Large flanking sequence effects in single nucleotide mismatch detection using fluorescent nucleoside $\text{C}^\text{f}$

Haraldur Gardarsson, Snorri Th. Sigurdsson *

University of Iceland, Science Institute, Dunhagi 3, 107 Reykjavik, Iceland

**ABSTRACT**

The first systematic study of flanking sequence effects on mismatch detection by a fluorescent nucleotide is described, using fluoroside $\text{C}^\text{f}$. Although a high degree of variance was observed in fluorescence intensity of mismatched duplexes between different flanking sequences, $\text{C}^\text{f}$ was able to distinguish a mismatch from the fully base-paired duplex in 13 out of 16 sequences, and even identify each mismatch in 10 of those flanking sequences. For the flanking sequences where fluoroside $\text{C}^\text{f}$ did not unambiguously determine its base-pairing partner, the experimental conditions were varied in an attempt to facilitate mismatch identification. No beneficial effect on the relative fluorescence intensities was achieved by changing the temperature, adding organic co-solvents or potassium iodide. In contrast, mercuric ions selectively quenched the fluorescence intensity of $\text{C}^\text{f}$ mismatch, effectively resolving the overlap of all emission spectra and thereby facilitating identification of all base-pairing partners in any flanking sequence by $\text{C}^\text{f}$. This is the first time mercuric ions have been used to selectively quench the fluorescence of a single mismatch. A noticeable characteristic of $\text{C}^\text{f}$ is that, unlike most fluorosides, the fluorescence intensity of $\text{C}^\text{f}$ was not quenched to a discernable degree by a flanking G–C pair.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Single nucleotide polymorphisms (SNPs) are single nucleotide mutations and the most abundant type of genetic variation in the human genome. Single SNPs have been linked to various diseases and disorders. For example, SNPs positioned in the 5′ flanking region or the regulatory intronic region have been reported to cause allele-specific effects on the expression of genes. The identification of SNPs, as disease-associated variants, is therefore an active area of research.

Various methods exist to detect SNPs, most of which rely on monitoring the hybridization of an oligonucleotide probe to a target sequence of interest. The probe binds selectively to the wild-type sequence and not to sequences containing a SNP site. Most commonly the hybridization is monitored by a fluorescent probe which is attached to the oligonucleotide probe. Hybridization assays require very specific conditions, that is, careful selection of the probe sequence and more importantly, the temperature for annealing must be carefully controlled. A fluorescent probe that could directly identify its base-pairing partner without the cumbersome hybridization process would therefore be advantageous.

In order to be able to use such a fluorescent probe, several aspects must be taken into consideration. First, the fluorescent probe must have appreciable fluorescence brightness for general use. Second, the excitation wavelength of the fluorescent probe must be different from where DNA absorbs. Third, it is preferable that the probe has a high Stokes shift (>100 nm) and that the fluorescence maximum is in the visible or near-infrared region, to facilitate detection and interpretation.

Our group has previously described the fluorescent nucleoside $\text{C}^\text{f}$ (Fig. 1), which, after incorporation into duplex DNA, is able to discriminate between fully base-paired and mismatch-containing duplexes. Such nucleosides have been called 'base-discriminating fluorescent nucleosides' (BDFs). However, BDFs have limitations, in that they can usually only distinguish one mismatch from the fully base-paired duplex. Furthermore many fluorescent probes experience quenching when incorporated into DNA, which can be severe when the fluorescent nucleoside is flanked by a G–C pair. Other methods of detecting a SNP site are for example, forced intercalation probes (FIT-probes) reported by Seitz et al. and internal charge transfer transfer probes (ICT-probes) reported by Wagenknecht, both of which can detect mismatches, but are unable to identify the mismatch in question.

