Labeling of DNA *via* rearrangement of *S*-2-aminoethyl phosphorothioates to *N*-(2-mercaptoethyl)phosphoramidates†

Menglin Chen and Kurt V. Gothelf*

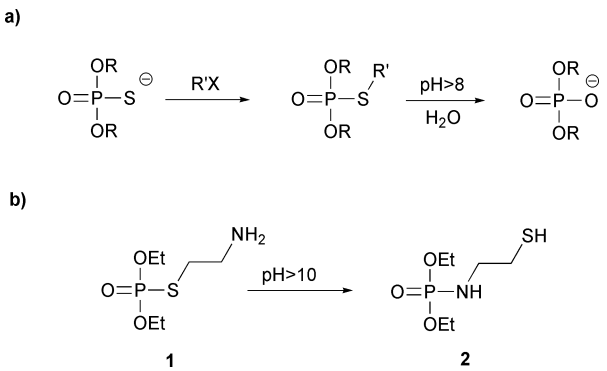
Received 14th November 2007, Accepted 18th December 2007
First published as an Advance Article on the web 24th January 2008
DOI: 10.1039/b717664c

The reaction of phosphorothioates in DNA with 2-bromoethylammonium bromide results in *S*-2-aminoethyl phosphorothioates, which can rearrange to *N*-(2-mercaptoethyl)phosphoramidates providing a facile method for the generation of site-specific thiol labeling of DNA sequences. The applicability of this method was demonstrated by conjugation of the thiolated DNA sequence with N^{α}-(3-maleimidylpropionyl) biocytin and Alexa Fluor 546 C5-maleimide.

Introduction

A number of strategies have been explored for the chemical functionalization of oligodeoxynucleotides with reactive functional groups. These have been applied to preparing DNA conjugates with fluorescent dyes, biochemical markers and linkers, peptides, nanoparticles and a variety of organic molecules.[1–4] In addition to amines, thiols have been extensively utilized for site-specific conjugation since they are excellent soft nucleophiles for aliphatic nucleophilic substitution or conjugate addition.[3] DNA has been labeled with thiols by both phosphoramidite[5] and *H*-phosphonate[6] reagents. Commercial reagents[7] are utilized for introducing a 3′- or 5′- thiol functionality to oligonucleotides *via* automated phosphoramidite oligonucleotide synthesis.[8,9] However, such methods are generally expensive and often place the conjugate group at some distance from the DNA sequence. Introduction of an internal thiol modification can be performed by using phosphoramidite nucleosides modified with a disulfide spacer at the nucleobase[10] or by post-synthetic modification of internally amine-modified oligonucleotides.[11–13] Alternatively, the thiol modification can be introduced at the phosphate backbone by the hydrogen phosphonate approach *via* a phosphoramidate analog.[14–17] The introduction not only allows the post-synthetic attachment of thiol-specific conjugate groups, but also provides modified DNA with better stability compared to the phosphorothioate triester DNA derivatives.[18]

Here we describe a new method for post-synthetically introducing a thiol modification at the DNA backbone by using a simple phosphorothioate oligonucleotide as the starting material. DNA sequences containing phosphorothioate diesters in specific positions are easily prepared during automated DNA synthesis by substitution of the I_2–pyridine oxidation reagents with a sulfurizing agent,[9] and it is used as an inexpensive standard modification in automated phosphoramidite oligonucleotide synthesis. Such internal phosphorothioate diesters provide nucleophilic sites available for reaction with appropriate alkylating agents as described by McLaughlin and Fidanza[19] (Scheme 1a). One major disadvantage of alkylated phosphorothioate triesters is lability towards basic hydrolysis. The labeled materials are relatively stable near neutral pH but undergo significant hydrolysis in solutions of increasing pH[20] (Scheme 1a). We have recently described that the reactivity of trialkyl phosphorothioates can be exploited for an intramolecular *S,N* rearrangement.[21] It was found that *S*-2-aminoethyl diethyl phosphorothioate **1** rapidly rearranges to diethyl *N*-(2-mercaptoethyl)phosphoramidate **2** (Scheme 1b). The reaction proceeds smoothly at room temperature under basic conditions.



Scheme 1 (a) Alkylation of phosphorothioate triesters and basic hydrolysis, R, R′ = alkyl group, X = Br, I. (b) Rearrangement of *S*-2-aminoethyl phosphorothioate **1** to *N*-(2-mercaptoethyl)phosphoramidate **2**.

