Published in final edited form as:

Nat Chem. 2017 November ; 9(11): 1043–1055. doi:10.1038/nchem.2859.

Fluorescent nucleobases as tools for studying DNA and RNA

Wang Xu, Ke Min Chan, and Eric T. Kool*

Department of Chemistry, Stanford University, Stanford, California 94305, USA

Abstract

Understanding the diversity of dynamic structures and functions of DNA and RNA in biology requires tools that can selectively and intimately probe these biomolecules. Synthetic fluorescent nucleobases that can be incorporated into nucleic acids alongside their natural counterparts have emerged as a powerful class of molecular reporters of location and environment. They are enabling new basic insights into DNA and RNA, and are facilitating a broad range of new technologies with chemical, biological and biomedical applications. In this Review, we will present a brief history of the development of fluorescent nucleobases and explore their utility as tools for addressing questions in biophysics, biochemistry and biology of nucleic acids. We provide chemical insights into the two main classes of these compounds: canonical and non-canonical nucleobases. A point-by-point discussion of the advantages and disadvantages of both types of fluorescent nucleobases is made, along with a perspective into the future challenges and outlook for this burgeoning field.

Fluorescent nucleobases are chemically modified DNA and RNA analogues that not only retain their chemical and biological functionalities, such as stacking, base pairing and enzyme incorporation, but also empower improved fluorescence properties for the analysis of nucleic acids. These molecules have emerged as an extraordinarily useful category of chemical and biological tools for the molecular-level understanding of nucleic acid structures, activities, locations and interactions¹. In 1969, Stryer reported a seminal study on the emissive compounds formycin and 2-aminopurine (2AP)². This research demonstrated that these fluorescent analogues can facilitate studies of nucleic acids conformation and interactions, and has led to the creation of a broad palette of fluorescent nucleobases. As the natural nucleobases are essentially non-fluorescent, significant modifications to nucleobase structure are required before they can be broadly useful as emissive tools. These modifications can either be made with an eye to preserving Watson–Crick-like pairing ability (that is, retaining ‘canonical’ pairing), or more dramatic changes to the nucleobase architecture can be made, resulting in non-canonical designs.

As part of their function as universal hereditary materials, a central feature of the canonical nucleobases in DNA and RNA is their ability to pair with each other: adenine with thymine/uracil and guanine with cytosine, through specific hydrogen-bonding interactions (Fig. 1a). This pairing preserves the standard purine/pyrimidine base pair architecture, resulting in base pairs with regular geometries. This conserved pairing structure has allowed organisms

* kool@stanford.edu.

Competing financial interests

The authors declare no competing financial interests.

to develop machinery to recognize the genetic information generically and pass it on through generations³. With careful design, fluorescent nucleobases can be crafted to retain some or most of these properties: hydrogen-bonded base pairing, enzyme and protein recognition, and purine/pyrimidine pairing architecture (Fig. 1b,c).

On the other hand, non-canonical nucleobases have also been developed for chemical and biological applications. In this case, the nucleobases are not necessarily purines or pyrimidines, but rather have more greatly varied structure since they are not constrained in their architecture or base-pairing abilities. This relative lack of constraints allows the non-canonical nucleobases to have more widely varied emissive properties, but also can limit their ability to be recognized by proteins and enzymes (Fig. 1d).

Here we describe the designs and properties of fluorescent nucleobases. Within the canonical class of fluorescent nucleobases, chemical modifications have been employed on both purine and pyrimidine structures. Ring fusion, ring structure modification as well as addition of ring substituents have all been studied and their optical properties examined. For non-canonical fluorescent nucleobases, a broad variety of molecular sizes and shapes have been constructed. In exchange for limited base-pairing ability, non-canonical fluorescent nucleobases can exhibit stronger π -stacking properties and optical versatility. Both categories of nucleobases remain broadly useful; we describe the similarities and differences of these two classes of tools, and their relative advantages and disadvantages for specific applications.

This Review is not meant to be comprehensive, but rather to illustrate the chief types of designs and properties engendered by various classes of fluorescent analogues. We limit ourselves here to a discussion of nucleobases in which the functional unit of the base itself is part of the fluorescence emitting structure; we will not describe nucleotide bases with fluorescent labels attached via a non-emissive linker^{4,5}. Although the latter class of molecules is broadly useful, fluorescent nucleobases offer some important advantages, including: (1) placement of the chromophore directly into the DNA/RNA helix, thus serving as a sensitive and spatially fixed probe of helix structure and interactions; and (2) at least in some cases, a lack of side chains that might otherwise perturb structure and protein–enzyme interactions^{6,7}.