Upon further examination it became evident that flanking sequence effects had an effect on the mismatch detection. Here we report a systematic study of the effects of all the flanking sequences on mismatch detection using fluorescent nucleoside $\text{C}^\text{f}$. We found a large variation in the fluorescence intensities of the mismatches.
between different flanking sequences. However, it was still possible to identify the base-pairing partner of $\text{Cf}^\prime$ in the majority of the sequences. In the sequences in which the fluoroside was unable to discriminate between the mismatches, we were able to alter the experimental conditions so that $\text{Cf}^\prime$ could identify any base-pairing partner in any sequence. Mercuric ions were especially useful as they enabled selective quenching of the fluorescence intensity of T-mismatched duplexes by utilizing the T-Hg$^+$ complex formation. This is the first example where mercuric ions have been utilized in this fashion, that is, selective quenching of a fluorophore in close proximity to T, and should be applicable with other fluorescent reporter groups.

2. Results and discussion

2.1. Synthesis of $\text{Cf}^\prime$ phosphoramidite

Our group has previously described the synthesis of $\text{C}$, a nitroxide spin label, and its incorporation into DNA. Upon reduction of the spin labelled DNA with Na$_2$S, highly fluorescent DNA was obtained, which was attributed to the amine derivative ($\text{Cf}^\prime$) of the spin label. As the amine is a precursor for the nitroxide, we sought to incorporate the phosphoramidite of $\text{Cf}^\prime$ directly, avoiding the oxidation and subsequent reduction steps (Scheme 1). A dimethoxytrityl group was introduced at the 5'-OH group of the fluoroside, followed by phosphitylation of the 3'-OH. The $\text{Cf}^\prime$ phosphoramidite was purified by repetitive precipitation from a CH$_2$Cl$_2$/n-hexane mixture and obtained in 66% yield over two steps.

2.2. Synthesis of modified oligonucleotides (ODNs)

$\text{Cf}^\prime$ phosphoramidite (2) was used in automated ODN synthesis to prepare the 16 ODNs for all possible immediately flanking sequences adjacent to the fluorescent nucleotide. The sequences of the 14-mer ODNs are the same except for the $\text{Cf}^\prime$-flanking bases. Syntheses of the modified ODNs was performed by following standard protocols, except for the incorporation of the modified phosphoramidite, for which the coupling time was extended (from 60 s + 30 s to 2/C2/300 s). All ODNs were purified by denaturing polyacrylamide gel electrophoresis (DPAGE) and their identity confirmed by MALDI-TOF mass spectrometry.

Table 1

<table>
<thead>
<tr>
<th>5'-d(GAC-CTC-N$^\text{G}$-TGTG)</th>
<th>$\lambda_{ab}$</th>
<th>$\lambda_{em}$</th>
<th>$\Phi_f$</th>
<th>$FB^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5'-d(GACG)</td>
<td>360</td>
<td>445</td>
<td>0.201 ± 0.007</td>
<td>1351</td>
</tr>
<tr>
<td>5'-d(GACG)</td>
<td>361</td>
<td>442</td>
<td>0.212 ± 0.005</td>
<td>1319</td>
</tr>
<tr>
<td>5'-d(GACG)</td>
<td>364</td>
<td>443</td>
<td>0.252 ± 0.005</td>
<td>2599</td>
</tr>
<tr>
<td>5'-d(GCTG)</td>
<td>361</td>
<td>442</td>
<td>0.215 ± 0.006</td>
<td>1328</td>
</tr>
<tr>
<td>5'-d(GCTG)</td>
<td>364</td>
<td>448</td>
<td>0.127 ± 0.003</td>
<td>921</td>
</tr>
<tr>
<td>5'-d(GCTG)</td>
<td>365</td>
<td>447</td>
<td>0.102 ± 0.002</td>
<td>761</td>
</tr>
<tr>
<td>5'-d(GCTG)</td>
<td>361</td>
<td>449</td>
<td>0.207 ± 0.004</td>
<td>1347</td>
</tr>
<tr>
<td>5'-d(GCTG)</td>
<td>362</td>
<td>449</td>
<td>0.119 ± 0.003</td>
<td>726</td>
</tr>
<tr>
<td>5'-d(GCTG)</td>
<td>359</td>
<td>450</td>
<td>0.176 ± 0.006</td>
<td>948</td>
</tr>
<tr>
<td>5'-d(GCTG)</td>
<td>362</td>
<td>449</td>
<td>0.156 ± 0.004</td>
<td>866</td>
</tr>
<tr>
<td>5'-d(GCTG)</td>
<td>364</td>
<td>449</td>
<td>0.179 ± 0.003</td>
<td>1219</td>
</tr>
<tr>
<td>5'-d(GCTG)</td>
<td>364</td>
<td>451</td>
<td>0.190 ± 0.004</td>
<td>1150</td>
</tr>
<tr>
<td>5'-d(GCTG)</td>
<td>364</td>
<td>450</td>
<td>0.140 ± 0.004</td>
<td>795</td>
</tr>
</tbody>
</table>

$^a$ Absorbance maxima.

