Pergamon

Tetrahedron Letters 43 (2002) 4581–4583

TETRAHEDRON LETTERS

A facile incorporation of the aldehyde function into DNA: 3-formylindole nucleoside as an aldehyde-containing universal nucleoside

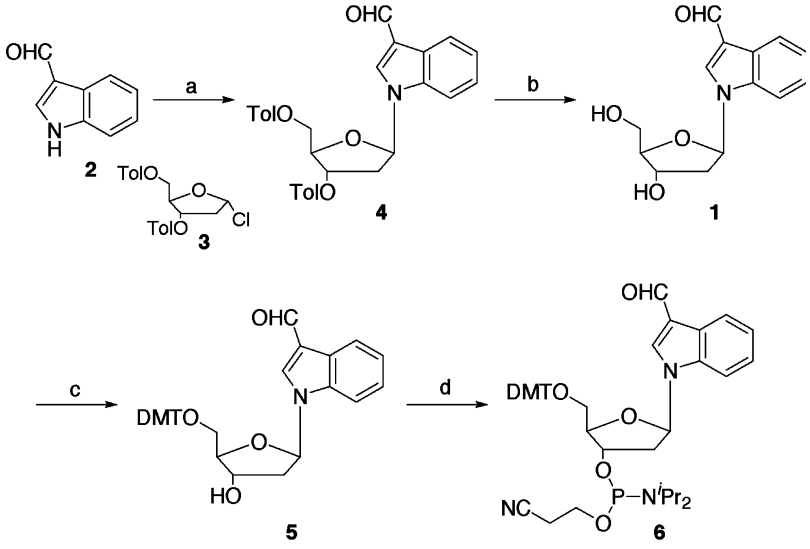
Akimitsu Okamoto, Kazuki Tainaka and Isao Saito*

Department of Synthetic Chemistry and Biological Chemistry, Faculty of Engineering, Kyoto University, CREST, Japan Science and Technology Corporation, Kyoto 606-8501, Japan

Received 25 March 2002; accepted 10 May 2002

Abstract—The first facile incorporation of an aldehyde function into DNA without any protection/deprotection of the aldehyde was achieved by the use of 3-formylindole 2′-deoxynucleoside (**1**). This building block is very easily prepared and efficiently incorporated into oligonucleotides. In addition, **1** acted as a universal nucleoside in duplex. The post-synthetic modification of oligonucleotides bearing **1** with functionalized molecules was also effective. © 2002 Published by Elsevier Science Ltd.

The assembling of functionalized oligonucleotides (ODNs) into complementary nucleic acids with target sequences has led to a rich variety of technologies exploiting new functionalities of ODNs.[1–5] The development of a new synthetic method for ODNs possessing diverse functionalities is of great interest. The aldehyde function is often used to couple biopolymers to other molecules by the process of reductive amination or the formation of adducts with hydroxylamines, hydrazines, and semicarbazides as nucleophiles.[6–11] However, aldehyde functions have rarely been applied for post-synthetic modification of ODNs, because suitable aldehyde



Scheme 1. Synthesis of 3-formylindole deoxynucleoside **1** and phosphoramidite **6**. (a) Sodium hydride, acetonitrile, 50°C, 1 h, and then **2**, acetonitrile, rt, 3 h (74%); (b) lithium hydroxide, THF–water (1:1), rt, 3.5 h (90%); (c) 4,4′-dimethoxytrityl chloride, *N*,*N*-dimethylaminopyridine, pyridine, rt, 5 h (73%); (d) $(^iPr_2N)_2PO(CH_2)_2CN$, tetrazole, acetonitrile, rt, 2 h (quant.).

Keywords: oligonucleotides; post-synthetic modification; aldehydes; universal nucleosides.

* Corresponding author. Tel.: +81-75-753-5656; fax: +81-75-753-5676; e-mail: saito@sbchem.kyoto-u.ac.jp

0040-4039/02/$ - see front matter © 2002 Published by Elsevier Science Ltd.
PII: S0040-4039(02)00888-2

building blocks for automatic ODN synthesis have not been available and such complicated and unwieldy processes as protection/deprotection and oxidative cleavage of diols were unavoidable during ODN synthesis.[10–14] It would be very useful to have an aldehyde building block mimicking nucleotides, especially if ODNs bearing multiple aldehyde functions are required. In addition, as a compatible aldehyde modifier, it is desirable that the aldehyde building block can be applied in the desired sequence, regardless of the type of complementary bases in duplex DNA.

