# Single base interrogation by a fluorescent nucleotide: each of the four DNA bases identified by fluorescence spectroscopy†

**Pavol Cekan and Snorri Th. Sigurdsson***

*Received (in Cambridge, UK) 1st February 2008, Accepted 15th May 2008*
*First published as an Advance Article on the web 5th June 2008*
**DOI: 10.1039/b801833b**

**Nucleoside Ç, which contains a rigid nitroxide spin label, is effectively reduced in DNA by sodium sulfide to the corresponding amine, yielding a fluorescent probe ($Ç^f$) that can report the identity of its base-pairing partner in duplex DNA.**

Single nucleotide polymorphisms (SNPs) are single nucleotide gene variations.[1] SNPs in a protein-encoding region can result in a mutant protein and subsequently predispose human beings to common diseases and disorders.[2,3] Thus, disease diagnosis and identification of targets for drug discovery by SNP detection is an active area of research. Most methods that have been developed for detection of SNPs rely on monitoring hybridization of an oligonucleotide probe to the sequence of interest; the probe binds to the wild-type sequence but not to DNAs containing an SNP.[4] Hybridization is usually detected by change in the fluorescence emission of a fluorophore that is linked to the oligonucleotide probe. Hybridization assays require a careful selection of the sequence of the oligonucleotide probe and the annealing conditions, especially the temperature. It would be advantageous if, instead of monitoring hybridization, fluorescent DNA probes that signal single-base mismatches within duplexes could be used.

A promising class of compounds for SNP detection are modified nucleosides having fluorophores that are sensitive to their proximal environment.[5–7] One strategy for the design of fluorescent nucleotides for SNP detection is conjugation of a fluorophore to the nucleobase through a linker.[8–11] For example, Saito and coworkers have conjugated pyrene to pyrimidine nucleosides and called them "base-discriminating fluorescent nucleosides" (BDFs) because they can detect and distinguish between mismatches. A different strategy is to extend the aromatic ring system of the bases in natural nucleosides and use the fluorescence of the base itself as a reporter group.[8,12–16] However, most of the fluorescent nucleobase analogs that have been reported to date are limited in their ability to differentiate between all the possible base-pairing partners.[9–12,14,17] In this communication we report a nucleoside containing a fluorescent base that is able to discriminate between the four bases of DNA.

We have recently described nucleoside **Ç** ("C-spin"), an analog of cytidine that contains a rigid spin label fused to the nucleobase, for the study of nucleic acid structure and dynamics by EPR spectroscopy (Fig. 1).[18] **Ç** is structurally related to the fluorescent nucleosides tC and $tC^O$.[19–21] Reduction of the nitroxide to the corresponding hydroxyl amine with DTT yielded a fluorescent nucleoside which made nucleoside **Ç** an attractive spectroscopic probe, because it can be used for biophysical studies of nucleic acids by both fluorescence and EPR spectroscopies. However, when studying the fluorescent properties of oligomers containing this modification we found that the hydroxyl amine was readily oxidized back to the nitroxide upon exposure to oxygen, thereby quenching the fluorescence. In the search for a more stable fluorescent analog of **Ç**, we sought to reduce the nitroxide to the corresponding amine.

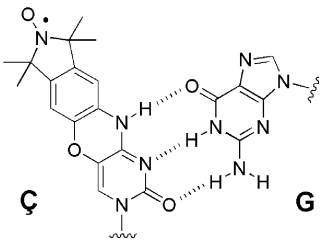
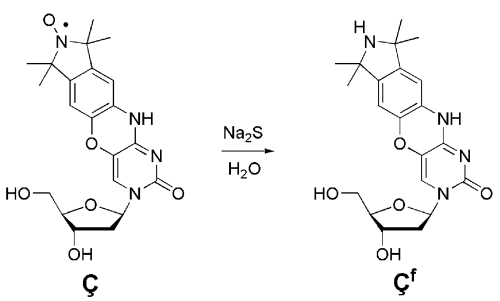



**Fig. 1** Structure and base pair properties of **Ç**.

Nitroxides have been reduced to amines by sodium sulfide in DMSO and DMF.[22] Limited solubility of nucleic acids in polar aprotic solvents prompted us to investigate this reaction in $H_2O$. Reduction of nucleoside **Ç** with $Na_2S$ in water at 45 °C did indeed proceed to give the highly fluorescent amine $Ç^f$ ($\Phi_F$ = 0.31) after 8 h and in nearly quantitative yield (Scheme 1).

