Pergamon

Bioorganic & Medicinal Chemistry Letters, Vol. 4, No. 8, pp. 995-1000, 1994

0960-894X/94 $6.00+0.00

0960-894X(94)E0089-W

AN APPROACH TOWARDS THIOL MEDIATED LABELLING IN THE MINOR GROOVE OF OLIGONUCLEOTIDES

Mark E. Douglas, Barbro Beijer and Brian S. Sproat*
European Molecular Biology Laboratory,
Meyerhofstrasse 1, 69012 Heidelberg, GERMANY

Abstract: The preparation of an appropriately protected 2′-*O*-(2-thioethyl)uridine phosphoramidite is reported along with its incorporation into an oligonucleotide and post-synthetic labelling.

The labelling of oligonucleotides with non-radioactive reporter groups and active probes has become a standard method for the detection and assay of these molecules. Labels are usually introduced by attachment to a free amino or thiol group within an oligonucleotide[1], and as such operations are carried out most conveniently at the end of the synthesis, the functionality should be protected until required and then be readily deblocked and made accessible to electrophiles.

Thiols have been used in this context for several years now[2], their activity being derived from the fact that they are very good nucleophiles and that there are a wide range of reactive labels (eg. haloacetyl and maleimido derivatives of dyes and photoactivatable compounds) with which they readily couple. There are various sites in a oligonucleotide where a thiol can be incorporated, the position being determined primarily by the type of probe with which it will react and the use to which the labelled oligonucleotide will be put.

The most accessible modifications are those that are situated at the 5′ or 3′- termini of the oligonucleotide. 5′-Terminal modifications that are incorporated during standard automated oligonucleotide synthesis are widely used[1], typical examples of these are the trityl protected ω-thioalkyl phosphoramidites of Connolly et al[3], the ω-thioalkyl H-phosphonates of Sinha[4] and the trityl protected 5′-thio-2′,5′-dideoxyribonucleoside 3′-*O*-phosphoramidites of Sproat[5]. All of these modifications are added to the completed oligonucleotide at the end of the synthesis and allow the thiol function to be selectively deblocked as desired.

For DNA oligonucleotides, 3′-terminal modifications are usually introduced by carrying out the synthesis on a solid support containing a disulphide bond between the support and oligonucleotide and then liberating the thiol function by reduction of the disulphide bond with DTT before or after the main ammonia deblock step[6-8]. The main advantage of modifications such as these is that they do not interfere with hybridisation, on the negative side however, they are not amenable to multiple labelling. There are, at present, two main ways in which a reactive thiol may be incorporated 'internally' within the oligonucleotide, these being modification of the phosphate backbone, and incorporation of modified residues with masked thiols present in the nucleobases.

There are two main approaches to the insertion of sulphur and thiol linkers within the phosphate backbone. The first uses standard phosphoramidite chemistry[9,10] and simply involves oxidation of the internucleotide phosphite bond with sulphur (either in the form of elemental sulphur or one of a number of commercially available sulphurising agents)[12] to give a phosphorothioate diester rather than a phosphodiester bond. As the sulphur has similar reactivity to a thiol, it can be coupled to a variety of probes[13].

The main advantages are that introduction of the sulphur is simple and that the modification can be placed at any position within the backbone. The main disadvantage is that the sulphurisation reaction produces two diastereoisomers at each internucleotidic bond, which may be troublesome when the oligonucleotide is to be used in biological systems.