We herein describe a convenient method for incorporating an aldehyde function into the desired position of an ODN. The aldehyde building block 3-formylindole 2′-deoxynucleoside (**1**) was easily prepared in a few steps, and its incorporation into ODN was achieved without any protection/deprotection of the formyl group during the synthesis of **1**-containing ODN. The synthesized ODNs formed stable duplexes regardless of the type of base opposite **1**. In addition, post-synthetic modification of ODN bearing an aldehyde function with functionalized hydrazides efficiently furnished the corresponding modified ODNs.

The synthetic route for **1** and its phosphoramidite is shown in Scheme 1. Commercially available 3-formylindole (**2**) was coupled with deoxyribosyl chloride (**3**)[15] to give **4** (74%). The hydrolytic deprotection of **4** afforded a free deoxynucleoside **1** (90%).[16] The primary alcohol of **1** was protected by the 4,4′-dimethoxytrityl group (73%) and quantitatively converted to cyanoethyl phosphoramidite **6**.

Amidite **6** was employed in the conventional solid-phase synthesis of ODNs. After ODN synthesis, deprotection was conducted with aqueous ammonia at 55°C for 16 h, and the resulting ODNs were purified by HPLC. The composition of the oligomer was determined by MALDI-TOF mass spectrometry (calcd 4048.67 for $[M-H]^-$, found 4048.14), and the product analysis obtained from enzymatic digestion with snake venom phosphodiesterase, nuclease P1, and alkaline phosphatase. An HPLC profile after enzymatic digestion, as shown in Fig. 1, exhibited five peaks corresponding to four natural bases and **1**.

The stability of the duplex ODN containing **1** was investigated by monitoring the melting temperatures (T_m) of the duplex. The duplexes 5′-d(GCGATG**1**GTAGCG)-3′/5′-d(CGCTAC*N*CATCGC)-3′ (N=C, G, T, and A) were prepared for T_m measurements, and the effects of the complementary bases on duplex stability were compared. In all T_m measurements, sigmoidal curves on the change of A_{254} were obtained, and the T_m values were calculated from the first derivative of the curve. The results of the T_m measurements are shown in Table 1. Selectivity of base-pairing with **1** was not observed for these duplexes. Their T_m values were less than the duplex containing the natural A–T base pair (59.2°C). Thus, the stability of a duplex containing **1** was not influenced by the base opposite **1**, and **1** in duplex ODN behaved like nitroindole[17,18] and methylbenzimidazole[19–21] nucleosides known as 'universal' nucleosides. Thus, the aldehyde function can be easily incorporated into any desired site of DNA by the use of **1**.

We next examined the modification of the formylindole **1** of ODN 5′-d(GCGATG**1**GTAGCG)-3′ with functionalized hydrazines (Fig. 2). The fully protected ODN after solid-phase synthesis on an automated DNA synthesizer was treated with 10 mM *N*,*N*-diphenylhydrazine in the presence of 10 mM sodium acetate in ethanol at 60°C for 24 h.[22] After incubation, the oligomer was deprotected by aqueous ammonia and then analyzed by HPLC. It was shown in an HPLC profile that a formyl group in the oligomer was converted to the corresponding diphenylhydrazone in 86% yield. The composition of the oligomer after HPLC

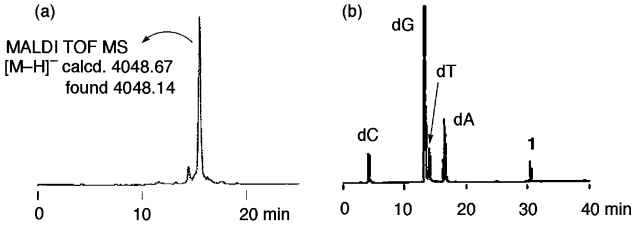


Figure 1. HPLC profiles for ODN containing 3-formylindole deoxynucleoside **1**. (a) crude oligomer 5′-d(GCGATG**1**GTAGCG)-3′; (b) nucleosides after enzymatic digestion (snake venom phosphodiesterase, nuclease P1, and alkaline phosphatase). Elution with a solvent mixture of 0.1 M triethylamine acetate (TEAA), pH 7.0, linear gradient over 20 min from 0 to 10% acetonitrile at a flow rate 1.0 mL/min.