Fluorescent nucleobase development

The first report of a fluorescent nucleobase dates back a half century, when Stryer reported fluorescence studies of aminated nucleobases: 2AP, formycin and 2,6-diaminopurine². Among these, 2AP (Fig. 2) emerged as a highly useful tool in nucleic acids research⁸. A few years later, Leonard discovered another brightly fluorescent purine analogue, ethenoadenine, which cannot form base pairs due to addition of two carbons, forming a ring on the previous pairing edge⁹. These molecules have been extensively studied and applied¹⁰. In the 1990s, pyrimidine fluorescent analogues began to emerge, with notable examples including m⁵K of McLaughlin¹¹ and the pteridines of Hawkins¹² (Fig. 2). The first C-glycosidic fluorescent nucleobases were synthesized around this time as well, as Kool reported the direct attachment of pyrene and other hydrocarbons to deoxyribose¹³. In 1998, Moreau expanded

the pyrimidine structure by fusion with phenyl rings, reminiscent of Leonard's earlier work, and reported BgQ as a larger-than-natural nucleobase¹⁴. In the 2000s, work in this field began to expand rapidly. Notable advances include the development of base-discriminating fluorescent nucleobases by Saito and Okamoto^{15–17}. Wilhelmsson synthesized a thiol-appended cytidine analogue (tC) at the same time¹⁸, and Hocek introduced extended fluorescent nucleobases as a facile way to modify the electronics of canonical nucleobases via conjugated linkers¹⁹. In the late 2000s, Sekine designed pyrimidopyrimidine base analogues²⁰ while Tor significantly expanded the isomorphous fluorescent nucleobase pool by designing a series of thieno-appended analogues^{21–25}. To date, hundreds of fluorescent nucleobases have been introduced, all of which exhibit distinct base stacking and pairing abilities or emission profiles²⁶. The work has not slowed; to the contrary, the past few years of research has resulted in more fluorescent nucleobase analogues than the previous decades combined.

Canonical fluorescent nucleobase design

Canonical nucleobase analogues as defined here are constrained by (1) the purine or pyrimidine underlying architecture and (2) retention of at least two Watson–Crick hydrogen-bonding groups, allowing them to potentially form hydrogen-bonded pairs with a complementary base. As a result of these constraints, the positions of modification are limited to the regions of the bases that do not block this pairing potential. Moreover, to design and tune the fluorescence properties, the electronics of the emissive π -system must take into account the underlying electronic arrangement of the purine or pyrimidine skeleton. The payoff for operating under these constraints can be the enhanced potential to base pair or act as an enzyme substrate.

For applications of fluorescent nucleobases *in vivo*, it is often desirable to modulate their excitation and emission wavelengths towards the red. This strategy helps avoid background emission by native compounds in the cell, as well as cytotoxicity arising from ultraviolet irradiation. Redshifted optical profiles can be achieved by extending π -systems and incorporating heteroatoms. However, these modifications may result in disrupted base pairing and distorted helical conformations. Thus careful structural design is important; structure–activity relationship studies and molecular simulations can be useful in guiding the designs.

Purine architecture modifications

Fluorescence-enhancing modifications of the purine scaffold mainly occur on the five-membered ring, namely the 8 position of adenine and the 7 and 8 positions of guanine (Fig. 3a,b). The 1, 2 or 6 positions of the six-membered ring are in most cases unmodified in the canonical class of fluorescent nucleobases because base pairing occurs at these sites. However, the 2 position of adenine has also seen several examples of extensions. To compare and discuss the purine modification patterns, we divide the modified nucleobases into four main categories: purine ring structure modifications; extended fluorescent scaffolds via conjugated linkers; purine substituent modifications; and purine ring fusions.

Modification of the purine ring is well exemplified by the pteridines, a class of nucleobases fusing pyrimidine and pyrazine six-membered rings (Fig. 3c). Both absorption and emission wavelengths of the pteridines exhibit significant bathochromic shifts to an extent of 100 nm compared with those of natural nucleobases, while their quantum yields increase by four orders of magnitude^{27–29}. The increased conjugation of the new six-membered ring lowers the highest occupied molecular orbital–lowest unoccupied molecular orbital gap between the ground and excited states³⁰. Further modifications can be conducted at the 1 and 8 positions of pteridines as 3-MI, 6-MI, 6MAP and DMAP all have methyl groups at one of these locations (Fig. 3c). Pteridines commonly experience quenching effects when near canonical nucleobases, especially purines. This can be limiting if bright fluorescence is required, whereas it can be of value if reporting on local structure is the goal³¹.

Extending the conjugated scaffold of purine structures is another important category of purine modification, with great potential to expand the emission spectrum. Most of the extensions are made at 2 or 8 positions because they are synthetically easy to achieve, and because base pairing can still occur in some cases. The pyrene fluorophore appended on adenine A^{PY}, A^P and G^P has allowed these nucleobases to be excited at long ultraviolet to short blue ranges (380 to 420 nm) while emitting blue light (450 to 480 nm) (Fig. 3d)^{32–34}. In some rarer cases, the 7 position can also be appended with fluorescence extensions, with replacement of the nitrogen at the 7 position with carbon to preserve aromaticity. One such example is a ruthenium bi-pyridyl fluorophore attached to the 7 position of guanine through an alkyne linker (Fig. 3d)¹⁹. In other cases, a phenyl linker and a diene linker are also reported, showing some versatility of conjugated linker selection (Fig. 3d, ABG–DABG)^{35–37}. These nucleobases preserve the conjugation between the extended moiety and the original ring electronic system, thus allowing the base pairing and local environment to affect the optical properties of the whole fluorophore. This environmental sensitivity, arising from the internal charge transfer between the extension and the natural nucleobase³⁸, results in redshifts and quantum yield changes³⁹. For example, A^P emits at 430 nm in MeOH with a quantum yield of 0.73, whereas when it is incorporated in a DNA strand, the emission wavelength redshifts to 450–480 nm with a substantial drop in quantum yield³².

