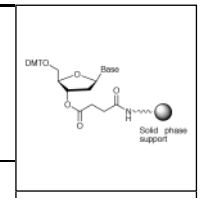
Synthesis of fluorescently labelled oligonucleotides and nucleic acids



Martin J. Davies,* Aneela Shah and Ian J. Bruce

School of Chemical and Life Sciences, University of Greenwich, Wellington Street, London, UK SE18 6PF

Received 23rd August 1999

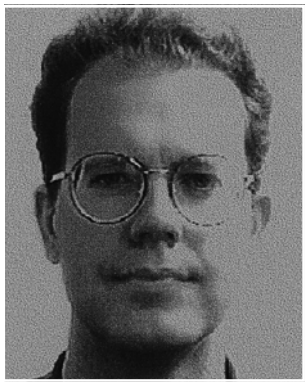
The demand for and application of fluorescently labelled nucleic acid materials are growing, driven by ventures such as the Human Genome Mapping Project and the advent of DNA-based clinical diagnostics. This article surveys the strategies that are currently available for conjugation of fluorescent molecules to oligonucleotides and nucleic acids, and demonstrates how the labelled materials have been used in new and innovative assays for probing nucleic acid structure and properties.

1 Introduction

In recent years, the detection, quantification and identification of nucleic acids have grown in importance, driven by ventures such as the Human Genome Mapping Project (HGMP), and the advent of nucleic acid-based clinical diagnostics. This has been coupled with a move away from traditional radioisotopic methods for nucleic acid detection, with a concomitant increase in the use of fluorescence for identifying nucleic acid species. This was initiated by the introduction in the late 1980's of fully automated nucleic acid sequencing systems using laser-based fluorescence detection,[1] which made researchers more aware of other benefits of fluorescence that traditional methods based on radioisotopic detection lacked. Using fluorescence detection means that problems of stability, storage and disposal of radioactive compounds are avoided.[2] Furthermore, the introduction of new hardware and detection systems ensures that fluorescence detection can be performed with the same levels of sensitivity as radioisotopic detection, and also offers new opportunities to probe molecular interactions, biochemical processes and cellular function by exploiting fundamental properties of the fluorescence process itself.

Thus, demand for and application of fluorescently labelled nucleic acid materials are growing. Automated, high-throughput fluorescent nucleic acid sequencing systems are run round-the-clock to generate sequence data for the HGMP, and National Forensic Services' DNA databases. Similarly, the polymerase chain reaction (PCR) amplification of nucleic acids is now regularly performed using fluorescently labelled oligonucleotide primers to produce an amplified DNA product that can be detected and quantified absolutely.[3] PCR can also benefit from a further advantage of fluorescent detection methods: a

Martin Davies received a BSc degree in chemistry from Imperial College, London, in 1987 and undertook postgraduate research in organic chemistry at Imperial under the supervision of Professor C. J. Moody. He obtained his PhD in 1990, and subsequently undertook post-doctoral research at the University of Texas, Austin, and the University of Greenwich, where he is now Senior Lecturer in Bio-organic Chemistry. His current research interests are the development of new recognition media for biomolecular purifications, and synthesis and applications of modified nucleic acids.

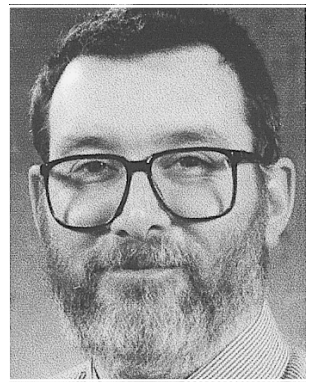


Martin J. Davies

Aneela Shah studied for both BSc and PhD degrees in the Department of Chemistry, University College, London, completing her PhD in organic chemistry in 1994. She has held post-doctoral research positions at the University of Durham, Hammersmith Hospital and Imperial College, London, and is now at the University of Greenwich where she is a senior research fellow in organic chemistry, studying the synthesis of novel multi-functional DNA labelling agents.

Aneela Shah

Ian Bruce gained his first degree and PhD from University College, London, and following research positions at UCL and Guy's Hospital, London, moved to Thames Polytechnic (now the University of Greenwich) in 1987. He was made Professor of Molecular Biology in 1992, and has more than a dozen years' experience in the area of methods development in molecular biology, collaborating closely with industry and other research institutions. His current research interests include molecular diagnostics and techniques for nucleic acid detection, and the environmental dissimilation of xenobiotics.



Ian J. Bruce

wide range of fluorochromes is now commercially available with spectral characteristics (λ_{max} excitation and λ_{max} emission) covering the wavelength range 350 to 700 nm, and into the near infra-red region of the electromagnetic spectrum. Thus, simultaneous, multiple detection of labelled molecules can be performed on the same sample, for example, following 'multiplex' PCR amplification of several nucleic acid sequences using pairs of oligonucleotide primers labelled with different fluorophores. Each pair gives rise to a separate amplified product that can be unambiguously identified due to its fluorescent label.[4]

Similarly, multiple, simultaneous fluorescent detection has enabled the technique of fluorescent *in situ* hybridisation (FISH) to become an important tool for clinical diagnosis and gene mapping.[5] Labelled oligonucleotide probes are used to locate specific nucleic acid sequences in cells and tissues, and with the number of fluorochromes available there is the potential to visualise several fluorescent signals relating to different genetic sequences within the same sample.

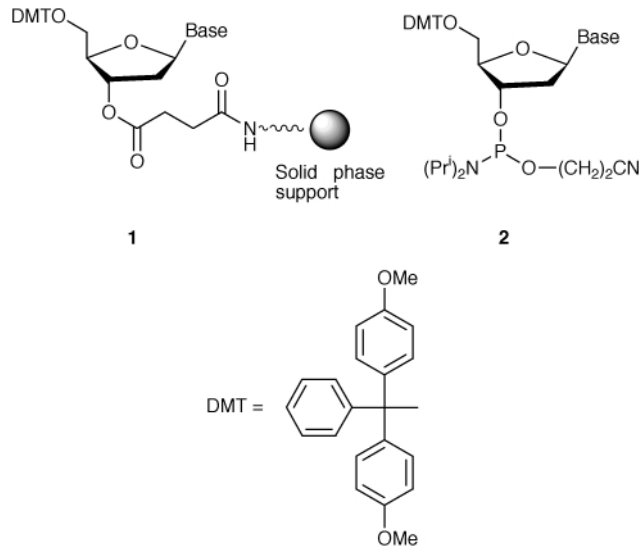
Labelled nucleic acids are also required to exploit the emerging technology of chip-based DNA arrays prepared using combinatorial chemistry methods.[6] High density oligonucleotide arrays are probed with fluorescently labelled nucleic acid species, for example, from a clinical sample, and the position of any hybridised, labelled nucleic acid identified using fluorescence microscopy. The position relates to a known oligonucleotide sequence immobilised at that part of the chip. This technology has found wide application in the sequencing by hybridisation (SBH) method, and is set to revolutionise disease diagnosis by monitoring of gene expression.