Table 1. T_m measurement of duplex containing **1**[a]
5′-d(GCGATG**1**GTAGCG)-3′
3′-d(CGCTAC*N*CATCGC)-5′

Entry	*N*	T_m (°C)
1	C	52.8
2	G	50.0
3	T	49.7
4	A	52.7

[a] 3 μM strand concentration, 10 mM PIPES buffer, 10 mM magnesium chloride, 100 mM sodium chloride (pH 7.0).

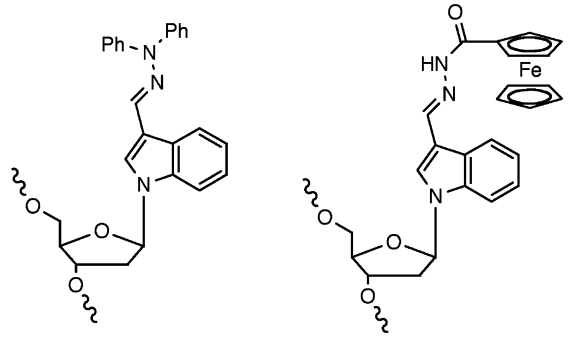


Figure 2. Products given by post-synthetic modification of 5′-d(GCGATG**1**GTAGCG)-3′.

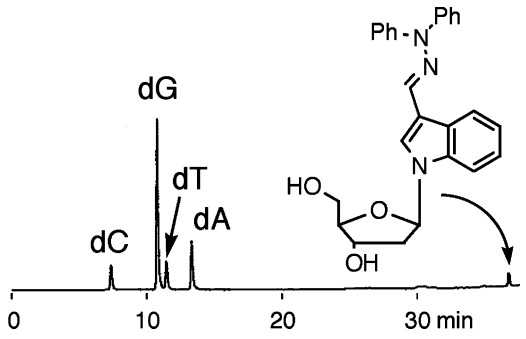


Figure 3. HPLC profiles after enzymatic digestion (snake venom phosphodiesterase, nuclease P1 and alkaline phosphatase) of **1**-containing ODN modified by diphenylhydrazine. Elution with a solvent mixture of 0.1 M triethylamine acetate (TEAA), pH 7.0, linear gradient over 20 min from 0 to 10% acetonitrile at a flow rate 1.0 mL/min.

purification was determined by MALDI-TOF mass spectrometry (calcd 4214.89 for $[M-H]^-$, found 4215.15) and the product analysis obtained from enzymatic digestion. An example of an HPLC profile after enzymatic digestion is shown in Fig. 3. The peaks shown in Fig. 3 suggest that *N*,*N*-diphenylhydrazine was effectively incorporated into **1** in ODN without any side reaction.[23] Similarly, ferrocenecarbohydrazide was also incorporated into **1** of ODN 5′-d(GCGATG**1**GTAGCG)-3′ in 50% yield (calcd 4274.72 for $[M-H]^-$, found 4275.23). Ferrocene-modified ODNs have been noted as excellent electrochemical probes in recent years.[24]

In conclusion, the facile incorporation of an aldehyde function into DNA without any protection/deprotection of the aldehyde was achieved by the use of 3-formylindole 2′-deoxynucleoside (**1**). This building block is very easily prepared and efficiently incorporated into oligonucleotides. In addition, **1** can be used as a universal nucleoside in duplex oligonucleotides. The post-synthetic modification of **1**-containing ODN with a variety of functionalized molecules is also possible. This post-synthetic modification method will be applicable to site-selective DNA labeling and bioconjugation with various functionalized groups.