$^b$ Emission maxima.

$^c$ Quantum yield of ODNs in H$_2$O.

$^d$ Fluorescence brightness, which is the multiplication product of the extinction coefficient at 365.5 nm and the quantum yield.
2.3. Photophysical properties of the fluorescent nucleoside $\text{C}^f$

The UV–vis and steady-state fluorescence emission spectra of the fluoroside were obtained in several solvents. Fluorescence emission quantum yields of $\text{C}^f$ were determined relative to a dilute solution of anthracene ($\Phi_F = 0.27$) in absolute ethanol. The quantum yields of $\text{C}^f$ in EtOH, water, 1,4-dioxane and CH$_3$CN were found to be 0.32, 0.31, 0.40 and 0.37, respectively. The normalized absorption and emission profile of the fluoroside in EtOH can be seen in Figure 2. The optical densities of the solutions used for quantum yield measurements were 0.02–0.07 and the quantum yield values were corrected with the refractive indexes of the solvents.

**Figure 3.** (A) The DNA sequence used for the flanking study, where N is G, C, A or T and X is $\text{C}^f$'s base-pairing partner. (B) The color codes of the fluorescence curves for the different $\text{C}^f$ base pairs. (C) The relative fluorescence intensity of all 64 duplexes. The nucleotides flanking the 5’- and 3’-side of $\text{C}^f$ are red and blue, respectively and change from G, C, A and T horizontally (5’-flanking) and in the same order vertically for 3’-flanking side. The sequences that can readily distinguish between all base-pairing partners have a green background, while those that are only able to identify a mismatch from the fully base-paired duplex have a yellow background. The sequences in which $\text{C}^f$ is unable to distinguish between the fully base-paired G and one of the mismatches have a red background. Each panel has been normalized separately by defining the emission intensity of the most fluorescent duplex as 1.00. Panels are highlighted green if the difference in area under the corrected emission curves exceeds 15% for all duplexes, excluding the A-mismatched duplex, which can easily be identified due to its unique 3-peak pattern in all duplexes. The panels are highlighted red, if any mismatched curve, excluding the A-mismatch, does not exceed 15% deviation from that of the fully base-paired duplex (black curve) and yellow if the deviation between mismatched curves, excluding the A-mismatch, does not exceed 15%.
vents.26 Degassing the solutions by sonication, or passing argon through the solution, before quantum yield determination had no significant effect on the emission intensity.

All UV–vis spectra contain one absorption band near the visible region, with the absorption maximum ranging between 359 and 364 nm. The fluorescence spectra lie within λ = 380–600 nm, with the emission maximum between 446 and 450 nm, yielding blue fluorescence. The molar extinction coefficient of the fluoroside in EtOH was calculated from Beer’s law and found to be ε_{365.5} = 11,330 ± 190 M^{-1} cm^{-1}. Therefore, C_1 is a relatively bright fluorescent nucleoside.10 Experiments in which the ion concentration was altered from 0 mM to 1000 mM NaCl, exhibited no significant change on the fluorescence intensity (data not shown).