The advent of fluorescence energy-transfer (FRET) based detection has made possible several improvements to nucleic acid monitoring and identification, by enhanced sensitivity of detection, or through the advent of real-time monitoring of PCR.[7] FRET relies upon the interaction of a 'reporter' fluorochrome and 'quencher' molecule, which may also be fluorescent. If the reporter and quencher molecules are in close proximity to one another, the quencher molecule absorbs the fluorescent signal emitted by the reporter upon excitation. However, when the two molecules are held apart from one another, fluorescence can be detected. For example, use of so-called 'molecular beacons'—oligonucleotide probes which possess a fluorochrome and non-fluorescent quencher at the 5′- and 3′-terminus respectively—allows the diagnostic detection of specific DNA sequences by fluorescence.[8] The probes possess hairpin-forming regions, and in the absence of a complementary nucleic acid strand, the fluorochrome and quencher are in close proximity to one another and quenching of the fluorescent signal results. When incubated with a target DNA that possesses the necessary complementary sequence, the probe anneals to the target such that the fluorochrome and probe are held apart from one another, and fluorescence can be detected signifying the presence of a particular DNA sequencing. FRET is also exploited in the Taqman® system pioneered by Perkin-Elmer for the quantitative real-time monitoring of PCR products.[9]

Application of the technologies described above, and others, requires access to the appropriately labelled nucleic acid or oligonucleotide. The purpose of this article is to survey the strategies that are currently available for conjugation of fluorescent molecules to oligonucleotides and nucleic acids, and to demonstrate with representative examples how labelled materials so produced have been utilised.

2 Oligonucleotide labelling

Oligonucleotide labelling techniques can be divided into two categories, namely automated and manual methods. Almost all synthetic oligonucleotides are prepared by automated solid-phase synthesis techniques using cyanoethyl phosphoramidite building blocks.[10] The first deoxynucleotide in a sequence to be prepared is attached to a solid phase support, either controlled-pore glass (CPG) or polystyrene (**1**). The nucleotide is protected



at its 5′-position with a 4,4′-dimethoxytrityl (DMT) group, which is removed using trichloroacetic acid to reveal the free 5′-hydroxy, to which cyanoethyl phosphoramidite derivatives (**2**) of subsequent nucleotides in the sequence are added. Further reactions convert the P(III) species to a protected phosphodiester P(V) species, and remove the next DMT group for attachment of the next nucleotide to the growing solid-supported sequence.

Colorimetric assay of the released DMT group also serves as a measure of coupling yield, which is usually at least 99%. Delivery of all reagents and reaction times are controlled through the automated synthesiser's programme. Upon completion of the synthesis, the oligonucleotide is cleaved from the support using aqueous ammonia, which also removes protecting groups from the phosphodiester and nucleobases. Automated fluorescent labelling can be performed using the same chemistry, by incorporating a fluorescent label directly into the synthetic sequence. Labels can be attached at the 5′-terminus, 3′-terminus, the nucleotide bases or the phosphate/sugar backbone, and automated labelling gives high yields of labelled material. These can be adequately purified using proprietary, single-use reversed-phase oligonucleotide purification cartridges, but HPLC purification is the best option if stringently purified materials are needed.

Manual labelling requires the automated synthesis of oligonucleotides bearing functional groups such as phosphate, amine or thiol at the position to be labelled, which can subsequently be conjugated with a fluorescent label post-synthesis. An advantage is that a far greater variety of labels suitable for manual conjugation to the oligonucleotide are available from companies such as Molecular Probes Inc. (Eugene, Oregon, USA) but yields are generally lower and the process is labour intensive. Two or even three rounds of HPLC purification are often required to separate labelled and unlabelled material, and to remove excess label, as fluorochromes can show high non-specific affinity for DNA.

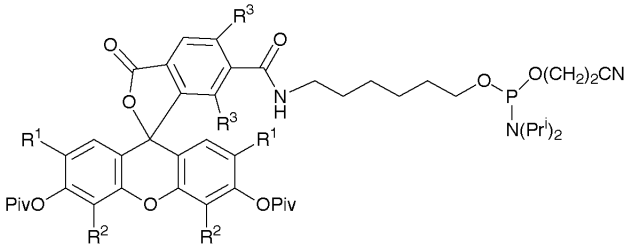
A number of companies now offer a range of specialist phosphoramidites for the introduction of a fluorochrome or other modification. These include Perkin-Elmer/Applied Biosystems (Foster City, California, USA; http://www2.perkin-elmer.com), Cruachem Ltd. (Glasgow, UK; http://www.cruachem.co.uk), Clontech Laboratories, Inc. (Palo Alto, California, USA; http://www.clontech.com) and Glen Research Corporation (Sterling, Virginia, USA; http://www.glenres.

com). The Internet websites of these companies are a particularly good source of information on the products offered and their applications and properties, together with references to original research papers. All of the commercially available compounds discussed below can be obtained from one or more of these companies or their distributors. Many more companies offer custom oligonucleotide synthesis services, and will supply labelled or modified materials to a particular level of purification, though it should be borne in mind that there is often a cost premium associated with such services.

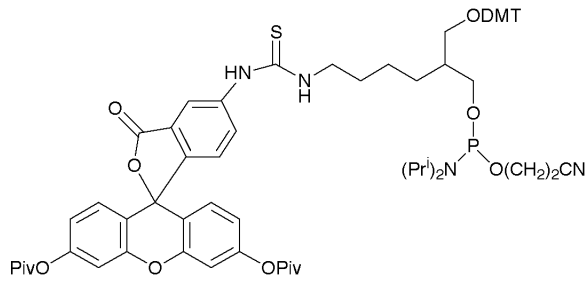
2.1 Labelling of oligonucleotides at the 5′-terminus

(i) Automated labelling. Automated label incorporation is the most direct route to oligonucleotides possessing fluorescent groups at the 5′-position, and is probably the most widely used of all labelling strategies. Fluorescent phosphoramidites can be coupled to the 5′-terminus of synthetic oligonucleotides at the last stage of a synthetic sequence, and by increasing the reaction time specified by the automated synthesiser's programme, possible problems of lower reactivity due to increased steric bulk of the fluorescent phosphoramidite can be overcome. Commercially available phosphoramidites of the most popular fluorochromes based on fluorescein (**3**) and (**6**), tetrachlorofluorescein (**4**), hexachlorofluorescein (**5**) and rhodamine (**7**) are shown. These are heavily used in four-colour automated fluorescent DNA sequencing, and indeed the tetrachloro- and hexachlorofluorescein derivatives were developed specifically for this purpose. Structures **6** and **7** both possess an additional branching arm terminated in a DMT-protected hydroxy group, which when removed can be used in further rounds of automated synthesis as an anchor for the attachment of more than one fluorochrome, or other reporter group. However, care should be exercised when using such a strategy for multiple label attachment, as the tolerance of the labels to the conditions encountered during automated synthesis decreases with each successive synthetic round.

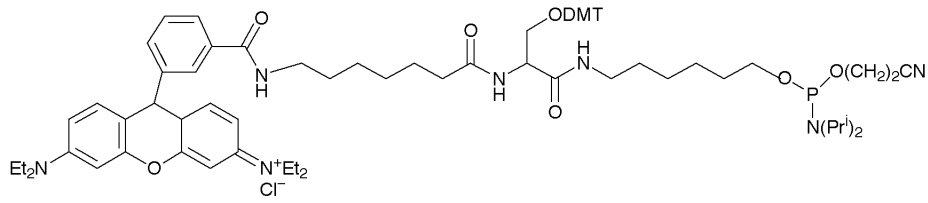
If more specialised compounds are required as a phosphoramidite, they must be synthesised in the laboratory. A prerequisite is that the label has a primary or secondary hydroxy group which can be converted to the cyanoethyl phosphoramidite, using a phosphitilating agent such as $ClP(OCH_2CH_2CN)N(Pr^i)_2$, which is readily available from several reagent suppliers.