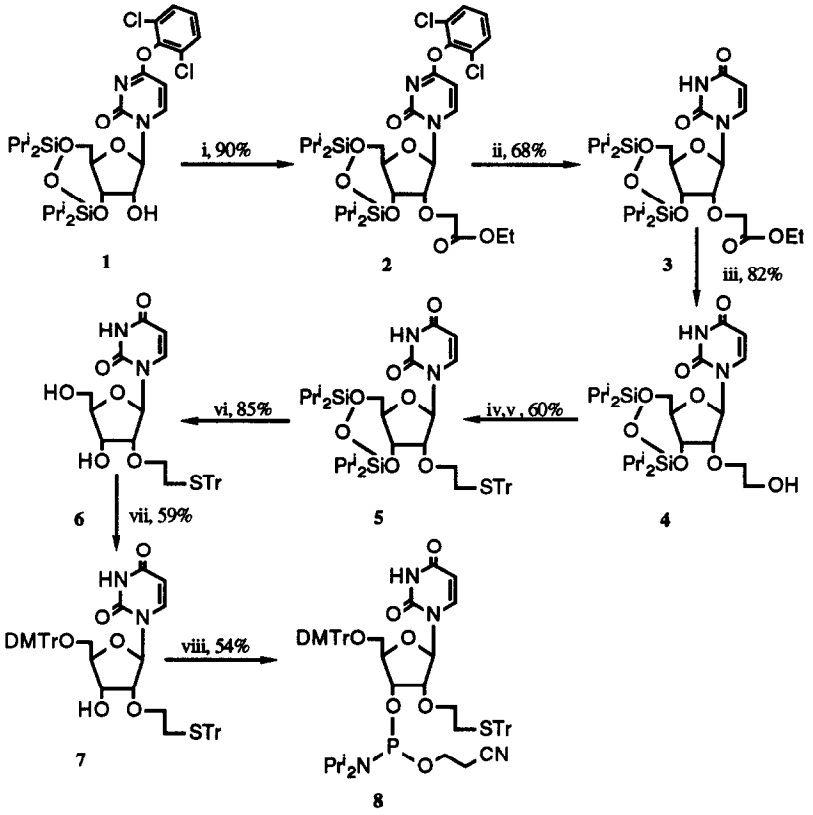
The second approach uses H-phosphonate chemistry[14] and relies on the oxidation of the internucleotide H-phosphonate diester bond with an alkylamine in carbon tetrachloride to give a phosphoramidate linkage. If the alkylamine has an internal disulphide bond then this can be reduced with DTT to give a thiol. A recent application has been demonstrated by McLaughlin[15] who used H-phosphonate chemistry to incorporate an alkylamine with a terminal protected amine and an internal disulphide bond. These can be deblocked separately to allow the incorporation of amine and thiol specific probes at different positions within the same sequence. This technique has been applied to incorporate two different dyes for Fluorescence Energy Transfer Spectroscopy of oligonucleotides[14]. Again, however, there is the problem of diastereoisomers.

Modified bases containing masked thiols can be incorporated at any position within a sequence but it is desirable that the modification does not interfere with hybridisation, thus the position for modification within the base is restricted to the 5 and 6-positions of pyrimidines[16,17] and the 8-position of purines.

We wish to report here the preparation of a nucleoside phosphoramidite containing a masked thioalkyl function in the 2′-position, and its incorporation into an oligonucleotide and subsequent labelling. The 2′-position in the sugar residue is ideal for the attachment of probes and labels as interference with hybridisation is unlikely and this position can be used to investigate the binding of enzymes and proteins to double stranded nucleic acid sequences. The key reaction in the synthesis is the 2′-*O*-alkylation of uridine derivative **1** (scheme) with ethyl bromoacetate in the presence of the strong non-nucleophilic base, 2-*tert*-butylimino-2-diethylamino-1,3-dimethylperhydro-1,3,2-diazaphosphorin (BDDDP)[18]. This, in combination with the base and sugar protection, allows the alkylation to be carried out directly on the nucleoside rather than on a sugar alone which then has to be coupled to a nucleobase, as was recently reported by Häner[19]. A similar alkylation leading to a modified adenosine residue carrying a 5-aminopentyl linker in the 2′-position has recently been reported by Cook *et al* [20].

