F.; Bourgeois, W. *Synthesis* **1990**, 945.

22. Nielsen, J.; Dahl, O. *Nucl. Acids Res.* **1987**, *15*, 3626. Barone, A. D.; Tang, J.-Y.; Caruthers, M. H. *Nucl. Acids Res.* **1984**, *12*, 4051.

23. Compound **11** was characterized: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.39 (m, 1 H, ArH), 7.28 (m, 10 H, ArH), 7.06 (d, J = 7.1 Hz, 1 H, C5-H), 6.77 (m, 4 H, ArH), 6.52 (m, 1 H, C2-H), 6.37 (m, 1 H, C1′-H), 4.70 (m, 1 H, C3′-H), 4.40 (s, 2 H, CH_2S), 4.23 (m, 1 H, C4′-H), 3.76 (s, 3 H, OCH_3), 3.75 (s, 3 H, OCH_3), 3.57 (m, 2 H, $NCH(CH_3)_2$), 3.33 (m, 2 H, C5′-H), 2.59 (t, J = 6.4 Hz, 2 H, OCH_2CH_2CN), 2.44 (t, J = 6.4 Hz, 2 H, OCH_2CH_2CN), 2.31 (s, 3 H, CH_3CO), 1.29 (d, J = 6.8 Hz, 6 H, $NCH(CH_3)_2$), 1.08 (d, J = 6.8 Hz, 6 H, $NCH(CH_3)_2$); ^{31}P NMR (162 MHz, $CDCl_3$) δ 149.1, 148.7; IR (CH_2Cl_2) ν_{max} 3444, 2963, 1690, 1607, 1509, 1443, 1250, 1250, 1177, 1033, 829, 753 cm^{-1}; EIMS, *m/z* 824 (M^+, 30), 781 (20), 750 (5), 304 (base).

(Received in USA 21 June 1993; accepted 26 August 1993)