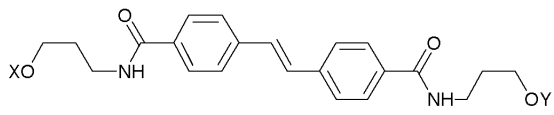
3 $R^1 = R^2 = R^3 = H$
4 $R^1 = R^3 = Cl$; $R^2 = H$
5 $R^1 = R^2 = R^3 = Cl$



6



7



8 a) X = Y = H
b) X = DMT; Y = H
c) X = DMT; Y = $P(N(Pr^i_2)(O(CH_2)_2CN)$

Letsinger and Wu synthesised a specialist phosphoramidite containing a fluorescent stilbenedicarboxamide bridge, which was subsequently found to be unusually effective in stabilizing short double and single stranded oligonucleotides.[11] This group also served as a useful means for monitoring the local environment of the bridge through fluorimetric detection, and thus for probing oligonucleotide assembly. For example, assemblies of oligonucleotides carrying this bridge in which two stilbene moieties were in close proximity, as was the case with mixed base duplexes, could be distinguished from those in which an isolated stilbene was present in a monomolecular hairpin structure. The former displayed excimer fluorescence, whilst the latter fluoresced as expected for stilbene. The stilbenedicarboxamide phosphoramidite **8c** was prepared by the conversion of *trans*-stilbene-4,4′-dicarboxylic acid, through the diacid chloride and diol (**8a**), to the mono-(4,4′-dimethoxytrityl) derivative (**8b**). This was phosphitilated to give **8c**, which was used to prepare stilbene–oligonucleotide conjugates labelled at the 5′-terminus using conventional automated synthesis techniques. The yield of the fluorescent conjugate was judged to be greater than 95%.

(ii) Manual labelling. *(a) via a 5′-amino functionality. The fluorescent labelling of an oligonucleotide at the 5′*-terminus by this route usually involves two steps, and most commonly proceeds as follows. An *N*-protected aminoalkyl phosphoramidite derivative is coupled to the 5′-terminus of the oligonucleotide at the final stage of automated DNA synthesis. Following cleavage and deprotection, an appropriately substituted fluorescent dye can then be manually coupled to the 5′-amino group. Typical derivatives are isothiocyanates, and *N*-hydroxysuccinimidyl (NHS) esters, and the latter are most favoured as they are less prone to competing hydrolysis in the aqueous, basic reaction medium, giving higher yields of labelled oligonucleotides (Fig. 1). Also, the thiocarbamate

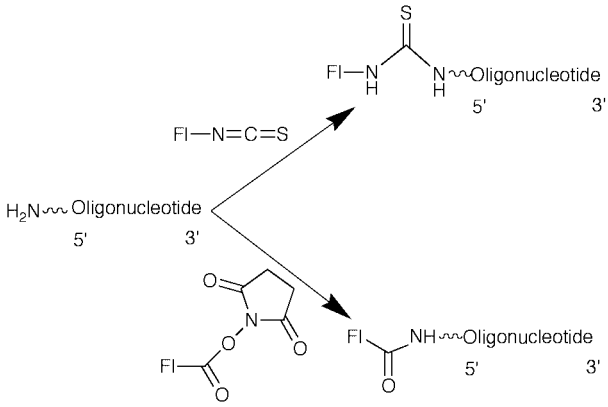
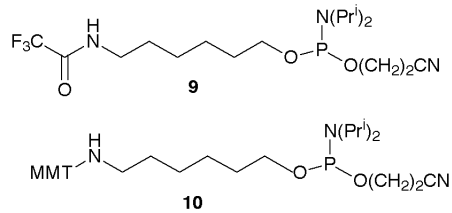


Fig. 1 Labelling of 5′-amino terminated oligonucleotides with isothiocyanate and *N*-hydroxysuccinimidyl ester derivatives of fluorochromes (Fl).

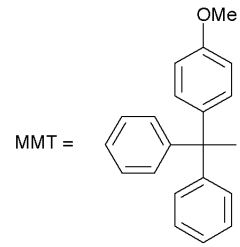
linkage formed by reaction with an ITC derivative is more prone to degradation during long-term storage. An excellent source of information on all aspects of manual modification of amino-terminated oligonucleotides (and other derivatives), as well as selection of appropriate fluorochromes, and fluorescence technology, is supplied by Molecular Probes Inc., in both hard copy and through an Internet website, on which the same information is duplicated.[12] This source also contains several thousand references to original research papers and is probably the most comprehensive resource of its kind currently available.

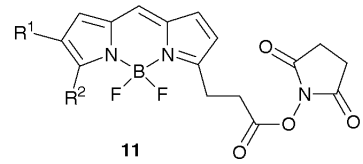
Many phosphoramidites containing protected amino groups have been developed in order to generate aliphatic amino groups at the 5′-terminus of oligonucleotides after cleavage and deprotection. Structures **9** and **10** are typical of these, and **9** is



9

10





11

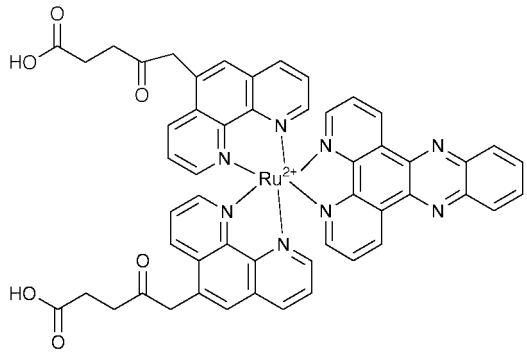
a $R^1 = CH_3$, $R^2 = CH_3$
b $R^1 = H$, $R^2 = C_6H_5$
c $R^1 = H$, $R^2 = -CH{=}CH\text{-}C_6H_5$
d $R^1 = H$, $R^2 = -CH{=}CH\text{-}CH{=}CH\text{-}C_6H_5$

probably the most widely used. The trifluoroacetyl group is removed during the post-synthesis treatment with aqueous ammonia. Compound **10** possesses a lipophilic group that can be retained after synthesis to aid purification of labelled material by reversed-phase (RP) HPLC. Mild acid treatment will subsequently remove the MMT group. Other variants are available which have longer hydrocarbon chains (to minimise steric interactions between label and oligonucleotide), and hydrophilic ethylene glycol linkages if the presence of a hydrophobic linker is not desired.

A recent application of this labelling method is described by Metzker *et al*. who prepared a range of oligonucleotides bearing the dipyrrollomethene boron difluoride (BODIPY) group[12] using the NHS ester of these dyes, which were attached through various linker arm modifications.[13] The labelled compounds (**11a**–**d**) were electrophoretically uniform in DNA sequencing gels, and showed improved spectral characteristics compared with compounds labelled with conventional fluorescein and rhodamine dyes. Their spectral overlap was narrower, and fluorescent intensities higher, meaning that more accurate and efficient high-throughput sequencing could be performed.

Modifications of these strategies need to be employed if the ITC or NHS ester derivative is not available. For example, the dipyridophenazine–ruthenium(II) complex **12** was coupled to an amino-linked oligonucleotide by a two-step reaction involving activation of the carboxylic acid group of the Ru(II) complex with dicyclohexylcarbodiimide and 1-hydroxybenzotriazole, followed by reaction with the oligonucleotide under basic (LiOH) conditions.[14] This derivative was prepared specifically to probe intramolecular interactions of the tethered label with complementary hybridised single stranded DNA. The Ru(II) complex is a strong, sequence-specific DNA intercalator, and fluoresces intensely in the presence of the complementary strand to its oligonucleotide partner, as intercalation shields the phenazine ring from quenching by water. In the absence of a complementary strand, negligible fluorescence is observed and thus the system serves as a method for unequivocal DNA sequence identification.

Many other groups have studied the use of chelated metal complexes tethered to oligonucleotides as fluorescent reporters,



12

particularly those in which complexation of a lanthanide ion is required for fluorescence to arise, and informative reviews of this subject are available.[15] Often, complexes are assembled in two stages: firstly, the complexing agent is covalently attached to the oligonucleotide, and this derivative treated with a solution of the lanthanide ion to produce the fluorescent metal–chelate complex.[16,17] The advantage of labelling by this method is that both chelate and lanthanide ion are not fluorescent by themselves and therefore background signals are very low.