Purine substituent modification is currently the most prevalent category of fluorescent nucleobase design. The modifications focus either on replacing the original purine substituents with new functional groups, or appending substituents on the purine core structure. As an early and long-useful modification, 2AP moves the 6-amino group of adenine to the 2 position, whereas DAP (2,6-diaminopurine) adds a second amino group at the 2 position of adenine (Fig. 3e). 2AP exhibits a ca. 50 nm redshift (relative to adenine) whereas DAP shows a 30 nm redshift². Among guanine analogues, the carbon at the 8 position can be replaced by nitrogen (8-AzaG)⁴⁰, whereas the nitrogen at the 7 position can be replaced by sulfur (4-thieno[3,2-*d*]-R) (Fig. 3e)⁴¹. 8-AzaG exhibits an environmentally sensitive emission band at 347 nm in buffer and 377 nm when incorporated into an RNA strand⁴². 4-thieno-R exhibits 351 nm emission in water. Although both nucleobases emit at similar wavelengths, their excitation maxima differ substantially. 8-AzaG absorbs at 256 nm in buffer, whereas 4-thieno-R absorbs at 294 nm, with the sulfur-substituted nucleobase requiring less energy to excite⁴³.

The optical profile of purines can also be readily tuned by appending small functional groups. For example, by adding a vinyl group to the 8 position of either adenine (8vA) or guanine (8vG), the emission can redshift to 382 nm or 400 nm, and result in remarkable quantum yields of 0.66 and 0.72, respectively (Fig. 3e)^{44,45}. Another important modification is the addition of external non-emissive aromatic structures to the natural nucleobases. The majority of modifications happen at the 8 position with some exceptions at the 6 position. When a thiophene group is placed at the 6 position of guanine, the nucleobase's absorption wavelength increases to 348 nm whereas emission increases to 434 nm (Fig. 3e, s)⁴⁶. Similarly, when the furan group is appended at the 8 position of adenine, its emission wavelength increases to 374 nm, whereas phenol or indole groups redshift the guanine emission to 390 nm (Fig. 3e, 8-Furan-A, 8-Phenol-G and 8-Indole-G)^{47–49}. Evidence (primarily environmental sensitivity) suggests that the internal charge transfer mechanism functions to provide enhanced quantum yields and redshifts in these cases. Although several of these substitutions are appealingly small and simple, they can affect pairing: for instance, 8-substituted purines have a tendency to flip to the *syn* conformation, inhibiting Watson–Crick pairing⁵⁰.

The last category of purine modification is ring fusion. Most such fusions are carried out at the 7 and 8 positions of the five-membered ring, to avoid perturbing base pairing (Fig. 3f). Multiple such fusion patterns with varied ring sizes or substituents have been designed, such as benzene or naphthalene groups. One important exception, by Leonard, is the benzopurine design, which inserts a phenyl group in between the five- and six-membered rings (Fig. 3f, xA and xG)⁵¹. Notably, introduction of tricyclic structure significantly expands the aromatic conjugation, thus leading to dramatic redshifts and increase of fluorescence intensity. As reported, lin-benzoadenine and lin-benzoguanine emit at 393 nm and 413 nm in buffer⁵¹, respectively, whereas ^{MDI}, ^{MDA} and ^{MDA} emit around 380–420 nm (Fig. 3f)^{17,52}. xA and xG are both strongly fluorescent with quantum yields above 40% (ref. 53), and ribo (RNA) variants of these two are also known⁵⁴. ^{MDI}, ^{MDA} and ^{MDA} are less fluorescent, with quantum yields being 0.12 or lower, possibly due to the rotational energy dissipation by methoxy group or to the structural distortion from the tetracyclic ring.

Pyrimidine scaffold modifications

In contrast to the purine analogues, there are fewer modification sites on the pyrimidine structure, because of the smaller size and greater simplicity of the ring system. Modifications concentrate on the 4, 5 and 6 positions, leaving the 2 and 3 positions open for pairing (Fig. 4a,b). Similar to the purine fluorescent analogues, the pyrimidine analogues can be divided into three categories: modified pyrimidine substituents; extended fluorescent scaffolds by use of conjugated linkers; and ring-fused pyrimidine ring systems.

Modifications of the pyrimidine substituents represent the largest group of fluorescent nucleobases based on pyrimidine structure. One of the earliest examples, m⁵K, removes the amino group at the 4 position of cytosine and adds a methyl group on the 5 position to afford a simple isomorphous structure (Fig. 4c). The absorption maximum only slightly redshifts (271 to 280 nm), but the emission maximum increases dramatically from 324 to 400 nm (ref. 11). The formation of a weak push–pull electronic structure, with methyl serving as a donor

and carbonyl as acceptor, is the cause of the redshift. In another strategy, by appending a furan group to the 5 position, the weakly fluorescent nucleobase C^{FU} can be produced with absorption maximum at 310 nm and emission maximum at 443 nm (Fig. 4c). Free rotation around the single bond between furan and cytosine likely dissipates the excited state energy, resulting in the low quantum yield of 0.01 (ref. 47). Other substituents employed on the 5 position in fluorescent uracil include thiophene, 4-methoxyphenyl and chlorinated phenyl groups (Fig. 4c, uracil derivatives)⁵⁵. Among these derivatives, the use of push–pull electronic designs yields especially useful results; for example, the 4-methoxyphenyl-appended nucleobase emits at 444 nm, whereas the phenyl-substituted nucleobase emits at a shorter wavelength (403 nm)^{55,56}. Such cases make it clear that addressing the specific electronic structure of the nucleobase is important in fluorophore design.