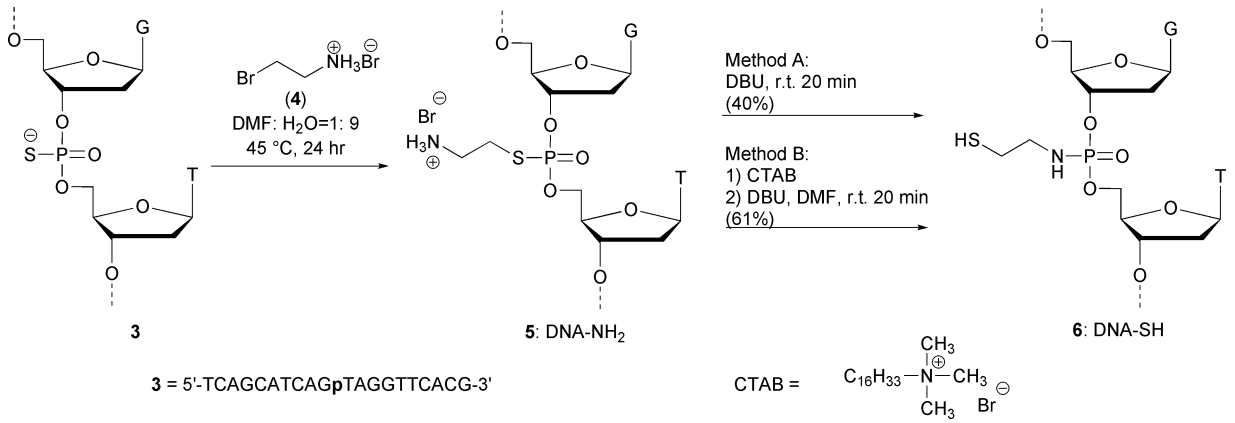
Results and discussion

The *S,N* rearrangement at phosphorus described above is applied here to oligonucleotides containing a phosphorothioate modification (Scheme 2). A 20-mer DNA sequence **3** with one phosphorothioate modification between the 10th and 11th bases (labeled as p), was alkylated to form the phosphorothioate triester DNA analog.

Alkylation of the phosphorothioate labeled DNA sequence **3** was tested with both 2-bromoethylammonium bromide (**4**) and 2-Fmoc-amino ethyl iodide (not shown). It turned out that alkylation with **4** was more efficient, which is probably due to the better solubility of **4** in H_2O and probably also because the

Centre for DNA Nanotechnology at iNANO and Department of Chemistry, University of Aarhus, Langelandsgade 140, DK-8000 Århus C, Denmark. E-mail: kvg@chem.au.dk; Fax: +45 8619 6199; Tel: +45 8942 3907

† Electronic Supplementary Information (ESI) available: HPLC chromatograms of compounds **5** and **6**, and MALDI-TOF mass spectra of compounds **5–9**. See http://dx.doi.org/10.1039/b717664c/



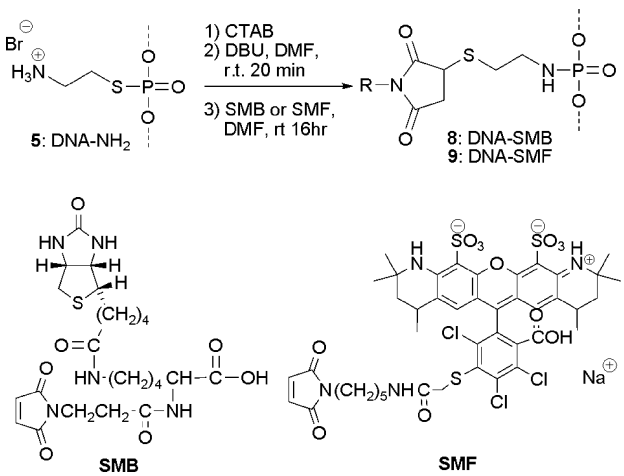
Scheme 2 Alkylation of phosphorothioate labeled oligonucleotide **3** and subsequent rearrangement into the corresponding *N*-(2-mercaptoethyl)-phosphoramidate **6**.

positive charge on **4** attracts it to the negatively charged DNA backbone. Alkylation of the DNA sequence with **4** in DMF–$H_2O = 1:9$, 20 mM HEPES, pH 6.5 was performed by incubation at 45 °C for 24 h to give a 78% yield of DNA–NH_2 **5** as verified by HPLC. Minor amounts of non-specifically alkylated DNA may be found in other fractions. However, an attempt to alkylate a DNA sequence without the phosphorothioate under similar conditions resulted in less than 10% of products. Since the phosphorothioate is chiral and the phosphorothioate-labeled DNA is a mixture of two diastereomers (non-distinguishable in HPLC), the alkylation also results in two diastereomers (*R*p and *S*p,) which in this case results in two different peaks in the HPLC chromatogram (Table 1). The faster and slower eluting diastereomers are denoted as diastereomers A and B, respectively. The identity of the alkylated product **5** was confirmed by MALDI-TOF MS (Table 1).