To study the reduction of **Ç** in DNA, an aqueous solution of the oligodeoxyribonucleotide 5′-d(GACCTCG**Ç**ATCGTG) was treated with $Na_2S$ at 45 °C for 14 h. Analysis of the crude reaction mixture by EPR spectroscopy revealed the disappearance of the nitroxide (data not shown). Furthermore,




**Scheme 1** Reduction of spin-labeled nucleoside **Ç**.

*University of Iceland, Science Institute, Dunhaga 3, 107 Reykjavik, Iceland. E-mail: snorrisi@hi.is; Fax: +354 5528911; Tel: +354 5254801*

† Electronic supplementary information (ESI) available: Synthetic protocols. Preparation, purification and characterization of DNA oligomers. See DOI: 10.1039/b801833b

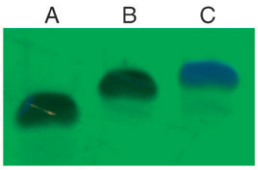



**Fig. 2** Analysis of DNA oligomers by DPAGE. **A**. 5′-d(GACCTCGCATCGTG), **B**. 5′-d(GACCTCGÇATCGTG) and **C**. reduced 5′-d(GACCTCGÇ$^{f}$ATCGTG).

denaturing polyacrylamide gel electrophoresis (DPAGE) showed that the $Na_2S$-treated DNA had become highly fluorescent (Fig. 2).

HPLC analysis of an enzymatically digested reduced oligomer revealed five peaks (Fig. S1†), corresponding to the nucleosides dC, dI (from enzymatic hydrolysis of dA), dG, dT and Ç$^{f}$, showing that the spin-labeled nucleotide was reduced by $Na_2S$ without affecting the other nucleotides. This result was also verified by MS analysis of the reduced DNA (Fig. S2†). Both the melting temperature ($T_M$) (Table 1) and CD spectrum (Fig. S3†) of a DNA duplex containing Ç$^{f}$ (Ç$^{f}$•**G**) are nearly identical to those of an unlabeled DNA duplex, indicating that the modification has a negligible effect on DNA duplex stability and conformation. CD spectra of the duplexes containing a mismatch were also consistent with B-form DNA (Fig. S3†).

To determine if Ç$^{f}$ could be used as a fluorescent probe for detecting single-base mismatches, four duplexes containing Ç$^{f}$ were synthesized (Ç$^{f}$•**G**, Ç$^{f}$•**A**, Ç$^{f}$•**C** and Ç$^{f}$•**T**, Table 1), each of which contained a different base opposite Ç$^{f}$ (G, A, C or T). The fluorescent spectra of the Ç$^{f}$-containing DNAs were recorded in an aqueous solution at 20 °C using excitation at 365.5 nm. The fluorescence of the four duplexes shows considerable variation, to the extent that they can all be clearly distinguished from each other (Fig. 3A). The duplex containing the Ç$^{f}$•**A** mismatch is the most fluorescent of the duplexes and is similar to that of the single strand. The fluorescence emission of the Ç$^{f}$•**C** and the Ç$^{f}$•**T** duplexes is between that of

**Table 1** Melting temperatures and photophysical properties of DNA 14-mer duplexes containing Ç$^{f}$[a]

$^{5}$G A C C T C G **M** A T C G T G$^{3'}$
$_{3'}$C T G G A G C **N** T A G C A C$_{5'}$

| Duplex | M | N | $T_M$[b]/°C | $\Delta T_M$[c]/°C | $\Phi_F$[d] | $\lambda_{em}$[e]/nm |
|---|---|---|---|---|---|---|
| **C•G** | C | G | 62.8 | | | |
| Ç$^{f}$•**G** | Ç$^{f}$ | G | 62.5 | −0.3 | 0.086 | 450 |
| **C•A** | C | A | 50.7 | | | |
| Ç$^{f}$•**A** | Ç$^{f}$ | A | 53.6 | +2.9 | 0.302 | 425, 449, 477 |
| **C•C** | C | C | 45.7 | | | |
| Ç$^{f}$•**C** | Ç$^{f}$ | C | 45.2 | −0.5 | 0.222 | 443 |
| **C•T** | C | T | 47.7 | | | |
| Ç$^{f}$•**T** | Ç$^{f}$ | T | 50.6 | +2.9 | 0.118 | 445 |

[a] 2 μM duplex in 10 mM sodium phosphate, 100 mM NaCl, 0.1 mM $Na_2EDTA$, pH 7.0. [b] The $T_M$ values are accurate within ±0.5 °C. [c] $\Delta T_M$ is the difference in the $T_M$ value between a duplex where either a **C** or Ç$^{f}$ is paired with the same nucleotide. [d] The fluorescence quantum yields ($\Phi_F$) were calculated[23] using anthracene as a standard[24] and the values are accurate within ±0.015. [e] Fluorescence emission maxima.

