Available online at www.sciencedirect.com
SCIENCE DIRECT®
Tetrahedron Letters
Tetrahedron Letters 46 (2005) 2235–2238

A cysteine-appended deoxyuridine for the postsynthetic DNA modification using native chemical ligation

Shuji Takeda, Shinya Tsukiji and Teruyuki Nagamune*

Department of Chemistry and Biotechnology, School of Engineering, The University of Tokyo, Tokyo 113-8656, Japan

Received 12 January 2005; revised 1 February 2005; accepted 3 February 2005

Abstract—A new deoxyuridine derivative **6** bearing a cysteine group at the C5 position was synthesized and incorporated into oligodeoxynucleotide (ODN) by phosphoramidite chemistry. The postsynthetic DNA modification of the cysteine-containing ODN using native chemical ligation with thioesters of biotin and green fluorescent protein variant was successfully demonstrated.

Modified oligonucleotides equipped with diverse functional molecules have numerous applications in structural biology, biotechnology, materials science, and medicine. The most common strategy for DNA modification involves phosphoramidite chemistry that enables the direct incorporation of functionalized nucleotides at defined sites within DNA strands during solid-phase oligonucleotide synthesis. A variety of phosphoramidite building blocks bearing, for example, unnatural bases,[1] fluorescent probes,[2] electroactive compounds,[3] and saccharides,[4] have been synthesized and incorporated into DNA. While this approach is general and widely used, many molecules of interest to be incorporated into DNA oligomers are not stable under the conditions used in DNA synthesis. Each new modified nucleotide must have sufficient stability against repetitive oxidation steps during oligomer synthesis and basic treatment for full deprotection and cleavage from a solid-phase support.[5] Thus, there remains a need for efficient methods for modifying DNA with functional molecules, which is unsuitable for the phosphoramidite chemistry, such as peptides and proteins.

An attractive alternative strategy for DNA modification is to incorporate a versatile building block bearing a reactive functional group, which allows conjugation of any desired molecule by chemical reaction after synthesis of the DNA strand. Several strategies currently used for the postsynthetic DNA modification involve incorporation of thiol- or amine-functionalized phosphoroamidites into DNA, at the 5′–/3′-terminus or internal positions, and its subsequent chemical modification using traditional bioconjugation techniques.[6] These strategies have found general applicability for attaching various compounds, ranging from small molecules to large proteins such as antibodies, to the oligonucleotides under physiological conditions. However, in preparation of protein–oligonucleotide conjugates, cross-reaction involving native reactive groups present on protein surfaces often generates conjugates with high heterogeneity. Recently, Dey and Sheppard developed the ketone-modified DNA as a new platform for DNA modification.[7] Although this approach offers a versatile and convergent method for the chemoselective modification of DNA with various aminooxy compounds by oxime formation, this method is still limited to the attachment of relatively small compounds synthetically accessible, thus not applicable to large proteins.

We have recently reported a convenient method for conjugating synthetic oligonucleotides to recombinant proteins specifically at their C-terminus.[8] Our approach involves the modification of 5′- or 3′-amine-functionalized oligonucleotides with a cysteine, followed by native chemical ligation (NCL)[9] with a recombinantly expressed protein containing a C-terminal α-thioester to form a stable amide bond. A similar approach has been reported by Gait and coworkers to prepare peptide–oligonucleotide conjugates.[10] These successful demonstrations motivated us to synthesize a new versatile phosphoramidite building block, which can be incorporated at any position within DNA strands during DNA

Keywords: Native chemical ligation; Intein; Deoxyuridine; DNA.
* Corresponding author. Tel.: +81 3 5841 7328; fax: +81 3 5841 8657; e-mail: nagamune@bio.t.u-tokyo.ac.jp

0040-4039/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tetlet.2005.02.018

synthesis, for an efficient postsynthetic DNA modification using NCL.