2.4. Photophysical properties of single stranded C_1-modified ODNs

The UV–vis and steady-state fluorescence emission spectra were obtained in H_2O with concentration of single stranded DNA ranging from 3.0 to 12.5 μM. The UV–vis and fluorescence spectra were similar to those of the fluorescent nucleoside in solution, that is, a single absorption peak near the visible region and the emission maximum around 450 nm. Table 1 lists the photophysical properties of the 16 C_1-modified ODNs. The fluorescence intensity of C_1 is generally higher when the fluoroside has a 5’-flanking C. The fluorescence brightness is especially high in the sequence 5’-d(CCA), attributed to an especially high molar extinction coefficient, in addition to the highest quantum yield of all modified ODNs. The related sequence 5’-d(ACC) also exhibits the highest emission intensity of its series, that is, when the fluoroside has a 5’-flanking A. The quantum yield, and therefore the brightness, is lower in sequences where there is a 5’-flanking C, but the quenching effects are not nearly as severe as described for other BDFs.15 Some of the observed variations in emission intensity could be due to secondary structure formation, however, no such stable structures were predicted by mFold (not shown).

2.5. Fluorescence of C_1-modified complementary and mismatched duplexes

Each of the 16 fluorescent ODNs was annealed to its complementary strand (G opposing C_1) as well as the three possible mismatches (A, C and T opposing C_1). All duplexes are expected to be stable at 20 °C, based on T_m values of 46–63 °C for the oligonucleotide with flanking sequence 5’-d(GCA).23 The relative fluorescence intensities of the 64 duplexes at 20 °C is shown in Figure 3.

It is clear from the data in Figure 3 that there is a high degree of variance in the order of fluorescence intensity for the individual base-pairings between the different flanking sequences. Some modified sequences are clearly superior to others when it comes to discrimination between mismatches, or simply to distinguish a mismatch altogether. The flanking sequences can be grouped into three distinct classes, based on the difference in fluorescence; spectra are considered to overlap one another if the difference in the area under the corrected emission curves is less than 15%, except for A-mismatched duplexes which can clearly be identified due to the unique 3 peak pattern in the emission spectra. In the first class (Fig. 3C green background), C_1 is able to distinguish and identify the mismatches. C_1 can identify its base-pairing partner in 10 out of 16 sequences. In the second class (Fig. 3C yellow background), the fluorescence probe can distinguish the fully base-paired duplex from a mismatch, but is unable to identify the mismatch in question (3 out of 16). In these sequences the fully base-paired duplex has comparatively lower fluorescence intensity than the mismatches. The reason why C_1 is incapable of identifying its base-pairing partner in these three flanking sequences is that the emission profiles of the C- and T-mismatch overlap. A third and final class of modified sequences are the red panels in Figure 3C (3 out of 16), which contain those sequences in which the fluoroside is incapable of distinguishing all mismatches from the fully base-paired duplex. In the red panels the T-mismatch and G-match emission profiles are nearly superimposable.

As mentioned above, the duplexes containing an A-mismatch exhibit a unique spectral profile in all sequences. Rather than one emission peak, three distinct peaks are observed at approximately 420, 450 and 480 nm, with an additional shoulder at 525 nm. The reason for the appearance of these vibrational peaks is unclear, but is possibly associated with a tautomer of the fluoroside which, if formed, would stabilize the A-mismatched duplex with an additional hydrogen bond. Another intriguing result is that the fluorescence intensity of the C-mismatch is consistently higher than that of the fully base-paired duplex. Therefore, C_1 can distinguish and discriminate the A- and C-mismatch from the fully base-paired duplex in all sequences.

All the flanking sequences that do not show either mismatch identification (green) or mismatch detection (yellow) have one thing in common: The emission of the T-mismatched duplex overlaps with the emission of another base-pairing partner. Figure 3 clearly shows that in 10 out of 16 sequences, C_1 can distinguish and identify the T-mismatch from the other possibilities. However, that leaves the 6 sequences 5’-d(TCT), 5’-d(ACG), 5’-d(CCA), 5’-d(ACG), 5’-d(GGA) and 5’-d(GCG), where the fluorescent probe is unable to unambiguously identify the T-mismatch. Therefore, we varied the conditions under which fluorescence was recorded, in an attempt to facilitate discrimination between all base-pairing partners.