The rearrangement was carried out by incubating the oligonucleotides DNA–NH_2 **5** (diastereomers A and B) in 2 M DBU in H_2O. After 20 min **5** disappeared, and DNA–SH **6** was obtained in 40% yield. The presence of the thiol in the rearrangement product was verified by Ellman's test.[22] Due to the low hydrolysis stability of phosphorothioate triesters at elevated pH, it is not surprising that the hydrolysis product DNA–OH **7** (identical to the native DNA sequence) came out as the main by-product.

To favor the rearrangement over hydrolysis, the reaction was attempted in an organic solvent by preparing the cetyltrimethylammonium bromide (CTAB) salt of the DNA sequence.[23–25] After adding CTAB in an aqueous solution, the CTAB salt of DNA–NH_2 **5** was precipitated, dried thoroughly and redissolved in dry DMF in the presence of molecular sieves. After treating with DBU, the rearrangement product **6** was obtained in 61% yield.

To investigate the usefulness of this new three-step labeling procedure, it was applied to the conjugation of N^{α}-(3-maleimidylpropionyl) biocytin (**SMB**) and Alexa Fluor 546 C5-maleimide (**SMF**) to DNA (Scheme 3). The CTAB salt of DNA–NH_2 **5** (diastereomers A and B) was prepared and subsequently treated with 2 M DBU in DMF, at room temperature for 20 min. Then 50% acetic acid in DMF was added slowly to adjust the pH to 7.0–7.5. Then excess of **SMB** or **SMF** was added and the sample



Scheme 3 Labeling of oligonucleotide **5** with N^{α}-(3-maleimidylpropionyl) biocytin (**SMB**) or Alexa Fluor 546 C5-maleimide (**SMF**).

Table 1 MALDI-TOF MS data and HPLC retention time for modified oligonucleotides

Oligonucleotide	Mass calculated	Mass measured	HPLC retention time/min
3	6132.0	6132.0	10.7[a]
DNA–NH_2 **5**	6176.1	6175.7	10.2[a] (A), 10.4[a] (B)
DNA–SH **6**	6176.1	6176.1	11.3[a] (A), 11.7[a] (B)
DNA–SMB **8**	6699.7	6699.9	12.6[a] (A), 13.7[a] (B)
DNA–SMF **9**	7184.2	7183.4	13.0[b] (A), 13.4[b] (B)
DNA–OH **7**	6116.0	6115.2	10.3[a], 9.7[b]

[a] Method a: HPLC linear gradient using 10–12% acetonitrile–TEAA 100 mM, pH 7.0 (16 min). [b] Method b: HPLC linear gradient using 0–65% acetonitrile–TEAA 100 mM, pH 7.0 (20 min).

was incubated at room temperature overnight. The crude labeled products DNA–SMB **8** and DNA–SMF **9** were purified by HPLC and identified by MALDI-TOF MS (Table 1 and Fig. 1). Both of the labeled products were obtained as mixtures of diastereomers A and B with 60% yield. The ratio of the two diastereomers is surprisingly 9 : 1. However, when labeling another oligonucleotide sequence in the same manner, a 1 : 1 ratio of the two diastereomers was obtained. Thus the ratio between the two diastereomers shows that the stereoselectivity of the rearrangement is highly dependent on the secondary structure of the oligonucleotide.