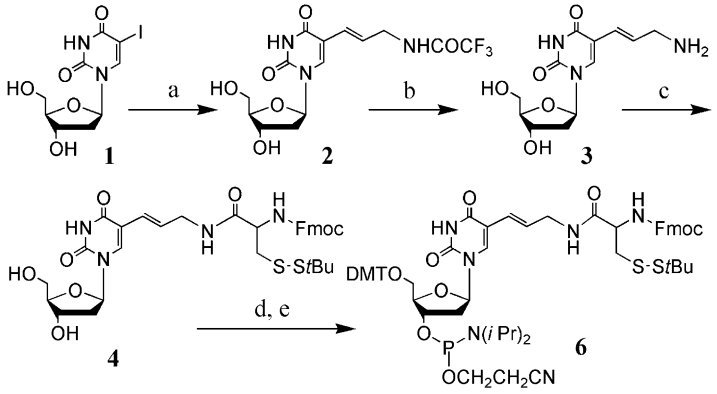
On the basis of previous studies, we designed a new deoxyuridine derivative **6** bearing a protected cysteine group at the C5 position via an allylamine linker. The C5 position of (deoxy)uridine is an appropriate site to introduce functionality without interfering with oligonucleotides duplex stability and recognition by RNA and DNA polymerases.[11] *N*-α-Fmoc-*S*-*tert*-butylthio-L-cysteine having its amino and thiol group protected with Fmoc and *tert*-butylthiol, respectively, was chosen as a protected cysteine to be introduced. Both protecting groups are stable under the conditions for DNA synthesis, but smoothly removed during full deprotection and NCL steps as reported previously.[8]

Phosphoramidite **6** was synthesized as shown in Scheme 1. 5-(3″-aminoallyl)-2′-deoxyuridine **3** was prepared in two steps from 5-iodo-deoxyuridine according to the procedure reported by Dey and Sheppard.[7] In brief, the palladium-catalyzed Heck reaction was used to couple 5-iodo-deoxyuridine to *N*-allyltrifluoroacetamide to give **2**, which was followed by treatment with 28% aqueous ammonia to remove the trifluoroacetamide group, affording **3**. Condensation of **3** with the *N*-α-Fmoc-*S*-*tert*-butylthio-L-cysteine pentafluorophenyl ester (Novabiochem) under standard conditions afforded the cysteine-appended deoxyuridine **4** in 49%.[12] Nucleoside **4** was further treated with 4,4′-dimethoxytrityl chloride in the presence of 4-(dimethylamino)pyridine to give the 5′-*O*-DMT-protected nucleoside **5** in 15%.[13] Subsequent 3′-*O*-phosphitylation with 2-cyanoethyl diisopropylchlorophosphoramidite yielded phosphoramidite **6** in a quantitative yield.[14]

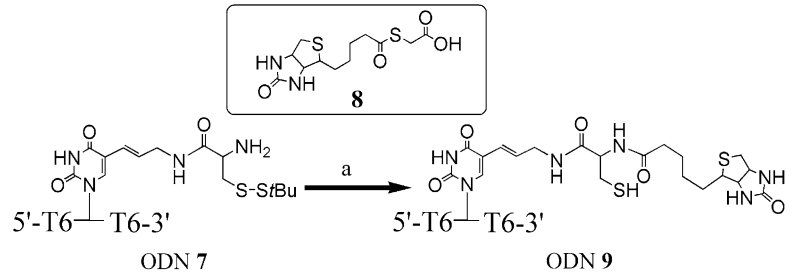
With compound **6** in hand, a cysteine-containing ODN (T6dU^{cys}T6, **7**), incorporating the modified base dU^{cys} at a central position, was readily prepared by standard automated DNA synthesis methods. After synthesis, the ODN was fully deprotected and cleaved from a solid-phase support by ammonia hydroxide treatment in good yield and purity. The crude ODN was purified by a reverse phase cartridge, followed by a reverse phase HPLC (RP-HPLC).[15] The product was satisfactorily characterized by MALDI-TOF-MS spectrometry.[16]

To investigate the chemical reactivity of the cysteine-containing ODN, we attempted NCL between the ODN **7** and the biotin-thioester **8** (Scheme 2).[17] A reaction mixture of ODN **7** and biotin-thioester was incubated at room temperature in a ligation buffer containing 2-mercaptoethanesulfonic acid (MESNA) and was analyzed by RP-HPLC.[15] Time course analysis revealed nearly quantitative formation of the expected biotin–ODN conjugate, verified by MALDI-TOF-MS analysis, within 12 h without significant side reactions (Fig. 1).[18]

Next we applied our strategy to modify oligonucleotides with proteins in a site-specific manner. Expressed protein ligation was performed to ligate the ODN **7** to the



Scheme 1. (a) CH_2=$CHCH_2NHCOCF_3$, $Na_2[PdCl_4]$, 50% DMF in NaOAc buffer (0.1 M, pH 5.2), 70 °C, 2 h (**2**, 59%); (b) NH_4OH (**3**, 100%); (c) *N*-α-Fmoc-*S*-*tert*-butylthio-L-cysteine pentafluorophenyl ester, Et_3N, DMF (**4**, 59%); (d) DMTCl, DMAP, dry pyridine (**5**, 15%); (e) (*i*-Pr_2N)($CNCH_2CH_2O$)PCl, DIEA, CH_2Cl_2.