2.6. Organic co-solvents, potassium iodide and temperature as external effectors to induce mismatch discrimination

As the emission intensity of the fluorescent nucleoside showed some variation in fluorescence with changes in solvent, we postulated that the addition of an organic co-solvent to the DNA sample would induce changes in the emission. Of the four organic co-solvents examined (EtOH, DMF, THF and 1,4-dioxane), only THF induced better discrimination between mismatches, but to a small extent (data not shown). That may be due to the fact that the amount of organic co-solvent could only be as high as ca. 25%, at which point the DNA samples started to precipitate. A KI titration was also performed to assess the solvent accessibility of the probe.27,28 The relative fluorescence intensity of matched and mismatched duplexes did not change upon addition of KI (data not shown), indicating a similar exposure of the fluorescent nucleoside to the organic solvent.

Temperature-dependence of fluorescence was also investigated. Fluorescence intensities were measured as a function of temperature using two different sequences, 5’-d(TCT) and 5’-d(TCA). These sequences were chosen because the former is an example of sequence where C_1 is unable to identify its base-pairing partner, but in the latter C_1 can distinguish and identify all base-pairing partners. While shifting the temperature from 10 °C to 40 °C, a gradual increase in mismatch discrimination was afforded in 5’-d(TCT) (Fig. 4A). However, the mismatch discrimination was lowered in 5’-d(TCA) (Fig. 4B). The emission intensity of the T-mismatch does not change significantly for these two flanking sequences upon increasing the temperature. The emission intensities of the A-mismatch and the fully base-paired duplex were lowered, while the emission intensity of the C-mismatch increased. Therefore, while temperature changes can enhance the discrimination in some sequences, it will reduce the discrimination in others.
2.7. The effects of mercuric ions (Hg\(^{2+}\)) on mismatch detection

As previously mentioned, the reason that the base-pairing partner of C\(^{\text{f}}\) cannot be identified in all flanking sequences, is due to overlap of the emission profile of the T-mismatch with other base-pairing partners. To enhance mismatch discrimination, we therefore sought a way to selectively change the emission properties of C\(^{\text{f}}\)-T. Mercuric ions bind to nucleotides, in particular T,\(^{29-31}\) and are also known to quench the fluorescence of aromatic hydrocarbons.\(^{32,33}\) The quenching effect is based on the spatial distance between the fluorophore and the mercuric ions.\(^{34}\) Therefore, we hypothesized that the selectivity of the T-Hg\(^{+}\) complex formation could be utilized to bring the mercuric ion into close proximity to C\(^{\text{f}}\) and thereby selectively quench the emission of the T-mismatch.

Upon titrating a 1 mM HgCl\(_2\) solution into the DNA duplexes of 5'-d(A\(^{\text{f}}\)T), only a gradual decrease was initially observed in the fluorescence intensity, presumably due to the presence of EDTA.

![Figure 4. Normalized fluorescence spectra of DNA oligomers 5'-d(T\(^{\text{f}}\)C) (A) and 5'-d(T\(^{\text{f}}\)G) (B) at 10 °C (left) and at 40 °C (right), C\(^{\text{f}}\) G (black), C\(^{\text{f}}\) C (red), C\(^{\text{f}}\) A (blue) and C\(^{\text{f}}\) T (green). All spectra are normalized in reference to the C\(^{\text{f}}\) C mismatched duplex at 40 °C, within each panel (A and B), as this shows the highest emission intensity.](image)

![Figure 5. Titration curves of the fluorescence intensity of duplexes as a function of mercuric ion concentration, using ODN 5'-d(A\(^{\text{f}}\)T). C\(^{\text{f}}\) G (black), C\(^{\text{f}}\) C (red), C\(^{\text{f}}\) A (blue) and C\(^{\text{f}}\) T (green). 10 µL aliquots of 1 mM Hg\(^{2+}\) solution were added to a solution of the oligonucleotides in a phosphate buffer (400 µL, 10 mM Na\(_2\)HPO\(_4\), 100 mM NaCl, 0.1 mM Na\(_2\)EDTA, pH 7.00).](image)