Prior to the development of aminating reagents for use in automated synthesis, introduction of amino functionality at the 5′-terminus was commonly achieved by initial treatment of an oligonucleotide obtained through automated synthesis with a phosphorylating enzyme, polynucleotide T4 kinase, giving a 5′-phosphate terminated species.[12] This can then be reacted with a diamine such as ethylene diamine in the presence of imidazole and a water-soluble carbodiimide (*e.g.* 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide, EDC) to give a 5′-amino terminated oligonucleotide (Fig. 2). More recently, reagents such as **13**

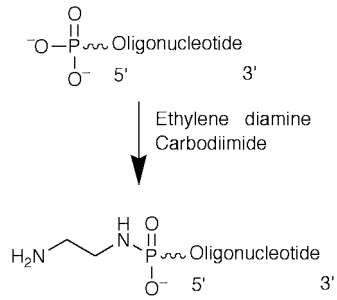
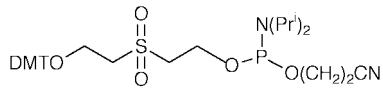


Fig. 2 Conversion of 5′-phosphate group to a 5′-amino group using ethylene diamine and a water-soluble carbodiimide.

have been developed which introduce the phosphate group during automated synthesis, and again these are available from the suppliers mentioned above.



13

Although less widely used today, this strategy can be useful for attaching novel labels without recourse to preparation of a specialised phosphoramidite derivative. Shinozuka has reported the preparation of novel bifunctional labels carrying acridine and fluorescein, and naphthalenesulfonic acid and tertiary amine groups, and their conjugation to 5′-phosphate terminated oligonucleotides using this approach.[18]

(b) via a 5-phosphate or 5′-hydroxy functionality. As described earlier, 5′-phosphorylation, followed by amination, has been a popular method for introducing functionality which can subsequently be modified with a fluorochrome. An obvious extension of this approach which avoids the need for what is essentially a three-step labelling strategy, is to use fluorochromes which already have suitable amine functionality attached. These can either be prepared in the laboratory, or are commercially available. For example, Molecular Probes sell ethylamine or cadaverine derivatives of some of their popular fluorochromes for direct conjugation to a 5′-phosphate terminated oligonucleotide.

In certain situations, synthesis and coupling of specialised fluorescent phosphoramidites can be problematic, and an alternative manual coupling strategy is required. A recent example is reported by McLaughlin and co-workers who attempted to prepare a phosphoramidite of the hexaethylene glycol-derivative of the fluorescent dye Hoechst 33258 (**14**), in order to study duplex stabilising properties of the resulting labelled oligonucleotide.[19] They found that the phosphoramidite could not be prepared under a variety of conditions, and reasoned that the benzimidazole rings were participating in a competing nucleophilic reaction with the phosphochloridite used to prepare the phosphoramidite. Instead, they adopted a 'reverse coupling' protocol, in which the oligonucleotide to which the dye was to be coupled was treated, in two successive reactions, with a mixture of (2-cyanoethoxy)bis(diisopropylamino)phosphine and tetrazole whilst still bound to the CPG support after synthesis (Fig. 3). This converted the 5′-hydroxy group to its phosphoramidite, which was then reacted with the hexaethylene-Hoechst 33258 compound through its terminal primary hydroxy group, followed by oxidation to a protected phosphodiester and post-synthesis cleavage and deprotection. Yields of the resulting conjugate varied between 50 and 75% after HPLC purification, but the method avoided the need for preparation of any protected version of the Hoechst dye.

(c) via a 5-thiol functionality. The thiol (–SH) functional group can be conjugated with reporter groups under similar reaction conditions to those used with 5′-amino terminated oligonucleotides. Instead of ITC or NHS ester derivatives, the thiol group reacts preferentially with maleimide or iodoacetamide derivatives of fluorochromes. Typical commercially available phosphoramidites that allow introduction of thiol groups at the 5′-terminus are **15** and **16**. Compound **15** carries a trityl group to aid RP-HPLC purification post-synthesis. Deprotection then reveals the free thiol. The disulfide link in compound **16** can be cleaved using dithiothreitol to yield the free thiol after synthesis and deprotection.

Manual approaches to the placement of thiol functionality at the 5-terminus have also been reported, which precede the introduction of phosphoramidites, and which were originally developed for modification of DNA and RNA rather than short oligonucleotides. However, these are also applicable to oligonucleotide modification, and a summary of these methods can be found in reference 12.

2.2 Labelling of oligonucleotides at the 3′-terminus

(i) Automated labelling. As with labelling at the 5′-terminus, a number of reagents are now commercially available for directly introducing a fluorescent label at the 3′-terminus of an oligonucleotide. These are almost all sold as the CPG-supported variant. The label is covalently attached to the CPG solid phase support and supplied pre-packed in an oligonucleotide synthesis cartridge, with DMT protecting groups as necessary, which are removed in the first steps of automated synthesis. The resulting free 5′-OH group acts as the starting point for oligonucleotide chain synthesis in the 3′ to 5′ sense. CPG-supported fluorescein, rhodamine, pyrene and dansyl groups are all commercially available.

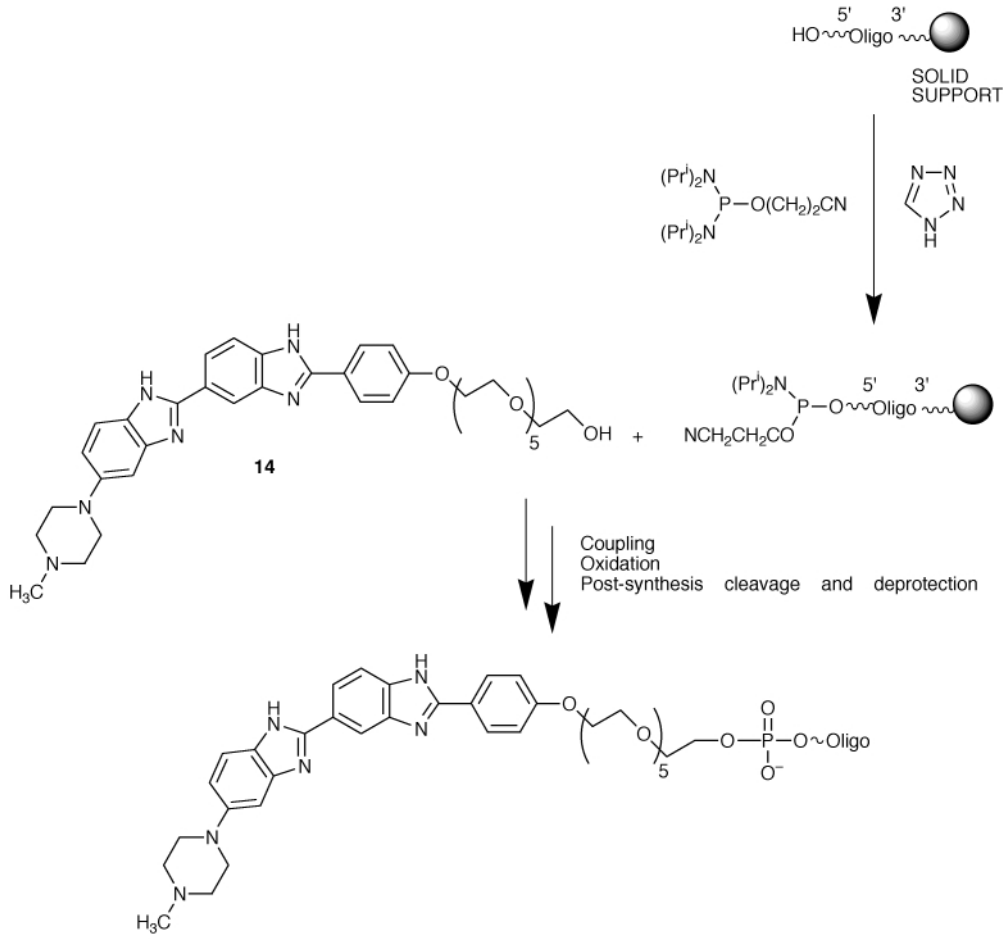
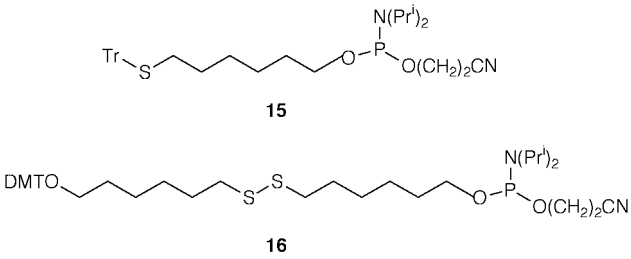
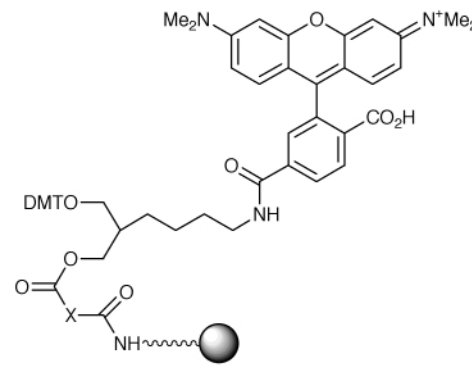