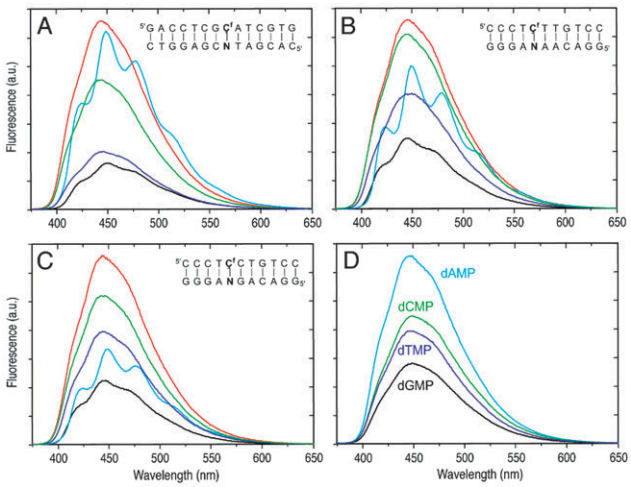



**Fig. 3** **A**–**C**. Fluorescence spectra of DNA oligomers; single-stranded (red), Ç$^{f}$•**A** (light blue), Ç$^{f}$•**C** (green), Ç$^{f}$•**T** (blue) and Ç$^{f}$•**G** (black) pairs in oligomers containing the flanking sequences 5′-d(-GÇ$^{f}$A-) (**A**), 5′-d(-TÇ$^{f}$T-) (**B**), and 5′-d(-TÇ$^{f}$C-) (**C**). **D**. Fluorescence spectra of Ç$^{f}$ in a 100 mM solution of either dAMP, dCMP, dTMP or dGMP.

Ç$^{f}$•**A** and the fully base-paired duplex (Ç$^{f}$•**G**). Although there is not a large difference in the fluorescence intensity between Ç$^{f}$•**T** and Ç$^{f}$•**G** at the higher wavelengths, Ç$^{f}$•**T** has a considerably higher emission at lower wavelengths (395–410 nm). As a result, the color emitted for Ç$^{f}$•**T** is light blue and green–blue for Ç$^{f}$•**G**.

To determine if the mismatch detection was dependent on the flanking sequence, we tested two other sequences (Fig. 3B and C). In both cases, the mismatches can be clearly distinguished from each other and from the fully base-paired sequence. Interestingly, the fluorescence intensity increases in the same order for Ç$^{f}$•**G**, Ç$^{f}$•**T** and Ç$^{f}$•**C** for all three flanking sequences (Fig. 3A–C). It is also noteworthy that there is a much larger difference between the fluorescence of Ç$^{f}$•**G** and Ç$^{f}$•**T** in the oligomers containing the pyrimidine flanking sequences (Fig. 3B and C). In contrast to the order observed for the Ç$^{f}$•**G**, Ç$^{f}$•**T** and Ç$^{f}$•**C** pairs, the relative fluorescence for Ç$^{f}$•**A** was different for all three oligomers. However, the Ç$^{f}$•**A** exhibits another striking difference in comparison to the other pairs, namely, 3–4 maxima are observed in the fluorescence spectra for all the oligomers. This unusual pattern does not appear to be flanking sequence-dependent. The reasons for this phenomenon, which is not observed for the other Ç$^{f}$-pairs, except to a much smaller degree for the other purine (Ç$^{f}$•**G**), are not clear at this time.

One possible explanation for the mismatch-sensing property of Ç$^{f}$ is the ability of the nucleobases to which it is paired to quench the fluorescence of Ç$^{f}$ to different extents. The fluorescence spectra of Ç$^{f}$ in 100 mM solutions of the nucleotides are shown in Fig. 3D, showing that the four spectra do not overlap. Furthermore, the fluorescence data for Ç$^{f}$•**G**, Ç$^{f}$•**T** and Ç$^{f}$•**C** in all the oligomers and for Ç$^{f}$ in the presence of dGMP, dTMP and dCMP show the same trend. This is an indication that the mismatch-sensing of Ç$^{f}$ originates in its fluorescence quenching by T, C and G. Further experiments are required to explain the fluorescence of Ç$^{f}$•**A**, relative to the


This journal is © The Royal Society of Chemistry 2008

other pairs, and the unusual shape of the $\mathbf{Ç^f}$•**A** fluorescence spectra.