Thus, compound **1** was rapidly alkylated with ethyl bromoacetate and BDDDP to give the 2′-*O*-ethoxycarbonylmethyl ether, **2** in 90% yield. This is in complete contrast to methylation of the 2′-hydroxyl group of compound **1** using iodomethane where the overall yield is generally around 70% and takes several hours to reach completion[18]. The 2,6-dichlorophenoxy nucleobase protection was then removed by treatment with oximate to give **3**. This is necessary otherwise reduction of the ester results in concomitant reduction of the heterocycle. The ester was then readily reduced to the 2-hydroxyethyl ether derivative **4**, by treatment with alcoholic sodium borohydride. Bromination was then carried out by treatment of **4** with carbon tetrabromide in the presence of triphenylphosphine in refluxing acetonitrile. Here it was necessary to add the carbon tetrabromide solution in acetonitrile dropwise to the refluxing solution containing nucleoside and

Scheme.



Key : i, BDDDP(1.5 eq), ethyl bromoacetate (2 eq), CH_3CN, rt , 1h; ii, 2- nitrobenzaldoxime (2.28 eq), 1,1,3,3- tetramethyguanidine (2.08 eq), CH_3CN, rt, 12h; iii, $NaBH_4$ (2.5 eq), CH_3OH (15 eq), t-BuOH, 80^0C, 1.25h; iv, Ph_3P (1.3 eq), CBr_4 (1.3 eq), CH_3CN, 90^0C, *then* v trityl mercaptan (1.3 eq), DBU (1.1 eq), benzene, 90^0C, 3h; vi, TBAF (2.2 eq), THF; vii, DMTrCl (1.2 eq), Et_3N (1.5 eq), pyridine, rt, 12h; viii, 2-cyanoethoxy-N,N-diisopropylamino chlorophosphine (1.8 eq), N,N-diisopropylethylamine (2.5 eq), CH_2Cl_2, rt, 28h.

DMTr = dimethoxytrityl. Pr^i = isopropyl. Tr = trityl.

triphenylphosphine so that the reactive brominating species was trapped by the alcohol before it could react with more triphenyl phosphine and thus become deactivated. At the end of the reaction, the volatile material was removed *in vacuo* and the gum redissolved in benzene and brought to reflux (CAUTION - TOXIC - all manipulations should be carried out wearing a lab coat, safety glasses and gloves in an efficient fume hood). The intermediate 2-bromoethyl ether was not isolated. Tritylmercaptan and 1,8-diazabicyclo[5.4.0]undec-7-ene were then added and the mixture was heated further giving the tritylmercapto derivative **5** which was then desilylated in the usual manner to yield compound **6** in good yield. Compound **6** was then dimethoxytritylated in reasonable yield giving compound **7**. Phosphitylation of compound **7** was carried out in accordance with standard procedures[22], however, even in the presence of a large excess of the phosphitylating agent, the reaction required about 26 h to reach completion, and because of this, a large amount of H-phosphonate was formed thereby somewhat reducing the yield of the desired product **8**. The slow phosphitylation is almost certainly due to steric hindrance caused by the bulky *O*-dimethoxytrityl and *S*-trityl groups. The purified phosphoramidite **8** was stored at -20°C and dried thoroughly by evaporation of benzene before use in solid-phase synthesis in order to remove traces of moisture and triethylamine which would impair the coupling yield.

Oligodeoxyribonucleotide synthesis was carried out on an Applied Biosystems model 394 DNA/RNA synthesiser using a standard 1 μmol synthesis cycle and trityl-on end procedure, except that the coupling time for the modified uridine phosphoramidite was increased to 10 min. Standard ammonia deprotection was carried out and the 5′-*O*-dimethoxytrityl, *S*-trityl protected oligonucleotide was purified by reversed phase HPLC. After selective removal of the 5′-*O*-dimethoxytrityl group with aqueous acetic acid, the *S*-trityl protecting group was cleaved with aqueous silver nitrate solution under neutral conditions; subsequently the silver ions were removed with DTT and labelling of the free thiol was performed following a standard procedure[1].