Further modifications of pyrimidine substituents can lead to a diversity of fluorescent nucleobases. C^{hPP} has a urea-containing six-membered ring appended on the 4 and 5 positions of cytosine to engender greater conjugation (Fig. 4d, C^{hPP}). Its absorption maximum increases from 271 nm of cytosine to 300 nm, whereas emission wavelength increases from 324 to 360 nm. The quantum yield of C^{hPP} shows an increase of four orders of magnitude over the natural base, to 0.12. It is generally the case that increased conjugation, combined with the rigidity induced by the use of ring structure, contributes greatly to improving optical properties⁵⁷. The ring extension principle can be amended by further substitution: for instance, when one more phenyl group is appended to the six-membered ring containing a sulfur atom, the important analogue tC is produced, exhibiting yet greater redshifts (absorption 375 nm, emission 500 nm) with a better quantum yield of 0.17 (Fig. 4d)¹⁸. This basic structure is also incorporated into the well-known ‘G-clamp’ analogue, which pairs especially strongly with cytosine (Fig. 4d)⁵⁸. Further functionalization of the phenyl substituent can result in additional fluorescent compounds, such as C^f that absorbs at 360 nm and emits at 450 nm (Fig. 4d)⁵⁹. BPP, another phenyl-appended pyrimidine derivative, absorbs at 347 nm and emits 390 nm, whereas emitting less brightly than tC (quantum yield = 0.04, Fig. 4d)¹⁶. Electronic effects can also be examined using the C^{hPP} derivatives. For example, when the R1 group on the C^{hPP} derivatives changes from electron-donating groups (OMe, SMe) to electron-withdrawing ones (CN, SO₂Me, NMe³⁺), the electron density on the fluorescent cytosine analogues decreases, and higher quantum yields are achieved (Fig. 4d, C^{hPP} derivatives)⁶⁰.

Extending fluorescent modules into the pyrimidine framework via conjugated linkers is another important approach to modify pyrimidines. The use of ethynyl and heterocyclic linkers allows a diversity of chromophores to be attached to the natural nucleobases, including pyrene, fluorenone or boron-dipyrromethene (BODIPY) dyes (Fig. 4e, 5-ethynylpyrene U)^{61,62}. To date, nearly all the extension has taken place at the 5 position of the pyrimidine ring to avoid perturbing the base pairing and to take advantage of the conjugation with the N1 nitrogen and 5,6 double bond. When the substituents are changed on the extended fluorophore, electronic effects can be studied. For instance, $5^{EP}U$ absorbs at 320 nm while emitting at 400 nm; when the *para*-position of phenyl ring is substituted by a donating dimethylamino group ($5^{EDMA}U$), the absorbance redshifts to 330 nm and the emission redshifts to 450 nm. On the other hand, blue shifts occur in the case of electron-withdrawing groups attached to phenyl ($5^{EBN}U$) (Fig. 4e, $5^{EP}U$, $5^{EDMA}U$ and $5^{EBN}U$)¹. In

the case of 5-ethynylpyrene U, the nucleobase enjoys the optical properties of pyrene, but with a redshifted absorption at 392 nm and emission at 400 and 424 nm (ref. 61). A more recent triazole linker not only allows extended conjugation, but also enables the convenience of using biorthogonal chemistry for the attachment of fluorescent modules (Fig. 4e, triazole appended cytosine)⁶³. It can be quite useful in the case of post-synthesis or post-transcriptional modification of the nucleic acid strands. However, in the currently reported cases of triazole-appended cytosine, only very weak fluorescence emission is observed, possibly due to the presence of bond rotations and/or the electron-deficiency of the triazole ring. Thus, the thiophene attached nucleobase absorbs at 330 nm and emits at 375 nm, with a quantum yield of 0.0019; the fluorenone attached nucleobase absorbs at 320 nm and emits at 385 nm, with a quantum yield of 0.0055. Given the low brightness of these compounds, further optimization would be helpful to enhance their practical application⁶³.

The last category of pyrimidine modification is ring fusion, which can be carried out on either the 5 and 6 positions or the 4 and 5 positions without adversely perturbing base pairing. The added ring can greatly influence or improve the fluorescent properties of the natural nucleobase. For instance, when thiophene is fused to the 5 and 6 positions of uracil, the new nucleobase (thieno-dU) absorbs at 304 nm and emits at 412 nm, with a robust quantum yield of 0.48 (Fig. 4f, thieno[3,4-*d*]-U)²³. In addition, when naphthalene is fused to thymine, BgQ exhibits remarkable fluorescence properties with 0.82 quantum yield at the emission maximum of 434 nm (Fig. 4f). Compared with another naphthalene fused purine analogue (NDA, Fig. 3f), which exhibits non-planar structure and therefore emits poorly with a quantum yield less than 0.1, the strong fluorescence of BgQ may be attributed to the planar and rigid tricyclic structure¹⁴. xC is an example of a ring-fused C-glycosidic nucleobase analogue, which can be excited at both short ultraviolet (260 nm) or long ultraviolet (320–330 nm) wavelengths while emitting at 388 nm with a quantum yield of 0.52 (Fig. 4f, xC). Notably, the xDNA series (of which xC is one member) can be copied by DNA polymerases in live bacterial cells, indicating their biocompatibility^{53,64}. The 4 and 5 positions of pyrimidines can also be readily expanded with pyr-role fusions. For example, pC displays a moderate quantum yield of 0.2 at 460 nm emission, when excited by 350 nm long ultraviolet light (Fig. 4f, pC)⁶⁵. The thiophene variant exhibits even greater red-shifted absorption (370 nm) and emission (471 nm) with enhanced quantum yield of 0.42 (Fig. 4f, thiophen-2-yl pC). Further substituents can be added to the thiophene ring⁶⁶. Finally, very recently, a dimethylamino-phenyl group was fused to the 5 and 6 positions of thymine, with the capability of chelating with mercury through electrostatic interactions with another thymine (Fig. 4f, ^{DMA}T). The metal-bridged base pair exhibits higher kinetic stability than natural base pairing, and the resulting fluorescent base–metal pair can be used to explore transition mercury metabolism in live cells⁶⁷.