In conclusion, we have shown that nucleoside $\mathbf{Ç^f}$ can distinguish between the four bases of DNA when incorporated into a DNA oligomer and hybridized to a target strand. This nucleoside, like other BDFs, can potentially be used in efficient homogenous assays for SNP detection that do not require enzymes or time-consuming steps, while avoiding hybridization errors. It is noteworthy that $\mathbf{Ç^f}$ can easily distinguish between the four bases of DNA with significant fluorescence intensity even when placed in a sequence adjacent to a **G**•**C** base-pair, which has been reported to substantially quench the fluorescence of other BDFs containing fluorescent bases.[12,13,17] Thus, it appears that the phenoxazine moiety, which **Ç** is based on, is a useful scaffold that could be used for synthesis of other fluorescent BDF derivatives. For example, BDFs that emit at higher wavelengths would enable the use of instruments that are currently being used in functional DNA technology for SNP analysis. Research along those lines, along with a more detailed structural and functional analysis of $\mathbf{Ç^f}$ will be reported in due course.

This work was supported by The Icelandic Research Fund (60028021) and by a doctoral fellowship to P. Cekan from the Eimskip Fund of the University of Iceland. We thank members of the Sigurdsson research group for helpful discussions.

## Notes and references

1. K. Nakatani, *ChemBioChem*, 2004, **5**, 1623–1633.
2. Y. Suh and J. Vijg, *Mutat. Res.-Fundam. Mol. Mech. Mutagen.*, 2005, **573**, 41–53.
3. M. Olivier, *Physiol. Genomics*, 2004, **16**, 182–183.
4. R. M. Twyman, *Curr. Top. Med. Chem.*, 2004, **4**, 1421–1429.
5. N. J. Greco and Y. Tor, *J. Am. Chem. Soc.*, 2005, **127**, 10784–10785.
6. J. N. Wilson and E. T. Kool, *Org. Biomol. Chem.*, 2006, **4**, 4265–4274.
7. U. Asseline, *Curr. Org. Chem.*, 2006, **10**, 491–518.
8. A. Okamoto, Y. Saito and I. Saito, *J. Photochem. Photobiol., C*, 2005, **6**, 108–122.
9. G. T. Hwang, Y. J. Seo and B. H. Kim, *J. Am. Chem. Soc.*, 2004, **126**, 6528–6529.
10. D. J. Hurley, S. E. Seaman, J. C. Mazura and Y. Tor, *Org. Lett.*, 2002, **4**, 2305–2308.
11. S. S. Bag, Y. Saito, K. Hanawa, S. Kodate, I. Suzuka and I. Saito, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 6338–6341.
12. A. Okamoto, K. Tainaka and I. Saito, *J. Am. Chem. Soc.*, 2003, **125**, 4972–4973.
13. A. Okamoto, K. Tanaka, T. Fukuta and I. Saito, *J. Am. Chem. Soc.*, 2003, **125**, 9296–9297.
14. K. Miyata, R. Tamamushi, A. Ohkubo, H. Taguchi, K. Seio, T. Santa and M. Sekine, *Org. Lett.*, 2006, **8**, 1545–1548.
15. C. Liu and C. T. Martin, *J. Mol. Biol.*, 2001, **308**, 465–475.
16. R. H. E. Hudson and A. Ghorbani-Choghamarani, *Synlett*, 2007, 870–873.
17. A. Okamoto, K. Kanatani and I. Saito, *J. Am. Chem. Soc.*, 2004, **126**, 4820–4827.
18. N. Barhate, P. Cekan, A. P. Massey and S. T. Sigurdsson, *Angew. Chem., Int. Ed.*, 2007, **46**, 2655–2658.
19. K.-Y. Lin, R. J. Jones and M. Matteucci, *J. Am. Chem. Soc.*, 1995, **117**, 3873–3874.
20. P. Sandin, L. M. Wilhelmsson, P. Lincoln, V. E. Powers, T. Brown and B. Albinsson, *Nucleic Acids Res.*, 2005, **33**, 5019–5025.
21. P. Sandin, K. Borjesson, H. Li, J. Martensson, T. Brown, L. M. Wilhelmsson and B. Albinsson, *Nucleic Acids Res.*, 2008, **36**, 157–167.
22. N. Kornblum and H. W. Pinnick, *J. Org. Chem.*, 1972, **37**, 2050–2051.
23. S. Fery-Forgues and D. Lavabre, *J. Chem. Educ.*, 1999, **76**, 1260–1264.
24. H. Du, R.-C. A. Fuh, J. Li, L. A. Corkan and J. S. Lindsey, *Photochem. Photobiol.*, 1998, **68**, 141–142.