Scheme 2. Biotin-thioester **8** (90 equiv), ligation buffer (20 mM Tris–HCl, 500 mM NaCl, 1 mM EDTA, 50 mM MESNA, pH 8.5), rt.

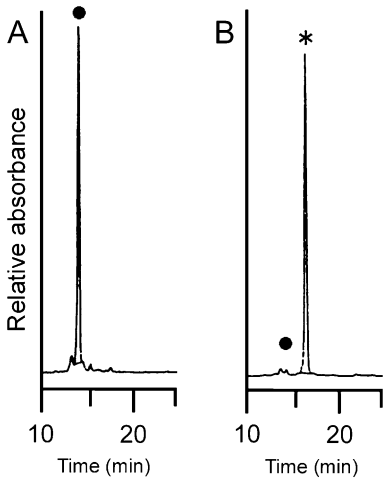


Figure 1. RP-HPLC analysis of the biotinylation of ODN **7**. (A) ODN **7** (B) ODN **7** after ligation reaction at 4 °C for 72 h. The ODN and ligation product are denoted by • and ∗, respectively.

C-terminus of an enhanced green fluorescent protein (EGFP) to give a T-shaped protein–DNA conjugate.[19] An α-thioester form of EGFP was prepared by intein-fusion technique and was purified by affinity chromatography over 95% homogeneity, as previously reported.[8] Progress of NCL between the EGFP-thioester and the ODN **7** was monitored by SDS-PAGE under a reducing condition. As shown in Figure 2, the ligation proceeded to approximately 80% completion after incubation at 4 °C for two days. In contrast, a control reaction of the EGFP-thioester with an ODN lacking the dU^{Cys} (T_{12}) under the same condition resulted in no change of the migrating band as expected (data not shown). Next we prepared a randomized ODN **10** (CGTTATCCCdU^{Cys}GATTCTGTGG) as described above,[20] and performed the NCL with the EGFP-thioester. Formation of the EGFP-ODN **10** conjugate was

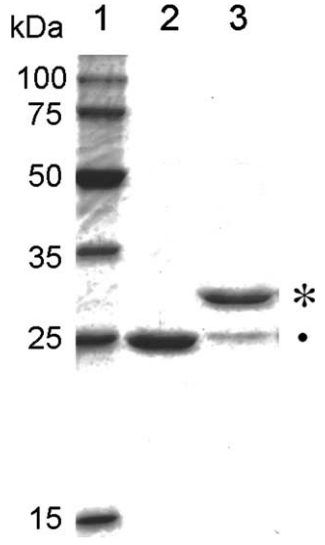


Figure 2. SDS-PAGE analysis. Lane 1, molecular weight marker; 2, EGFP-thioester; 3, EGFP-thioester mixed with ODN 7. The EGFP-thioester and ligation product are denoted by • and ∗, respectively.

clearly observed, however, the reaction yield was low (approximately 1%) (data not shown). This is probably due to folding of the single strand ODN and thus less accessibility of the cysteine group, which is a general issue in handling single strand DNAs. Future study will be aimed at optimizing the reaction condition for more efficient chemical ligation, for instance by performing the reaction in the presence of denaturant or complementary DNAs.

In conclusion, we have developed a new deoxyuridine derivative to incorporate a cysteine group at any position in oligonucleotides by automated DNA synthesis. A cysteine-appended DNA should be a general scaffold for the postsynthetic DNA modification with various thioester-tagged molecules, ranging from small compounds to recombinant proteins, using NCL.

Acknowledgements

We are grateful to Professor Kazuhiko Saigo and Dr. Yasuhiro Ishida for their supports with the NMR measurement, as well as Professors Makoto Komiyama and Hiroyuki Asanuma for allowing us to use their DNA synthesizer. Dr. Akinori Kuzuya is acknowledged for his helpful assistance. We thank Mr. Satoshi Goda for providing a synthetic intermediate.