Design of non-canonical fluorescent nucleobases

Non-canonical fluorescent nucleobases further break the boundaries of natural nucleobases by avoiding the need to adhere to the molecular skeleton of purines and pyrimidines, and even the base pairing between them. By bypassing these structural limitations, researchers can engender a broad range of fluorescence properties. First, by avoiding the limits on purine/pyrimidine size, one can generate nucleobase analogues that exhibit emission further

to the red, thus better avoiding toxicity and background emission from ultraviolet excitation. Second, some well-known aromatic hydrocarbons exhibit excited-state behaviour that is not as well documented in the smaller, heterocyclic canonical nucleobase analogues⁶⁸. For example, pyrene, one of the earliest non-canonical nucleobase analogues, exhibits efficient excimer fluorescence when two such residues are adjacent in a sequence¹³. In addition, non-canonical hydrophobic DNA bases can undergo base stacking that is much more favourable than that of canonical nucleobases (Fig. 5a)⁶⁹. Of course, this escape from the purine/pyrimidine framework can also present some limitations. Notably, bases that are too large to fit into the double-stranded helix context are unlikely to be efficient substrates for polymerase enzymes⁷⁰. Nevertheless, multiple laboratories have demonstrated strategies for the use of polymerases to incorporate varied designs of non-canonical fluorescent nucleobases into DNA^{71–74}.

For convenience sake, we divide this growing new class of fluorescent compounds into two categories based on their structures: (1) nucleobases made of polycyclic hydrocarbons (Fig. 5a); and (2) nucleobases comprising planar heterocyclic fluorophores (Fig. 5b). Polycyclic hydrocarbons were first incorporated into the DNA backbone in the 1990s, when fluorophores such as pyrene, phenanthrene, perylene and benzopyrene were introduced as new C-glycosidic nucleobases^{13,75}. Subsequent studies also reported the incorporation of pyrene into the RNA backbone⁷⁶. The α -anomers of these fluorescent nucleobases were synthesized using organometallic C–C coupling, which can be epimerized (if desired) to reveal the natural β -anomers. Owing to the strong π -interaction between neighbouring hydrocarbon nucleobases, exotic optical phenomena, such as intense excimers and energy transfer from such excited states, can be easily observed. This unusual emission behaviour has led to applications as biophysical probes of enzyme activity^{77,78}. Importantly, the energy transfer at the excited states leads to far-redshifted emission wavelengths. For example, pyrene can be excited in the long ultraviolet range (341 nm) and emits at 395 nm with a quantum yield of 0.12. Its excimer, however, emits around 500 nm with even higher quantum efficiency. Similarly, perylene and benzopyrene also display remarkable excimer or exciplex emission that redshift almost 100 nm compared with their original emission wavelengths. To date, several dozen hydrocarbon fluorescent nucleobase analogues have been developed (Fig. 5c)^{75,79,80}. Such compounds have been incorporated into libraries of thousands of short oligomers entirely composed of non-canonical nucleobase analogues, with the goal of discovery of unusual fluorescence properties such as multicolour emission with a single excitation^{81–83}.

Another example of unusual fluorescence properties afforded by non-canonical nucleobases is found in phenethynylpyrene deoxyriboside, which can be used in a fluorescence photoswitch. When two of the fluorescent nucleobases are adjacent, they yield green excimer fluorescence. However, on irradiation at 365 nm, the dimeric dye undergoes a rapid colour change as the alkynyl groups undergo a cycloaddition, blocking the excimer state (Fig. 5d). When combined with a redshifted dye acceptor, the dimeric fluorophore yields dramatic colour changes in stained human cells where they are exposed to light⁸⁴.

Heterocyclic non-canonical fluorophores are also an important category of fluorescent nucleobases. In 1972, arguably the first heterocyclic non-canonical fluorescent nucleobase,

ethenoA, a highly emissive adenine derivative, was introduced⁹. Although adenine is part of its structure, we categorize it as non-canonical because its added ring blocks its ability to undergo standard hydrogen-bonded base pairing. Another important and beautiful example has been the development of coumarin nucleobase analogues⁸⁵. Nile Red, porphyrin and many other fluorophores have also been added to this family (Fig. 5e)^{86,87}. A third example of heterocycle-containing non-canonical nucleobase is imidazophen-anthrene, used in detection of duplex DNA repair activity (Fig. 5f)⁸¹. Recently, the principle of polymerase shape complementarity⁸⁸ was also utilized to develop non-canonical fluorescent nucleobase pairs with high fidelity for enzymatic DNA synthesis. For instance, Dss-Px pair can be readily used for site-specific labelling of nucleic strands (Fig. 5g)⁸⁹. Other fluorescent nucleobase pairs developed include Ds-Px and Dss-Pa (not shown here), where Ds is 7-(2-thienyl)imidazo[4,5-b]-pyridine. The development of varied new heterocyclic fluorescent nucleobases has also triggered the need for developing alternative strategies for facile attachment of nucleobases to the ribose scaffold. Researchers have previously utilized the standard C–N bond in natural and most canonical fluorescent nucleobases, while the C–C bond has been used as well. However, in some cases heterocyclic fluorophores cannot be easily linked via C–N or C–C bonds. Hence chemists have utilized other linkers to the ribose sugar, such as C–O bonds in acetal chemistry, amide or ester linkages, and others^{90,91}. This has allowed a more diverse range of fluorophores to be directly stacked in the DNA.