In our initial investigations we attempted to use the *tert*-butyl- and cyano- groups to protect the mercaptan function however both methods tended to give irreproducible results and as a consequence we switched to the rather bulky yet readily cleavable trityl group. We envisage that this methodology will be particularly useful for the incorporation of photoactivatable groups such as phenyl azides in synthetic oligoribonucleotides for mapping interactions with RNA binding proteins.

Acknowledgement
The authors would like to thank Samantha O'Loughlin and Viviane Adam for oligonucleotide synthesis and purification.

References and Notes

1. Sinha, N.D.; Striepeke, S. *Oligonucleotides and Analogues, A Practical Approach;* Eckstein, F., Ed.; IRL Press: Oxford, **1991**; pp.185-208.
2. Goodchild, J. *Bioconjugate Chem.* **1990**, *1*, 165.
3. Connolly, B.A; Rider, P. *Nucleic Acids Res.* **1985**, *13*, 4485.
4. Sinha, N.D.; Cook, R.M. *Nucleic Acids Res.* **1988**, *16*, 2659.
5. Sproat, B.S.; Beijer, B.; Rider, P.; Neuner, P. *Nucleic Acids Res.* **1987**, *15*, 4837.
6. Bonfils, E.; Thuong, N.T. *Tetrahedron Lett.* **1991**, *32*, 3053.
7. Zuckerman, R.; Corey, D.; Schultz, P. *Nucleic Acids Res.* **1987**, *15*, 5305.
8. Gupta, K.C.; Sharma, P.; Kumar, P.; Sathyanarayana, S. *Nucleic Acids Res.* **1991**, *19*, 3109.
9. McBride, L.J.; Caruthers, M.H. *Tetrahedron Lett.* **1983**, *24*, 245.