References and notes

1. (a) Hirao, I.; Ohtsuki, T.; Fujiwara, T.; Mitsui, T.; Yokogawa, T.; Okuni, T.; Nakashima, H.; Takio, K.; Yabuki, T.; Kigawa, T.; Kodama, K.; Yokogawa, T.; Nishikawa, K.; Yokoyama, S. *Nat. Biotechnol.* **2002**, *20*, 177; (b) Tanaka, K.; Tengeiji, A.; Kato, T.; Toyama, N.; Shionoya, M. *Science* **2003**, *299*, 1212; (c) Meggers, E.; Holland, P. L.; Tolman, W. B.; Romesberg, F. E.; Schultz, P. G. *J. Am. Chem. Soc.* **2000**, *122*, 10714.
2. (a) Tyagi, S.; Kramer, F. R. *Nat. Biotechnol.* **1996**, *4*, 303; (b) Tong, A. K.; Li, Z.; Jones, G. S.; Russo, J. J.; Ju, J. *Nat. Biotechnol.* **2001**, *19*, 756; (c) Heyduk, T.; Heyduk, E. *Nat. Biotechnol.* **2002**, *20*, 171.
3. (a) Anne, A.; Blanc, B.; Moiroux, J. *Bioconjugate Chem.* **2001**, *12*, 396; (b) Yu, C. J.; Wang, H.; Wan, Y.; Yowanto, H.; Kim, J. C.; Donilon, L. H.; Tao, C.; Strong, M.; Chong, Y. *J. Org. Chem.* **2001**, *66*, 2937; (c) Navarro, A.-E.; Spinelli, N.; Moustrou, C.; Chaix, C.; Mandrand, B.; Brisset, H. *Nucleic Acids Res.* **2004**, *32*, 5310; (d) Mori, K.; Subasinghe, C.; Cohen, J. S. *FEBS Lett.* **1989**, *249*, 213; (e) Whittemore, N. A.; Mullenix, A. N.; Inamati, G. B.; Manoharan, M.; Cook, P. D.; Tuinman, A. A.; Baker, D. C.; Chamber, J. Q. *Bioconjugate Chem.* **1999**, *10*, 261.
4. (a) Sheppard, T. L.; Wong, C.-H.; Joyce, G. F. *Angew. Chem., Int. Ed.* **2000**, *39*, 3662; (b) Matsuura, K.; Hibino, M.; Ikeda, T.; Yamada, Y.; Kobayashi, K. *Chem. Eur. J.* **2004**, *10*, 352; (c) Maier, M. A.; Yannopoulos, C. G.; Mohamed, N.; Roland, A.; Fritz, H.; Mohan, V.; Just, G.; Manoharan, M. *Bioconjugate Chem.* **2003**, *14*, 18.
5. 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 Enzymol.* **1987**, *154*, 287.
6. (a) MacMillan, A. M.; Verdine, G. L. *Tetrahedron* **1991**, *47*, 2603; (b) Shinozuka, K.; Kohgo, S.; Ozaki, H.; Sawai,