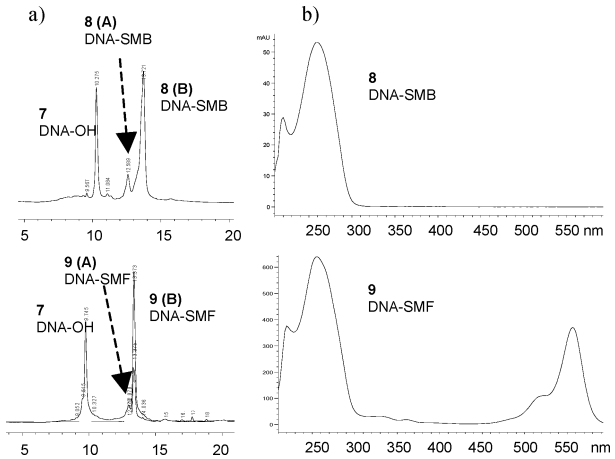


Fig. 1 (a) HPLC chromatogram of the reaction leading to DNA–SMB **8** (260 nm) and to DNA–SMF **9** (260 and 550 nm), (b) UV–vis spectra of DNA–SMB **8** and DNA–SMF **9**.

All the resulting phosphoramidate DNA conjugates are stable under slightly acidic and under basic conditions. As reported, hydrolysis of the phosphoramidate linkage to the phosphate diester linkage requires harsh conditions such as 85% formic acid at 95 °C.[26] Thus, these phosphoramidate conjugates are significantly more stable than phosphorothioate conjugates.

The introduction of the modifications at the phosphate backbone described above may have structural implications and could possibly alter the stability of the DNA duplex. Therefore melting temperatures (T_m) of the labeled phosphoramidate oligonucleotides **8** and **9** (diastereomer mixtures) were measured to investigate their thermal stability (Table 2). Gratifyingly, the labeling of a single phosphorothioate diester according to this new rearrangement method did not alter dramatically the stability and specificity of a 20 bp duplex. Attachment of **SMB** to the phosphorus residue has less influence on the melting temperature ($\Delta T_m \sim 1$ °C) than labeling with **SMF** ($\Delta T_m \sim 2$–3 °C). This may be due to the smaller size and the longer linker of **SMB** compared to **SMF**. The ΔT_m between the fully matched and the A–G mismatched duplexes is 5 °C for unlabelled **3** and 4–5 °C for **8** and **9**.

Table 2 Melting temperatures for labeled oligonucleotides

Complex[a]	T_m/° C	ΔT_m/°C
Unlabeled DNA **3** and **c3**	66.1	—
SMB labeled DNA **8** and **c3**	65.3	−0.8[b]
SMF labeled DNA **9** and **c3**	63.1	−3.0[b]
Unlabeled DNA **3** and **m3**	61.1	−5.0[c]
SMB labeled DNA **8** and **m3**	60.0	−5.3[c]
SMF labeled DNA **9** and **m3**	59.0	−4.1[c]

[a] **c3** (5′-CGTGAACCTACTGATGCTGA) is the sequence fully complementary to **3**, while **m3** (5′-CGTGAACCTAATGATGCTGA) contains a single-base A–G base mismatch. [b] Compared to T_m of the unlabeled DNA **3** and **c3**. [c] Compared to T_m of the same sequence paired with **c3**.

Conclusions

In conclusion, we have developed a new strategy for the labeling of DNA *via* an inexpensive phosphorothioate modification. This procedure involved alkylation of the phosphorothioate with 2-bromoethylammonium bromide (**4**). By adding a base, the phosphorothioate rearranged to the phosphoramidate DNA derivative containing a free thiol functional group. The rearrangement is accompanied by some degree of hydrolysis, however by performing the reaction in an organic solvent a yield of 61% was obtained. The usefulness of this strategy for conjugation of important functionalities to DNA was demonstrated by the successful labeling with biocytin and a fluorophore.

Experimental

HPLC was performed on a Hewlett Packard Agilent instrument with an autosampler and fraction collector on an XTerra C18 column (Waters #186000602) using acetonitrile–TEAA 100 mM, pH 7.0 mixtures as eluent. Oligonucleotides were analyzed by MALDI-TOF mass spectrometry on a Bruker AutoFlex instrument in a 3-hydroxypicolinic acid (HPA)–ammonium citrate matrix using the negative ion reflector mode on an AnchorPlate target. Oligonucleotides in the aqueous solution were precipitated by NH_4OAc–EtOH according to the precipitation protocol of Maniatis *et al.*[27] T_m values were obtained in 10 mM sodium phosphate (pH 7.0) and 500 mM sodium chloride at duplex concentrations of 1.0 μM. Absorbance values were measured with a Cary 100 Bio UV–visible spectrophotometer equipped with a Cary temperature controller attached to a Cary 1 thermostattable multicell block. The solution temperatures were measured directly with a thermistor probe (Cary thermometer probe series II). Absorbance and temperature data were collected after analog to digital conversion (Cary thermal analysis software[28]). T_m values were determined from first- and second-order derivatives of the absorbance *vs.* temperature plots.