As pointed out above, a potential limitation of employing non-canonical structure is that standard Watson–Crick pairing is lost, and in many cases, the ability to be incorporated into DNA or RNA via polymerase enzymes. However, researchers have been creative in overcoming this limitation. Large non-canonical bases can stack especially strongly, thus allowing one to form strong base pairs even without hydrogen bonding⁹². In addition, the use of polymerase enzymes that require no DNA template allows for the incorporation of surprisingly large fluorophores⁷³. Another successful strategy has been the fitting of non-canonical bases with size- and shape-complementary non-canonical partners so that even standard polymerases can function with the base pair^{92,93}. These strategies are described in more detail below.

Labelling methods

Site-specific incorporation of fluorescent nucleobases into DNA or RNA is a key step required for the large majority of applications. The labelling methods can be conceptually divided into three categories: (1) direct oligonucleotide synthesis using synthesizer and phosphoramidite chemistry; (2) post-synthesis or post-transcriptional modification using mild coupling conditions; and (3) enzymatic incorporation using primers that contain fluorescent nucleobases or fluorescent nucleotides in the pool (Table 1).

The most prevalent approach to chemically introduce fluorescent nucleobases into oligonucleotides is solid-phase synthesis using phosphoramidite chemistry coupled via nucleic acid synthesizer (Fig. 6a)⁹⁴. The clear advantage of this method is the synthetic power to readily and site-specifically introduce nucleotides of greatly varied structures and properties into the oligonucleotide strand. The disadvantage is also very prominent: the cost of completing one cycle is much higher than enzymatic methods, since scales are usually

larger. Furthermore, large, unnatural nucleobases may exhibit lower synthesis yields compared with natural nucleobases due to steric inhibition by the bulky structure. Also important is the limitation that many researchers who would like to apply fluorescent nucleobases do not have the expertise to synthesize such highly specialized nucleobase analogues. Moreover, even if they are available, they may not have access to specialized DNA synthesis on automated synthesizers. Finally, this approach has the limit of producing modified DNAs of roughly 100 nt or shorter, and ~50 nt or less for RNA.

Post-synthesis or post-transcriptional modification is a potentially powerful method for the site-specific labelling of DNAs and RNAs with fluorescent nucleobases, although this approach is currently in early stages of research⁹⁵. The chemical methodologies must be mild enough to limit unwanted modification or destruction of native DNA or RNA, while performing a bond connection in high yield in an aqueous reaction environment. To date, two reaction strategies in this direction are noteworthy: the Suzuki and Stille coupling methods (Fig. 6b). In these approaches, natural nucleobases are modified with halogens, for example 8-bromoguanine, 5-iodothymine and 2-iodoadenine. Fluorescent extensions are equipped with either organoboron or organostannane groups, which can be coupled to the oligonucleotides through palladium-catalysed conditions. For Suzuki coupling, 8-bromoguanine has been initially investigated to react with fluorescent extensions under the catalysis of palladium diacetate, heated to 70 °C in a water–acetonitrile mixture. This has been used to produce fluorescent phenyl-guanine or benzothiophene-guanine analogues⁹⁶. A more reactive 5-iodothymine has also been explored that could undergo the coupling reaction at lower temperatures⁹⁷. For Stille coupling, both 5-iodothymine and 2-iodoadenine have been explored to react with high excess of organostannane under the catalysis of palladium(0) triarylphosphine, heated to 60 °C in dimethylformamide^{98,99}. In these examples, the fluorescent extensions are directly conjugated to the nucleobases. The advantages of these chemistry-based post-synthesis modifications are: (1) site-specific labelling of nucleic acid strands with lengths of up to ~1,000 nt and potentially at any given position; (2) less expensive incorporation of fluorescent nucleobases than direct chemical synthesis. The disadvantages of this approach mainly derive from its early development stage, with the lack of diverse coupling methods being its biggest bottleneck.

The third labelling method involves incorporation of fluorescent nucleotides into enzymatically produced DNAs and RNAs, with the nucleobases residing either in a primer or in the nucleotide pool. In this regard, canonical fluorescent nucleobases can also be labelled either near the 5' end of the strand or in the middle of the strand where its pairing partner resides. For example, a benzo-fused 7-deazaadenosine was incorporated by KOD XL DNA polymerase as the base pairing partner of T (ref. 100). In another case, thieno[3, 4-*d*]U was transcribed into single strand RNA by T7 RNA polymerase¹⁰¹. In addition, a very recent report demonstrated efficient and selective incorporation of size-expanded nucleobases; that is, xA, xG, xT and xC, into DNA with human DNA polymerase θ (ref. 102). The large-sized nucleobases exhibit base-pairing and enhanced base stacking⁶⁹.