Fig. 3 Preparation of a Hoechst 33258–oligonucleotide conjugate using a reverse coupling strategy.



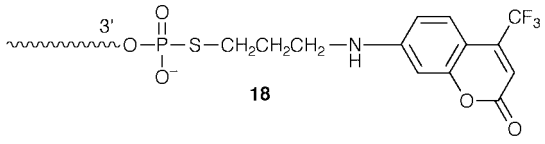


If labelling with other fluorescent groups at the 3′-terminus is required then again, a specialised reagent must be prepared in the laboratory. These must then be attached to a CPG support for packing in a column assembly. Common procedures involve prior derivatisation of the CPG with an aminotrialkoxysilane, and subsequent reaction of the primary amine with succinic anhydride to give a carboxy-terminated CPG. The carboxy group serves as the anchoring point for attachment of the 3′-modifying agent. For example, Mullah and Andrus synthesised the tetramethylrhodamine (TAMRA) solid supports **17**.[20] This support was used to good effect in the preparation of doubly labelled oligonucleotides in which the 5′-terminus carried a fluorescein group. FRET between the fluorescein (reporter) and rhodamine (quencher) groups was measured as a method for monitoring hybridisation of the probe to a complementary single stranded target. FRET decreased when hybridisation occurred, as the two probes were held apart from one another when hybridised to the single stranded DNA.[21]

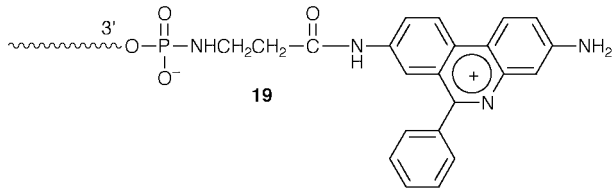
(ii) Manual labelling. As with 5′-labelling, manual attachment of a label at the 3′-terminus is most commonly achieved by conjugation with a suitable functional group introduced by automated synthesis. CPG derivatives can be purchased which, after post-synthesis cleavage and deprotection, lead to the introduction of amine, thiol, or phosphate functionality at the 3′-terminus. These derivatives contain essentially the same structures as the 5′-modifying agents already described, but with the necessary extra succinyl linkage for attachment to the CPG. These groups can then be conjugated with fluorochromes using reaction conditions identical to those already discussed for 5′-manual labelling. A CPG support that can be used for the introduction of a carboxylic acid at the 3′-terminus is also available from Glen Research, which can be converted to its NHS-ester *in situ* and reacted with amine-containing labels. This represents a useful reversal of the strategy for conjugation of labels *via* a 3′-amine group.

Mergny *et al*. used a manual labelling strategy in the preparation of 3′-labelled probes for use in fluorescence energy

transfer studies of hairpin structures in DNA fragments.[22] Oligonucleotides carrying a phosphorothioate group at the 3′-end were synthesised from a modified derivatized support, by a route involving the addition of a nucleoside 3′-phosphorothioate-triester to a disulfide bond. Coupling of the 3′-thiophosphate derivative with 4-(trifluoromethyl)-7-(ω-bromopropylamino)coumarin gave the fluorescent compound **18**. An



18



19

aminoalkylated derivative of ethidium was obtained by treating ethidium bromide with excess 3-(*N*-trifluoroacetyl)aminopropionyl chloride and trifluoroacetic anhydride. Synthetic oligonucleotides were converted to their cetyltrimethylammonium salts by precipitation with aqueous cetyltrimethylammonium bromide. The cetylammonium salts were then treated with the ethidium derivative, 2,2′-dipyridyl disulfide and triphenylphosphine in dimethyl sulfoxide, to give 3′-substituted ethidium oligonucleotides **19**.

2.3 Labelling at the nucleobases

Strategies for the introduction of fluorescent molecules *via* a modified nucleobase again fall into the categories of phosphoramidite and manual labelling, following synthesis with a suitably functionalised phosphoramidite bearing a group on one (or more) of the bases which can be conjugated with a dye derivative.

(i) Automated labelling. There are a limited number of commercially available phosphoramidites that carry a fluorescent tag on a modified base, such as structure **20**. However, phosphoramidites such as **6** based upon the (aminobutyl)propanediol skeleton can function as reagents for introduction of labels at sites internal to the oligonucleotide sequence in a fashion that mimics the positioning of a tag *via* a nucleobase. The three-carbon fragment separating maintains the internucleotide-phosphate distance, and the linker arm bearing the label is reasonably far away from the oligonucleotide backbone such that steric interactions do not hinder hybridisation or enzymatic manipulation processes. Duplex destabilisation is minimised with respect to oligonucleotides labelled using other methods.[23]

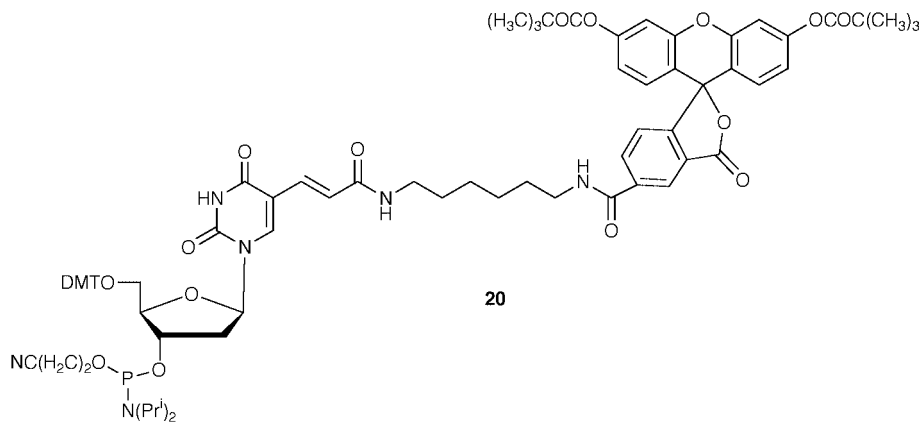
The preparation of fluorescently-modified nucleotides which function satisfactorily in automated synthesis has been, and continues to be, a productive area.[24,25] An interesting recent example is reported by Rosler and Pfleiderer who synthesised a series of condensed areno[*g*]lumazine derivatives **21**–**24** which can be regarded as structural analogues of thymidine.[26] These were used to synthesise lumazine-*N*-2′-deoxyribonucleosides, which were converted to both cyanoethyl phosphoramidites **25** and succinyl-CPG derivatives **26**.