References

1. Goodchild, J. *Bioconjugate Chem.* **1990**, *1*, 165–186.
2. Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 543–590.
3. Beaucage, S. L.; Iyer, R. P. *Tetrahedron* **1993**, *49*, 1925–1963.
4. Beaucage, S. L.; Iyer, R. P. *Tetrahedron* **1993**, *49*, 6123–6194.
5. Verma, S.; Eckstein, F. *Ann. Rev. Biochem.* **1998**, *67*, 99–134.
6. Coombs, M. M.; Livingston, D. C. *Biochim. Biophys. Acta* **1969**, *174*, 161–173.
7. Talpaert-Borlé, M.; Liuzzi, M. *Biochim. Biophys. Acta* **1983**, *740*, 410–416.
8. Vasseur, J.-J.; Peoc'h, D.; Rayner, B.; Imbach, J.-L. *Nucleosides Nucleotides* **1991**, *10*, 107–117.
9. Proudnikov, D.; Mirzabekov, A. *Nucleic Acids Res.* **1996**, *24*, 4535–4542.
10. Trévisiol, E.; Renard, A.; Defrancq, E.; Lhomme, J. *Tetrahedron Lett.* **1997**, *38*, 8687–8690.
11. Tilquin, J.-M.; Dechamps, M.; Sonveaux, E. *Bioconjugate Chem.* **2001**, *12*, 451–457.
12. Kremsky, J. N.; Wooters, J. L.; Dougherty, J. P.; Meyers, R. E.; Collins, M.; Brown, E. L. *Nucleic Acids Res.* **1987**, *15*, 2891–2909.
13. Urata, H.; Akagi, M. *Tetrahedron Lett.* **1993**, *34*, 4015–4018.
14. Dechamps, M.; Sonveaux, E. *Nucleosides Nucleotides* **1998**, *17*, 697–709.
15. Hoffer, M. *Chem. Ber.* **1960**, *93*, 2777–2781.
16. Spectral data of **1**: ^{1}H NMR (DMSO-d_6) δ 9.92 (s, 1H), 8.55 (s, 1H), 8.09 (d, 1H, $J=7.6$ Hz), 7.70 (d, 1H, $J=8.4$ Hz), 7.29 (m, 2H), 6.42 (t, 1H, $J=6.4$ Hz), 5.36 (d, 1H, $J=4.0$ Hz), 4.97 (t, 1H, $J=5.2$ Hz), 4.37 (m, 1H), 3.87 (q, 1H, $J=4.0$ Hz), 3.55 (m, 2H), 2.49 (m, 1H), 2.35 (m, 1H); MS (FAB, NBA/CH_2Cl_2) *m*/*e* 261 $[M+H]^+$; HRMS (FAB) calcd for $C_{14}H_{16}O_4N$ $[M+H]^+$ 262.1080, found 262.1074.
17. Loakes, D.; Brown, D. M. *Nucleic Acids Res.* **1994**, *22*, 4039–4043.
18. Loakes, D. *Nucleic Acids Res.* **2001**, *29*, 2437–2447.
19. Seela, F.; Bourgeois, W.; Rosemeyer, H.; Wenzel, T. *Helv. Chim. Acta* **1996**, *79*, 488–498.
20. Guckian, K. M.; Morales, J. C.; Kool, E. T. *J. Org. Chem.* **1998**, *63*, 9652–9656.
21. Morales, J. C.; Kool, E. T. *Biochemistry* **2000**, *39*, 12979–12988.
22. Post-synthetic modification via on-column conjugation: A solution of the coupling reagents (10 mM diphenylhydrazine hydrochloride, and 10 mM sodium acetate in 1.5 mL ethanol) was prepared. To the solution was added the resin-bound DNA (containing ~0.5 μmol); it was then shaken at 60°C for 24 h, and then the solvent was evaporated in vacuo. The resin was treated with 28% aqueous ammonia (1 mL) for 16 h at 55°C and concentrated under vacuum, then filtered. Oligonucleotides were purified by reverse phase HPLC on a 5-ODS-H column (10×150 mm, elution with a solvent mixture of 0.1 M triethylamine acetate (TEAA), pH 7.0, linear gradient over 40 min from 5 to 100% acetonitrile at a flow rate 3.0 mL/min).
23. The T_m of 5′-d(GCGATG**X**GTAGCG)-3′/5′-d(CGCTACCCATCGC)-3′ (X denotes *N*,*N*-diphenylhydrazine-modified **1**) was 54.2°C.
24. Yu, C. J.; Wan, Y.; Yowanto, H.; Li, J.; Tao, C.; James, M. D.; Tan, C. L.; Blackburn, G. F.; Meade, T. J. *J. Am. Chem. Soc.* **2001**, *123*, 11155–11161.