With careful design, non-canonical nucleobase analogues can also be incorporated into DNAs and RNAs site-specifically by polymerases. Perhaps the earliest example for this was the pyrene nucleobase paired opposite abasic sites in DNA. In this 'base pair' design, the

pyrene is as large as a full pair, and thus the absence of a pairing partner (that is, an abasic site) provides space for the 'pair' (which is as stable as an A–T base pair) to form. In 1999, it was shown that DNA polymerases could incorporate the fluorescent pyrene nucleotide (dPTP, deoxynucleoside 5'-triphosphate derivative of pyrene (P) deoxynucleoside) efficiently and selectively opposite abasic sites in DNA^{92,93}. A recent prominent and important example is the Dss–Px unnatural nucleobase pair, with Dss being a fluorescent nucleobase and Px a quencher (Fig. 6c)¹⁰³. In this case, Dss is labelled in the primer and it pairs specifically with Px. Through polymerase chain reaction with an engineered polymerase, the pairs can be introduced site-specifically into longer DNAs. In one study, Px was site-specifically paired with Dss, quenching its fluorescence emission. The Ds–Pa pair was also used with T7 RNA polymerase to site-specifically label RNA transcripts¹⁰⁴. Such a system can in principle be expanded to other fluorescent nucleobases, as long as a specific base pair can be exploited. The advantage of this system is the ability to site-specifically locate even a single fluorophore while otherwise maintaining native structure and sequence. Such an approach enables inexpensive labelling, and suggests potential biological compatibility.

Finally, it is also worth noting that template-independent DNA polymerases (such as terminal deoxynucleotidyl transferase) can incorporate multiple fluorophores at the end of a single-stranded DNA or RNA primer⁷⁴. This allows both canonical fluorescent nucleobases, and non-canonical nucleobases to be strung together as polymeric fluorescent tags or strings at the 3' end of DNA or RNA. In one study, researchers took advantage of the excimer fluorescence of pyrene nucleobase analogue to yield a colour-change homogeneous assay of DNA fragment ends, a common result of apoptosis⁷³.

Emerging applications

Fluorescent nucleobases have been employed in a great diversity of chemical, structural, biophysical and biochemical implementations, including single nucleotide polymorphism detection^{16,17,20,105–108}, microenvironment monitoring^{28,109,110}, structural and morphological measurements^{111–115}, as well as enzyme (polymerase) activity testing^{24,116–118}. The choice of which fluorescent nucleobase is most appropriate for the given application depends on the context of the experiment. For instance, in the design of reporters of enzymatic activity, fluorescent nucleobases might either interact with an enzyme substrate or serve as the substrate themselves; thus on enzymatic reactions, their fluorescence signal changes can be detected. This approach requires careful design or choice of the fluorescent nucleobase. In one recent example, a fluorogenic probe utilizing a pyrene nucleobase and the quenching properties of positively charged 1-methyladenine was used to probe DNA repair in real time. The excited energy level of the fluorophore lies above the highest unoccupied electronic level of 1-methyladenine, allowing quenching by photoinduced charge transfer. This probe provides a simple way to measure the intracellular activity of ALKBH3, an important demethylation enzyme involved in tumorigenesis. This probe can further be used to screen and identify enzyme inhibitors (Fig. 7a)¹¹⁹. In a second example, a fluorescent adenosine analogue thieno[3,4-*d*]-6-aminopyrimidine (thA) was designed to analyse the activity of adenosine-to-inosine RNA editing adenosine deaminase acting on RNA (ADAR) enzymes. The fluorescent nucleobase is well designed to fit into the

enzyme pocket and retain the activities; thus after reaction, the 4-position amine transforms to a ketone group, which results in redshifted fluorescence emission. thA is able not only to monitor the ADAR2 deamination process, but also to probe the enzyme editing site (Fig. 7b)¹²⁰.

In addition to *in vitro* studies, researchers have begun to extend the application of fluorescent nucleobases into reporting on the biology of living cells. In such circumstances, one must carefully consider the strategy for delivery of the nucleoside or nucleotide into the cell, and its subsequent metabolism. Negatively charged nucleotides are typically not taken up, and so neutral nucleosides may instead be used, as long as they can be substrates for subsequent cellular kinase enzymes. For example, a recently developed fluorescent thymidine analogue ^{DMA}T was directed for the measurement of mercury. The kinetic stability of the T–Hg^{II}–T metallo–base pair and engendered fluorescence change allows the investigation of mercury induced DNA metabolism *in vivo* (Fig. 7c)⁶⁷. The modification takes place at the 5 and 6 positions of thymine, thus not strongly affecting base pairing. A simple dimethylamino group is mounted on the *para*-position of the appended benzene ring, forming an electron donor–acceptor pair that significantly redshifts the emission maximum beyond 500 nm. The simple design incorporates both compatibility with base pairing and redshifted emission for analysis. In another case, a furan-fused uracil analogue was developed to conduct functional detection of human concentrative nucleoside transporters, which determine the flux of nucleosides through the plasma membrane. Although the specificity of the nucleoside transporter is not determined, this simple but efficient design not only allows the visualization of transporters in living cells, but also provides a robust way to predict nucleoside-derived drug sensitivity (Fig. 7d)¹²¹. This work could inspire a broad range of fluorescence analysis related to nucleic acid–protein interactions.