DNA–NH_2 (5)

2-Bromoethylammonium bromide (**4**) (0.4 mg, 0.02 mmol) in 2 μL DMF was added to DNA **3** (10 nmol) in 18 μL H_2O, 20 mM HEPES, pH 6.5 in an Eppendorf tube at 45 °C and the reaction was incubated for 24 h. After ethanol precipitation, DNA–NH_2 **5** was purified in 78% yield (48.1 μg) by reversed-phase HPLC [linear gradient using 10–12% acetonitrile–TEAA 100 mM, pH 7.0 (16 min), XTerra C18 column, 0.5 mL min^{-1}, 260 nm]. MALDI-TOF MS: calcd for **5**, 6176.1 $[M - H]^-$; found, 6175.7 $[M - H]^-$.

DNA–SH (6)

Method A. A solution of 2 M DBU in 20 μL H_2O was added to dry DNA–NH_2 **5** (5 nmol) in an Eppendorf tube and incubated at room temperature for 20 min. Then it was precipitated with

ethanol and redissolved in H_2O. The thiol-tethered DNA–SH **6** was purified in 40% yield (12.3 μg) by reversed-phase HPLC [linear gradient using 10–12% acetonitrile–TEAA 100 mM, pH 7.0 (16 min), XTerra C18 column, 0.5 mL min^{-1}, 260 nm]. MALDI-TOF MS: calcd for **6**, 6176.1 $[M - H]^-$; found, 6176.1 $[M - H]^-$.

Method B. To DNA–NH_2 **5** (5 nmol in 10 μL H_2O) was added 5 μL 8% CTAB aqueous solution, and the supernatant removed. The resulting CTAB salt of DNA–NH_2 **5** was thoroughly dried and treated with 2 M DBU in dry DMF (10 μL) at room temperature in the presence of 4 Å molecular sieves for 20 min. After $LiClO_4$ precipitation, it was redissolved in H_2O and precipitated again in ethanol. The thiol-tethered DNA–SH **6** was purified in 61% yield (19.2 μg) by reversed-phase HPLC [linear gradient using 10–12% acetonitrile–TEAA 100 mM, pH 7.0 (16 min), XTerra C18 column, 0.5 mL min^{-1}, 260 nm]. MALDI-TOF MS: calcd for **6**, 6176.1 $[M - H]^-$; found, 6176.1 $[M - H]^-$.

Biocytin labeled DNA–SMB (8)

The CTAB salt of DNA–NH_2 **5** (2 nmol) was thoroughly dried and dissolved in a solution of 2 M DBU in dry DMF (10 μL) at room temperature in the presence of 4 Å molecular sieves. After 20 min, 50% acetic acid was added slowly to the reaction solution to adjust the pH to 7.0–7.5. The N^{α}-(3-maleimidylpropionyl) biocytin, **SMB** (1.0 mg, 0.02 mmol) in DMF (1 μL) was added to the reaction solution which was incubated at room temperature overnight. After precipitation with $LiClO_4$ it was redissolved in H_2O and precipitated again in ethanol. Finally it was redissolved in H_2O, TEAA 100 mM, pH 7.0, and purified in 60% yield (8.0 μg) by reversed-phase HPLC [linear gradient using 10–12% acetonitrile–TEAA 100 mM, pH 7.0 (16 min), XTerra C18 column, 0.5 mL min^{-1}, 260 nm]. MALDI-TOF MS: calcd for DNA–SMB **8**, 6699.7 $[M - H]^-$; found, 6699.9 $[M - H]^-$.

Alexa Fluor 546 labeled DNA–SMF (9)

The CTAB salt of DNA–NH_2 **5** (2 nmol) was thoroughly dried and dissolved in a solution of 2 M DBU in dry DMF (10 μL) at room temperature in the presence of 4 Å molecular sieves. After 20 min, 50% acetic acid was added slowly to the reaction solution to adjust the pH to 7.0–7.5. Then Alexa Fluor 546 C5 maleimide, **SMF** (2.0 mg, 0.02 mmol) in DMF was added to the reaction solution which was incubated at room temperature overnight. After precipitation with $LiClO_4$ it was redissolved in H_2O and precipitated again in ethanol. Finally it was redissolved in H_2O, TEAA 100 mM, pH 7.0, and purified in 60% yield (8.6 μg) by reversed-phase HPLC [linear gradient using 0–65% acetonitrile–TEAA 100 mM, pH 7.0 (20 min), XTerra C18 column, 0.5 mL min^{-1}, 260 nm]. MALDI-TOF MS: calcd for DNA–SMF **9**, 7184.2 $[M - H]^-$; found, 7183.4 $[M - H]^-$.