Using the lumazine moiety as a probe of intermolecular interactions, the authors found that when incorporated at sites internal to the nucleotide sequence, the areno[*g*]lumazines conferred enhanced duplex stability on a self-complementary 18 base sequence. Interestingly, if the labels were placed within three bases of the 5′- or 3′-terminus, or if two labels were incorporated adjacent to one another, then duplex destabilisation was observed. These effects could have importance in the design of modified oligonucleotides for use in antisense DNA technology.

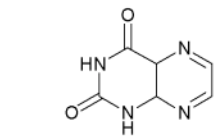
(ii) Manual labelling. Labels can be attached at amino or thiol functional groups tethered by a suitable linking chain to the base. Phosphoramidite reagents such as **27** which are commercially available can be used to introduce a primary amine group. Fluorochrome coupling methods are again performed using the appropriate ITC or NHS derivative. A compound (**28**) is also available (Glen Research) that introduces a carboxylate group on a modified base after cleavage and deprotection, which again offers the opportunity for conjugation of labels which carry a primary amine group.

Meyer and Hanna reported the preparation of a novel phosphoramidite (**29**) that can be used for introducing a thiol-modified uridine into an oligonucleotide sequence, using standard automated synthesis protocols.[27] The 2,4-dinitrophenyl group can be removed after synthesis to reveal the free –SH moiety, which can be reacted with iodoacetamide derivatives of fluorescein (or other species such as spin-label molecules).

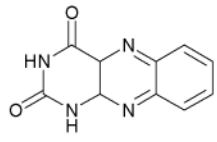
In a number of reports, Mathies and co-workers have used a base-labelling approach to prepare oligonucleotides bearing a fluorescein derivative at the 5′-end and a fluorochrome-labelled thymidine base at different positions within the sequence, using phosphoramidite **27b**.[7,28] Fluorochromes were attached to this modified base using the NHS ester strategy. The doubly labelled oligonucleotides have been studied extensively in FRET-based sequencing and detection. The 5′-fluorescein acts as a common donor, and other fluorescein and rhodamine derivatives linked



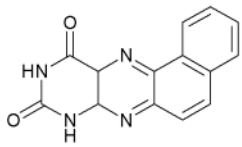
20



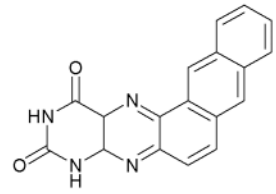
21



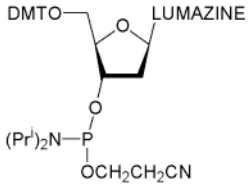
22



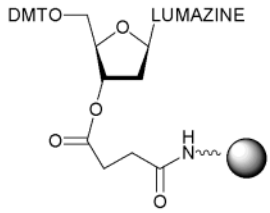
23



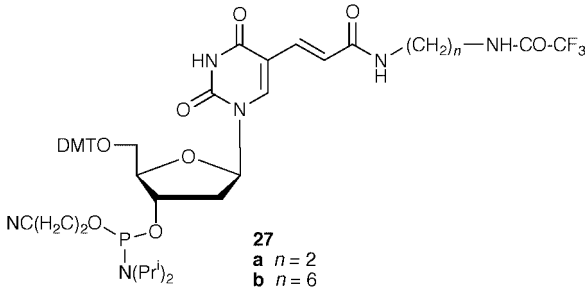
24



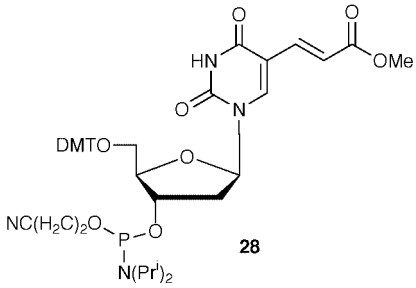
25



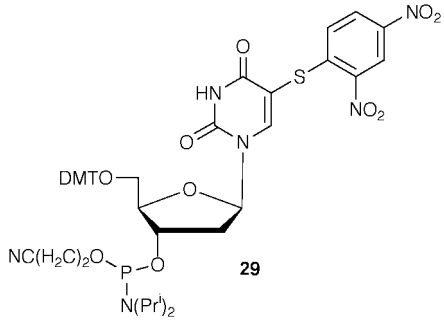
26



27
a $n = 2$
b $n = 6$



28



29

through the modified thymidine act as acceptor. By appropriate positioning of the acceptor labels with respect to the donor label, four separate oligonucleotide primers for use in DNA sequencing reactions have been generated with a common excitation wavelength and four separate emission wavelengths. These primers and the DNA sequencing fragments generated using them display similar electrophoretic properties to the singly labelled primers commonly used in 'four-colour' sequencing, but display fluorescence intensities two to fourteen times greater, depending on the distance between donor and acceptor. A ten base spacing appears to be the optimum, and the increased sensitivity means that DNA sequencing can be performed using only a quarter of the usual amount of template DNA.

2.4 Labelling of oligonucleotides at the phosphate backbone

A disadvantage of using modified bases to tether fluorescent reporter groups is that the tether itself can cause steric hindrance with other atoms in the oligonucleotide sequence. There are a number of potential advantages with backbone labelling in comparison with end-labelling and base-labelling procedures. The use of internal phosphodiester residues instead of terminal phosphates means that the desired functionality can be placed at virtually any position within the sequence. Furthermore, since the phosphodiester residues are not involved in interstrand DNA base pairing, the presence of a label on the DNA backbone may not drastically alter the stability or structure of duplex nucleic acids. The advent of 'antisense' DNA technology has led to the widespread synthesis and use of phosphorothioate derivatives (Fig. 4) which are resistant to nucleases.[29] The majority of

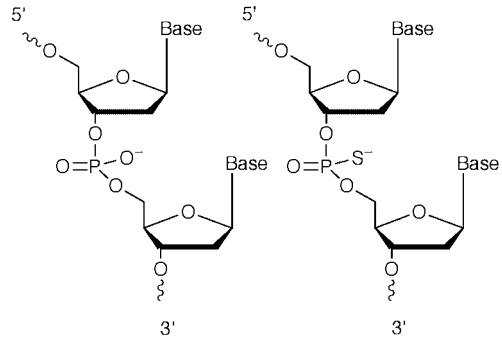


Fig. 4 Comparison of the structures of a phosphodiester linkage (left) and phosphorothioate linkage (right).

labelling strategies at the phosphate backbone involve modification of the sulfur group of a phosphorothioate residue with a sulfur-reactive probe such as a haloacetamide, aziridine, sulfonamide or bromo-α,β-unsaturated carbonyl.[30]

2.5 Labelling of oligonucleotides at ribose

This is the least common of all strategies for oligonucleotide labelling, as labelling at the other sites already mentioned addresses the majority of needs in terms of ease and subsequent use of labelled materials. However, some approaches to labelling by this strategy have been published to address particular analytical needs.[31]