10. Sinha, N.D.; Biernat, J.; McManus, J.; Koster, H. *Nucleic Acids Res.* **1984**, *12*, 4539.
11. Eckstein, F.; Gish, G. *TIBS*, **1989**, *14*, 97.
12. Zon, G.; Stec, W. *Oligonucleotides and Analogues, A Practical Approach;* Eckstein, F., Ed.; IRL Press: Oxford, 1991; pp.87-103.
13. Agrawal, S.; Zamecnik, P.C. *Nucleic Acids Res.* **1990**, *18*, 5419.
14. Blackburn, G.M.; Cohen, J.S.; Todd, A.R. *J. Chem. Soc.* **1966**, 239.
15. Fidanza, J.A.; McLaughin, L.W. *J. Org. Chem.* **1992**, *57*, 2340.
16. Bradley, D.H.; Hanna, M.M. *Tetrahedron Lett.* **1992**, *33*, 6223.
17. Goodwin, J.T.; Glick, G.D. *Tetrahedron Lett.* **1993**, *34*, 5549.
18. Sproat, B.S.; Beijer, B.; Iribarren, A. *Nucleic Acids Res.* **1990**, *18*, 47.
19. Keller, T.H.; Häner, R. *Helv. Chim. Acta*, **1993**, *76*, 884.
20. Manoharan, M.; Guinosso, C.J.; Cook, P.D. *Tetrahedron Lett.* **1991**, *32*, 7171.
21. Calzada, J.G.; Hooz, J. *Organic Synthesis, Collective Vol VI;* Noland, W.E., Ed.; Wiley **1988**; pp. 634-637.
22. Atkinson, T.; Smith, M. *Oligonucleotide Synthesis* , Gait, M.J., Ed.; IRL Press: Oxford; **1985**; pp. 35-81.
23. Compound **2**. ^{13}C NMR spectrum ($CDCl_3$) δ: 169.84 (C-4), 169.72 (acetyl CO), 154.56 (C-2), 144.40 (C-6 and dichlorophenyl C-1), 128.79 (dichlorophenyl C-2 and C-6), 128.53 (dichlorophenyl C-3 and C-5), 126.92 (dichlorophenyl C-4), 93.72 (C-5), 89.39 (C-1′), 82.06 (C-2′), 81.53 (C-4′), 67.54 (C-3′), 67.29 (acetyl CH_2), 60.53 (ethyl CH_2), 59.27 (C-5′), 17.23, 17.10, 16.74, 16.57 (isopropyl CHs), 13.28, 12.84, 12.73 and 12.10 ppm (isopropyl CH_3s).
TLC. Ascending mode on aluminium backed silica plates containing 254 nm fluor; Rf: 0.31 (3:1 v/v petroleum ether/ethyl acetate).
24. Compound **3**. ^{13}C NMR spectrum ($CDCl_3$) δ: 169.82 (acetyl CO), 163.50 (C-4), 149.84 (C-2), 139.63 (C-6), 101.54 (C-5), 88.82 (C-1′) , 82.48 (C-2′), 81.54 (C-4′), 68.45 (C-3′), 67.64 (acetyl CH_2), 60.81 (ethyl CH_2), 58.33 (C-5′), 18.51, 17.45, 17.37, 17.21, 16.96, 16.74 (isopropyl CHs), 14.15 (ethyl CH_3), 13.46, 13.10, 12.90 and 12.41 ppm (isopropyl CH_3s).
TLC. As above; Rf: 0.29 (2:1 v/v hexane/ethyl acetate).
25. Compound **4**. ^{13}C NMR spectrum ($CDCl_3$) δ: 163.45 (C-4), 150.31 (C-2), 139.13 (C-6), 101.76 (C-5), 89.49 (C-1′), 82.92 (C-2′), 81.81 (C-4′), 73.07 (OCH_2), 68.29 (C-3′), 61.64 (CH_2OH), 59.32 (C-5′), 17.46, 17.37, 17.28, 17.21, 17.05, 16.97, 16.80 (isopropyl CHs), 13.48, 13.08, 12.91 and 12.63 ppm (isopropyl CH_3s).
TLC. As above; Rf: 0.55 (1:1 v/v hexane/ethyl acetate).
26. Compound **5**. ^{13}C NMR spectrum ($CDCl_3$) δ: 164.08 (C-4), 149.81 (C-2), 144.75 (trityl C-1s), 139.47 (C-6), 129.48 (trityl C-3s and C-5s), 127.64 (trityl C-2s and C-6s), 126.38 (trityl C-4s), 101.28 (C-5), 88.83 (C-1′), 82.23 (C-2′), 81.44 (C-4′), 69.60 (C-3′), 68.02 (C-5′), 66.39 (trityl quaternary C), 59.31 (OCH_2), 31.92 (CH_2S), 17.38, 17.26, 17.09, 16.90, 16.73 (isopropyl CHs), 13.28, 12.96, 12.74 and 12.43 ppm (isopropyl CH_3s).
TLC. As above; Rf: 0.55 (2:1 v/v hexane/ethyl acetate).
27. Compound **6**. ^{13}C NMR spectrum ($CDCl_3$) δ: 164.01 (C-4), 150.21 (C-2), 144.33 (trityl C-1s), 141.34 (C-6), 128.08 (trityl C-3s and C-5s), 127.74 (trityl C-2s and C-6s), 126.38 (trityl C-4s), 101.69 (C-5), 89.00 (C-1′), 84.55 (C-4′), 81.38 (C-2′), 69.07 (C-3′), 68.27 (C-5′), 66.71 (trityl quaternary C), 60.55

(OCH_2) and 31.67 ppm (CH_2S).

TLC. As above; Rf: 0.45 (1:9 v/v ethanol/dichloromethane).

28. Compound **7**. TLC as above; Rf: 0.63 (1:19 v/v ethanol/dichloromethane containing 1% triethylamine).
29. Compound **8**. ^{31}P NMR spectrum ($CDCl_3$) δ: 146.60 and 146.39 ppm.

    TLC. As above; Rf: 0.73 (1:19 v/v ethanol/dichloromethane containing 1% triethylamine).

(*Received 24 September* 1993)