![Figure 6. The relative fluorescence spectra of the sequences titrated with a 1 mM HgCl\(_2\) solution. In the absence (left) and presence (right) of mercuric ions. (A) 5'-d(T\(^{\text{f}}\)C)(T)\(_{10}\), (B) 5'-d(T\(^{\text{f}}\)C)(A)\(_{10}\), (C) 5'-d(C\(^{\text{f}}\)A)(T)\(_{10}\), (D) 5'-d(C\(^{\text{f}}\)A)(A)\(_{10}\), (E) 5'-d(G\(^{\text{f}}\)C)(T)\(_{10}\), (F) 5'-d(G\(^{\text{f}}\)C)(A)\(_{10}\). The final concentrations of added Hg\(^{2+}\) ions were 91, 70, 91, 48, 91 and 111 µM, respectively. The color codes are the same for all spectra, C\(^{\text{f}}\) G (black), C\(^{\text{f}}\) C (red), C\(^{\text{f}}\) A (blue) and C\(^{\text{f}}\) T (green).](image)
in the buffer (Fig. 5). After that, a dramatic decrease in the emission intensity of the \(C^T-T\) duplex was observed. The emission of the C-mismatched duplex was also quenched, but to a smaller degree, as C has a lower affinity for mercuric ions than T. The A-mismatch and the fully base-paired duplex were affected much less. At the highest concentration of added mercuric ions (200 \(\mu M\)), all duplexes experienced considerable quenching, due to the general quenching effects of mercury. All flanking sequences where mismatch identification had been problematic, 5'-d(TCIT), 5'-d(CATC), 5'-d(CCAC), 5'-d(CCTC) and 5'-d(CGTC), were treated with mercuric ions (Fig. 6). All flanking sequences show the same trend as described for the 5'-d(ACT) mercuric titration, that is, a dramatic decrease in the emission of the \(C^T-T\) duplex, while the C-mismatch is quenched to a lesser extent. Furthermore, the two sequences 5'-d(CGTC) and 5'-d(CGTC), which fall very close to the 15% cutoff, were subjected to mercuric titrations. They followed the same trend as other sequences (Supplementary data). Thus, by the addition of mercuric ions, the mismatch identification problems in all flanking sequences were resolved.

3. Conclusion

The development of a fluorescent probe that can identify all base-pairing partners is of considerable interest due to their possible use in SNP assays. Herein, we have described the emission properties of fluorescent nucleoside \(C^T\), its incorporation into DNA and the emissive properties of the \(C^T\)-labelled ODNs and duplexes. In particular we have examined the effects of flanking sequence on mismatch detection. These studies have shown that \(C^T\) is a promising probe for SNP assays using synthetic oligonucleotides. However the applicability of identifying all mismatches in small amounts of unknown samples remains to be seen.

When reviewing the results of the mismatch detection, it is apparent that there is a high degree of variance in the order of fluorescence intensity for the individual base-pairings between the different flanking sequences. The sequences can be categorized into three groups. First, in 10 out of 16 sequences, \(C^T\) is able to distinguish and identify a mismatch. Second, in 3 out of 16 sequences, \(C^T\) is able to distinguish a mismatch from the fully base-paired duplex. In the remaining 3 sequences, \(C^T\) is unable to distinguish all mismatches from the fully base-paired duplex. In all sequences the A-mismatched duplex has a unique peak pattern, which effectively makes \(C^T\) an ideal SNP probe for the A-allele of any SNP site. Furthermore, the fluorescence intensity of \(C^T\) is not severely quenched by a flanking G–C pair, unlike many other BDFs.

For the sequences in which \(C^T\) was incapable of uniquely identifying its base-pairing partner, mercuric ions proved extremely useful to facilitate discrimination. The mercuric ions selectively quenched the emission of the T-mismatched duplexes, eliminating any spectral overlap. This enabled mismatch detection in all the sequences and shows that \(C^T\) is a probe that identifies individual mismatches, independent of the flanking sequence. To our knowledge, this is the first time that mercuric ions have been used to selectively quench the fluorescence of a probe that is proximal to T, and should find use in other fluorescence-based assays.

Acknowledgments

We thank the Icelandic Research Fund (060028021) for financial support, Dr. Ajay Kumar Kale for technical assistance, Dr. Pavol Cekan for valuable discussions, and the reviewers of this paper for valuable suggestions.

Supplementary data

Supplementary data (experimental procedures, characterization data and photophysical data) associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2010.06.060.

References and notes