Acknowledgements

This study was funded in part by iNANO through the Danish Research Councils, the Carlsberg Foundation, and the Danish National Research Foundation.

Notes and references

1 J. Goodchild, *Bioconjugate Chem.*, 1990, **1**, 165.
2 C. H. Tung and S. Stein, *Bioconjugate Chem.*, 2000, **11**, 605.
3 P. Virta, J. Katajisto, T. Niittymaki and H. Lonnberg, *Tetrahedron*, 2003, **59**, 5137.
4 N. Venkatesan, S. J. Kim and B. H. Kim, *Curr. Med. Chem.*, 2003, **10**, 1973.
5 D. A. Stetsenko and M. J. Gait, *J. Org. Chem.*, 2000, **65**, 4900.
6 N. D. Sinha and R. M. Cook, *Nucleic Acids Res.*, 1988, **16**, 2659.
7 Commercially available from the company of Glen Research, Sterling, Virginia, USA (www.glenres.com).
8 M. D. Matteucci and M. H. Caruthers, *J. Am. Chem. Soc.*, 1981, **103**, 3185.
9 W. J. Stec, G. Zon, W. Egan and B. Stec, *J. Am. Chem. Soc.*, 1984, **106**, 6077.
10 H. A. Held, A. Roychowdhury and S. A. Benner, *Nucleosides, Nucleotides Nucleic Acids*, 2003, **22**, 391.
11 G. T. Hermanson, *Bioconjugate Techniques*, Academic Press, San Diego, 1996, pp. 652.
12 C. J. Loweth, W. B. Caldwell, X. G. Peng, A. P. Alivisatos and P. G. Schultz, *Angew. Chem., Int. Ed.*, 1999, **38**, 1808.
13 J. W. Hojfeldt, P. Blakskjar and K. V. Gothelf, *J. Org. Chem.*, 2006, **71**, 9556.
14 J. A. Fidanza and L. W. McLaughlin, *J. Org. Chem.*, 1992, **57**, 2340.
15 M. Endo and T. Majima, *J. Am. Chem. Soc.*, 2003, **125**, 13654.
16 M. Endo, T. Shiroyama, M. Fujitsuka and T. Majima, *J. Org. Chem.*, 2005, **70**, 7468.
17 E. Privat and U. Asseline, *Bioconjugate Chem.*, 2001, **12**, 757.
18 S. Peyrottes, J. J. Vasseur, J. L. Imbach and B. Rayner, *Nucleic Acids Res.*, 1996, **24**, 1841.
19 J. A. Fidanza and L. W. McLaughlin, *J. Am. Chem. Soc.*, 1989, **111**, 9117.
20 J. A. Fidanza, H. Ozaki and L. W. McLaughlin, *J. Am. Chem. Soc.*, 1992, **114**, 5509.
21 M. Chen, A. Maetzke, S. J. K. Jensen and K. V. Gothelf, *Eur. J. Org. Chem.*, 2007, **35**, 5826.
22 G. L. Ellman, *Arch. Biochem. Biophys.*, 1959, **82**, 70.
23 A. S. Boutorine, D. Brault, M. Takasugi, O. Delgado and C. Hélène, *J. Am. Chem. Soc.*, 1996, **118**, 9469.
24 A. S. Boutorine, D. Trung Le, J. P. Battioni, D. Mansuy, D. Dupré and C. Hélène, *Bioconjugate Chem.*, 1990, **1**, 350.
25 D. M. McLoughlin, J. O'Brien, J. J. McManus, A. V. Gorelov and K. A. Dawson, *Bioseparation*, 2000, **9**, 307.
26 B. C. Froehler, *Tetrahedron Lett.*, 1986, **27**, 5575.
27 T. Maniatis, J. Sambrook and E. F. Fritsch, *Molecular Cloning—A Laboratory Manual,* Cold Spring Harbor Press, Cold Spring Harbor, New York, 1982, pp. 76–85.
28 The Cary Win UV Bio Package (version 3.0 (182)), Varian Inc., Palo Alto, CA, USA, 2002.