H. *Chem. Commun.* **2000**, 59; (c) de la Torre, B. G.; Morales, J. C.; Aviñó, A.; Iacopino, D.; Ongaro, A.; Fitzmaurice, D.; Murphy, D.; Doyle, H.; Redmond, G.; Eritja, R. *Helv. Chim. Acta* **2002**, *85*, 2594.
7. Dey, S.; Sheppard, T. L. *Org. Lett.* **2001**, *3*, 3983.
8. Takeda, S.; Tsukiji, S.; Nagamune, T. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2407.
9. Dawson, P. E.; Kent, S. B. H. *Annu. Rev. Biochem.* **2000**, *69*, 923.
10. Stetsenko, D. A.; Gait, M. J. *J. Org. Chem.* **2000**, *65*, 4900.
11. (a) Ahmadian, M.; Zhang, P.; Bergstrom, D. E. *Nucleic Acids Res.* **1998**, *26*, 3127; (b) Lee, S. E.; Sidorov, A.; Gourlain, T.; Mignet, N.; Thorpe, S. J.; Brazier, J. A.; Dickman, M. J.; Hornby, D. P.; Grasby, J. A.; Williams, D. M. *Nucleic Acids Res.* **2001**, *29*, 1565; (c) Latham, J. A.; Johnson, R.; Toole, J. J. *Nucleic Acids Res.* **1994**, *22*, 2817; (d) Ahmadian, M.; Zhang, P.; Bergstrom, D. E. *Nucleic Acids Res.* **1998**, *26*, 3127; (e) Sawai, H.; Ozaki, A. N.; Satoh, F.; Ohbayashi, T.; Masud, M. M.; Ozaki, H. *Chem. Commun.* **2001**, 2604.
12. The nucleoside **4** was satisfactorily characterized by ^{1}H NMR: ^{1}H NMR (*d*-DMSO) δ 11.41 (s, 1H), 8.24 (s, 1H), 7.98 (s, 1H), 7.89 (d, 2H), 7.68–7.75 (m, 2H), 7.42 (t, 2H), 7.32 (t, 2H), 6.40–6.45 (m, 1H), 6.13–6.18 (m, 2H), 5.23 (m, 1H), 5.05 (t, 1H), 4.34 (m, 1H), 4.22–4.30 (m, 5H), 3.76–3.79 (m, 2H), 3.58–3.61 (m, 2H), 3.32 (m, 1H), 2.91–3.09 (m, 2H), 2.11 (t, 2H), 1.30 (s, 9H).
13. The nucleoside **5** was characterized by ^{1}H NMR and MALDI-TOF-MS: ^{1}H NMR ($CDCl_3$) δ 9.82 (s, 1H), 7.65–7.74 (m, 2H+1H), 7.58 (m, 2H), 7.34–7.39 (m, 2H), 7.23–7.29 (m, 9H+2H), 6.82 (d, 4H), 6.73 (m, 1H), 6.47 (m, 1H), 6.22–6.27 (m, 1H), 5.90 (d, 1H), 5.54 (d, 1H), 4.54 (m, 1H), 4.34–4.39 (m, 2H+1H), 4.20 (m, 1H), 4.10 (m, 1H), 3.77 (s, 6H), 3.58 (m, 2H), 3.43 (dd, 1H), 3.34 (dd, 1H), 3.03 (m, 2H), 2.54 (m, 1H), 2.31 (m, 2H) 1.30 (s, 9H): MALDI-TOF-MS (Matrix: dithranol) calcd for **5** $[M+Na]^+$ = 1021.35, obsd. 1021.67.
14. The phosphoramidites **6** was identified by ^{32}P NMR, and MALDI-TOF-MS. **6**: ^{32}P NMR δ 149.56, 149.96: MALDI-TOF-MS (Matrix: dithranol) calcd for **6** $[M+Na]^+$ = 1221.46, obsd. 1221.03.
15. HPLC was carried out on a μBondasphere C18 column (5 μm, 150 × 3.8 mm, Waters) eluted with 0.1 M ammonium acetate buffer (pH 7.0) containing 10–20% acetonitrile in a linear gradient over 30 min at a flow rate of 0.5 mL/min, detected at 260 nm.
16. ODN **7** was identified by MALDI-TOF-MS (Matrix: 3-hydroxypicolinic acid) ODN **7** calcd for $[M+H]^+$ = 4141.43, obsd 4137.74.
17. Biotin-thioester **8** was synthesized according to the following paper: Tolbert, T. J.; Wong, C.-H. *Angew. Chem., Int. Ed.* **2002**, *41*, 2171.
18. Biotinylation reaction was carried out in a total volume of 10 μL of ligation buffer (20 mM Tris–HCl, 500 mM NaCl, 1 mM EDTA, 50 mM MESNA, pH 8.5). Final concentrations of biotin-thioester **8** and ODN **7** were 9 mM and 100 μM, respectively. The ligation products were characterized by MALDI-TOF-MS (Matrix: 3-hydroxypicolinic acid). ODN **9** calcd for $[M+H]^+$ = 4261.76, obsd 4265.80.
19. Expressed protein ligation was carried out in a total volume of 10 μL of ligation buffer. Final concentrations of EGFP-thioester and ODN **7** were 9 and 100 μM, respectively.
20. ODN **10** was identified by MALDI-TOF-MS (Matrix: 3-hydroxypicolinic acid) ODN **10** calcd for $[M+H]^+$ = 6617.11, obsd 6614.46.