For example, Defrancq and co-workers devised a new one-step derivatisation of oligonucleotides at an abasic hemiacetal site.[32] The abasic position was treated with a rhodamine label containing an oxyamino group. The oxime ether linkage formed was stable under physiological conditions, and did not require a subsequent chemical reduction step (Fig. 5). Trimer **31a** [d(GXA)] and undecamer **31b** [d(CGCACXCACGC)], in which X represents the abasic site, were prepared. Reaction of

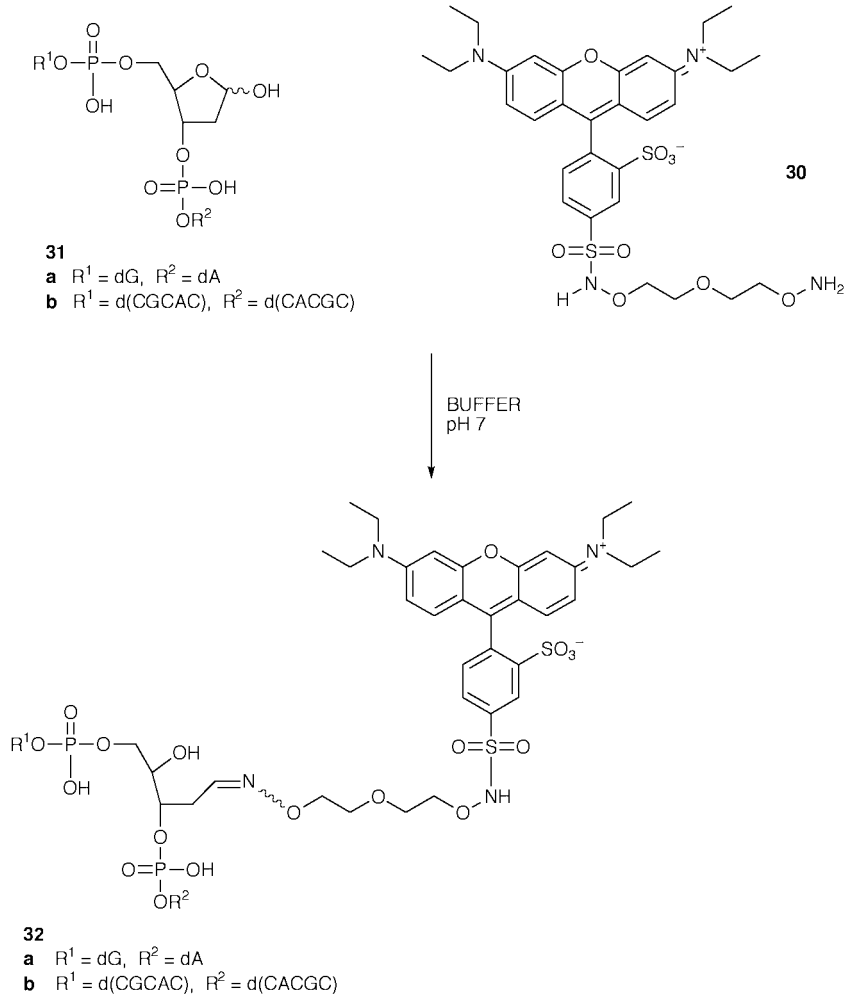


Fig. 5 Abasic site labelling using an aminooxy derivative of a fluorochrome.

the fluorescent probe **30** with oligonucleotides **31a** and **b** at pH 7 gave the conjugated products **32a** and **32b** respectively. This model system demonstrated that the labelling technique could be used for quantifying, by fluorimetry, the abasic content of larger nucleic acids.

3 Nucleic acid labelling

As has been described, single-step oligonucleotide labelling can be performed with precision at a particular point in an oligonucleotide by appropriate programming of an automated synthesiser. Equal precision can be accomplished in the labelling of native nucleic acids, restriction fragments or amplified products, but relies on the specificity of enzymes, or targeting of a particular feature inherent in the nucleic acid structure. Many manipulations originally developed for radioisotopic labelling of nucleic acids have been adapted for fluorescent labelling. As mentioned earlier, the literature supplied by Molecular Probes[12] is an excellent source of information on applications and references to the primary literature on all aspects of manual fluorescent labelling, and this applies to nucleic acids as well as oligonucleotides. Labelling approaches can be categorised as enzymatic labelling and chemical labelling.

3.1 Enzymatic labelling

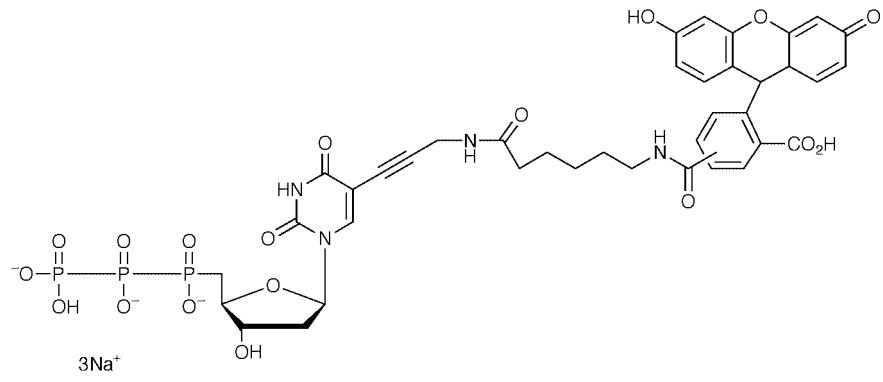
Enzymatic modification can take place at the 5′- and 3′-terminus, and also at sites internal to the nucleic acid sequence. Each of these modifications uses a different enzyme but can utilise the same substrate, a modified deoxyuridine triphosphate, a typical example being **33**. Such compounds are available from a number of companies including Roche Molecular Biochemicals (Mannheim, Germany; http://biochem.roche.com), and Molecular Probes. The long linker arm reduces interaction between the label and the nucleic acid. Other modified dUTPs based on rhodamine, coumarin and other fluorochromes are also available.

For 5′-modification, the enzyme polynucleotide T4 kinase is used. For introduction of the label at the 3′-terminus, deoxynucleotidyl transferase is used. For internal site labelling, most commonly used when probes for *in situ* hybridisation are to be prepared, the process of ‘nick’ translation is used. This method uses two enzymes, DNAse and a DNA polymerase. The DNAse cuts, or ‘nicks’ one strand of a double stranded DNA molecule, then labelled nucleotides are incorporated as the polymerase proof-reads the nicked site. The modified dUTP replaces dTTP in this process. Average DNA lengths of 200–500 base pairs are produced and the labelled dUTP is introduced at a rate of about one every 30–50 base pairs, meaning that multiple labelling can occur.

For each of these methods, the label is not introduced quantitatively, so there will be a mixture of labelled and unlabelled material in the product mixture. Excess labelled material can be removed by gel filtration chromatography, or by precipitation of the DNA.

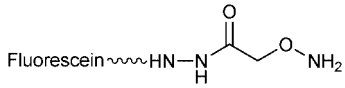
3.2 Chemical labelling

Many of the methods described for manual oligonucleotide modification are also applicable to nucleic acids, for example,



33

conversion of the 5′-phosphate to a phosphoramidate using a diamine, the free-NH_2 group of which can be modified with an NHS or ITC fluorochrome derivative. Targeting of abasic sites in DNA can also be carried out using probes similar to those described earlier, which carry an aminooxy functionality. In a recent report, a fluorescein-conjugated aminooxyacetylhydrazide derivative (**34**) was prepared, which was reacted with



34

open chain aldehydes generated at abasic sites in DNA through depurination, forming a stable oxime bond.[33] In this work, the authors found that abasic sites could be almost quantitatively labelled with the fluorescent agent, and the fluorescent quenching interaction of two fluorochromes on adjacent sites could be used to measure the total abasic content of a DNA sample. This multiple labelling technique has now been applied to detection of base mismatches in double stranded DNA.[34]

A multiple labelling applicable to both RNA and DNA can be achieved by first treating cytidine residues within the sequence with sodium bisulfite to form a sulfonate, which can subsequently be coupled with a hydrazide or diamine. This method was first reported some thirty years ago, and in recent years has been adapted for fluorescent labelling by using fluorochromes bearing appropriate hydrazide or amine groups for reaction with the sulfonate. Alternatively, a two-step procedure can be used if the necessary derivative is not available. For example, the preparation of pyrene labelled DNA probes for use in detection of complementary DNA sequences has been reported using this approach. Pyrene labels were attached to a small fraction of cytidine residues within a poly(C) fragment that had been modified with ethylenediamine groups through the bisulfite reaction, and pyrenesulfonyl chloride was used as the reactive label.[35]

With many of these chemical modification methods, separation of the unincorporated label can be more labour intensive compared to enzyme modification. The labels used tend to show higher non-specific affinity for nucleic acids, and often the only satisfactory way to completely remove them is to purify the labelling mixture by polyacrylamide gel electrophoresis (PAGE). Labelled nucleic acid can be visualised as a fluorescent band under UV light, excised, and homogenised to release it into solution. As with enzymatic modification procedures, chemical modification is never quantitative so a mixture of labelled and unlabelled material will in fact be isolated from the gel, as there is not a significant difference in their size for them to be resolved by PAGE.

4 Conclusions

This article has given an overview of the common strategies currently used for the introduction of a fluorescent label at a specified point in an oligonucleotide or nucleic acid molecule, and has described how the labelled materials have been used to develop new assays or probe molecular interactions in ways not previously considered. Unlike the development of most other methods used in the field of molecular biology, which have been driven by the researcher at the bench, the introduction of fluorescent methods was initiated by equipment manufacturers looking for further commercial benefit. However, researchers took this as a cue to explore the benefits of fluorescence more widely, leading to the development of much new and innovative chemistry and applications, which has been made more readily available by specialist reagent suppliers. The advances in this area now underpin many of the exciting developments in clinical diagnosis that will be applied routinely in years to come.

References

1 L. M. Smith, J. Z. Sanders, R. J. Kaiser, P. Hughes, C. Dodd, C. R. Connell, C. Heiner, S. B. Kent and L. E. Hood, *Nature (London)*, 1986, **321**, 674.
2 E. S. Mansfield, J. M. Worley, S. E. Mckenzie, S. Surrey, E. Rappaport and P. Fortina, *Mol. Cell. Probes*, 1995, **9**, 145.
3 H. A. Erlich, D. Gelfand and J. J. Sninsky, *Science*, 1991, **252**, 1643.
4 F. F. Chehab and Y. W. Kan, *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 9178.
5 M. J. Yang and J. Cao, *Prog. Biochem. Biophys.*, 1998, **25**, 333.
6 R. J. Lipshutz, S. P. A. Fodor, T. R. Gingeras and D. J. Lockhart, *Nat. Genet.*, 1999, **21**, 21.
7 J. Ju, I. Kheterpal, J. R. Scherer, C. Ruan, C. W. Fuller, A. N. Glazer and R. A. Mathies, *Anal. Biochem.*, 1995, **231**, 131, and references therein.
8 S. Tyagi and F. R. Kramer, *Nature Biotechnol.*, 1996, **14**, 303.
9 C. T. Wittwer, M. G. Herrmann, A. A. Moss and R. P. Rasmussen, *Biotechniques*, 1997, **22**, 130.
10 T. Brown and D. J. S. Brown, in *Oligonucleotides and Analogues. A Practical Approach*, ed. F. Eckstein, IRL Press, Oxford, 1991, p. 1.
11 R. L. Letsinger and T. Wu, *J. Am. Chem. Soc.*, 1995, **117**, 7323.
12 R. P. Haugland, *Handbook of Fluorescent Probes and Research Chemicals*, *6th edn.*, Molecular Probes, Eugene, 1996.
13 M. L. Metzker, J. Lu and R. A. Gibbs, *Science*, 1996, **271**, 1420.
14 Y. Jenkins and J. K. Barton, *J. Am. Chem. Soc.*, 1992, **114**, 8736.
15 P. G. Sammes and G. Yahioglu, *Nat. Prod. Rep.*, 1996, **13**, 1.
16 A. Chan, E. Diamandis and M. Krajden, *Anal. Chem.*, 1993, **65**, 158.
17 J. Coates, P. G. Sammes, G. Yahioglu, R. M. West and A. J. Garman, *J. Chem. Soc., Chem. Commun.*, 1994, 2311.
18 K. Shinozuka, Y. Seto, H. Kawata and H. Sawai, *Bioorg. Med. Chem. Lett.*, 1993, **12**, 2883; K. Shinozuka, Y. Seto and H. Sawai, *J. Chem. Soc., Chem. Commun.*, 1994, 1377.
19 S. B. Rajur, J. Robles, K. Wiederholt, R. G. Kuimelis and L. W. McLaughlin, *J. Org. Chem.*, 1997, **62**, 523.
20 B. Mullah and A. Andrus, *Tetrahedron Lett.*, 1997, **38**, 5751.

21 N. Ota, K. Hirano, M. Warashina, A. Andrus, B. Mullah, K. Hatanaka and K. Taira, *Nucleic Acids Res.*, 1998, **26**, 735.
22 J.-L. Mergny, A. S. Boutorine, T. Garestier, F. Belloc, M. Rougee, N. V. Bulychev, A. A. Koshkin, J. Bourson, A. V. Lebedev, B. Valeur, N. Thuong and C. Helene, *Nucleic Acids Res.*, 1994, **22**, 920.
23 P. S. Nelson, M. Kent and S. Muthini, *Nucleic Acids Res.*, 1992, **20**, 6253.
24 S. L. Beaucage and R. P. Iyer, *Tetrahedron*, 1993, **49**, 1925.
25 W. T. Markiewicz, G. Groger, R. Rosch, A. Zebrowska, M. Markiewicz, M. Klotz, M. Hinz, P. Godzina and H. Seliger, *Nucleic Acids Res.*, 1997, **25**, 3672.
26 A. Rosler and W. Pfleiderer, *Helv. Chim. Acta*, 1997, **80**, 1869.
27 K. L. Meyer and M. M. Hanna, *Bioconjugate Chem.*, 1996, **7**, 401.
28 J. Ju, C. Ruan, C. W. Fuller, A. N. Glazer and R. A. Mathies, *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 4347.
29 E. Uhlmann and A. Peyman, *Chem. Rev.*, 1990, **90**, 543.
30 A. J. Stewart, C. Pichon, P. Midoux, R. Mayer, M. Monsigny and A. C. Roche, *New J. Chem.*, 1997, **21**, 87.
31 K. Yamana, Y. Ohashi, K. Nunota and H. Nakano, *Tetrahedron*, 1997, **53**, 4265.
32 D. Boturyn, E. Defrancq, V. Ducros, C. Fontaine and J. L'Homme, *Nucleosides, Nucleotides*, 1997, **16**, 2069.
33 G. M. Makrigiorgos, S. Chakrabarti and A. Mahmood, *Int. J. Radiat. Biol.*, 1998, **74**, 99.
34 G. Maulik, S. Botchway, S. Chakrabarti, S. Tetradis, B. Price and G. M. Makrigiorgos, *Nucleic Acids Res.*, 1999, **27**, 1316.
35 J. Yguerabide, E. Talavera, J. M. Alvarez and M. Afkir, *Anal. Biochem.*, 1996, **241**, 238, and references therein.

Review a905230e