Concluding remarks and future prospects

The field of fluorescent nucleobases has undergone rapid growth, both in synthesis and in applications. Future opportunities for development in the design and synthesis will no doubt include expanding the optical properties of nucleoside analogues. For example, there remains a need for bright fluorophores that are shifted beyond the blue and green parts of the spectrum¹²². A major design challenge is to enhance conjugation for redshifting, while keeping the conjugated structures manageably small so that the perturbation of the DNA structure is not a severe problem. Alternatively, extended conjugation will take place at positions that do not adversely affect the biological structures or activities under investigation. Another significant target in this field is the development of fluorescent nucleobase analogues that are bright enough, and sufficiently resistant to photobleaching, to be broadly useful in single-molecule analyses and super-resolution imaging¹²³. Beyond these, a third area for expansion is the development of new environmentally sensitive nucleobase analogues that report with high sensitivity on structure and interactions, but which retain high quantum yields.

A second design challenge for the fluorescent nucleobase field is the development of improved strategies and methods for incorporating such analogues into larger DNAs and RNAs. Although purely synthetic methods now make it relatively simple to produce short

labelled DNAs and RNAs (up to roughly 50–100 nt), the development of additional methods for labelling larger RNAs and DNAs made by polymerases (for example, 100–1,000 nt sizes) would be welcome, especially by biologists who do not have access to synthetic chemistry methods. This may well require the development of new enzyme-friendly fluorescent nucleotide and base-pair analogues, and also may benefit from the development of new polymerase enzymes that can better accept modified nucleotides as substrates. Further in the future, the development of fluorescent nucleotide analogues that can be delivered into cells and be incorporated into cellular RNAs and DNAs would be exciting and highly useful in biological and biomedical research. A number of challenges must be addressed to make this a reality, including not only cell permeability, but also metabolic stability. Modified nucleosides must escape the cellular machinery that recognizes damage both in the nucleotide pool and in the DNA once incorporated there. It also seems likely that, in addition to the chemical post-synthesis modification, gene-editing techniques and ligation strategies will play key roles in the fluorescent labelling of biologically derived nucleic acids (Fig. 6b). Enzymatic assembly and incorporation of relatively large fragments of DNA is becoming increasingly common, and inclusion of fluorescence-labelled segments could incorporate designed fluorescent cassettes into large DNAs.

The field of fluorescent nucleobase development is driven not only by synthesis but also the fluorescence properties that can result. The utility of a given analogue depends not only on its properties, but also on how those properties vary in different structural contexts: single stranded versus double stranded, and with varied neighbouring bases. A better theoretical understanding of the effects of environment on fluorescence of such compounds will in the future lead to predictive designs and will greatly enhance future applications. Similarly, given that there are now over 100 fluorescent nucleobases to choose from, we are beginning to see an increase in the combination of multiple fluorescent analogues into DNA and RNA strands, resulting in interesting and potentially useful new fluorescence properties that emerge from the electronic interactions between the fluorophores^{83,124}. A better theoretical and physical understanding of these electronic interactions will ultimately lead to more effective designs and a broad range of applications in imaging, sensing and biological investigation.

Finally, possibly the biggest opportunity for future expansion of the fluorescent nucleobase field may be their biological and biomedical applications^{125–128}. There are numerous possible uses of such compounds in biological imaging, and this has only just begun to be explored. Similarly, one can readily envision the use of such analogues in numerous biological sensing and reporting applications. Although *in vitro* studies are now rapidly expanding in this direction, the use of such compounds in living cells and organisms is in its infancy.

Acknowledgments

We thank the US National Institutes of Health (GM067201, GM110050, CA217809) for support.

References

1. Wilhelmsson, M., Tor, Y. *Fluorescent Analogues of Biomolecular Building Blocks: Design and Applications*. John Wiley & Sons; 2016.
2. Ward DC, Reich E, Stryer L. Fluorescence studies of nucleotides and polynucleotides. I Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives. *J Biol Chem*. 1969; 244:1228–1237. [PubMed: 5767305]
3. Watson, JD. *Molecular Biology of the Gene*. 7. Pearson; 2013.
4. Hocek M, Fojta M. Nucleobase modification as redox DNA labelling for electrochemical detection. *Chem Soc Rev*. 2011; 40:5802–5814. [PubMed: 21625726]
5. Prober JM, et al. A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science*. 1987; 238:336–341. [PubMed: 2443975]
6. Fuller CW, et al. The challenges of sequencing by synthesis. *Nat Biotechnol*. 2009; 27:1013–1023. [PubMed: 19898456]
7. Lacenere C, Garg MK, Stoltz BM, Quake SR. Effects of a modified dye-labeled nucleotide spacer arm on incorporation by thermophilic DNA polymerases. *Nucleosides Nucleotides Nucleic Acids*. 2006; 25:9–15. [PubMed: 16440981]
8. Lobsiger S, Blaser S, Sinha RK, Frey HM, Leutwyler S. Switching on the fluorescence of 2-aminopurine by site-selective microhydration. *Nat Chem*. 2014; 6:989–993. [PubMed: 25343604]
9. Secrist JA III, Barrio JR, Leonard NJ. A fluorescent modification of adenosine triphosphate with activity in enzyme systems: 1,*N*⁶-ethenoadenosine triphosphate. *Science*. 1972; 175:646–647. [PubMed: 4257930]
10. Rist MJ, Marino JP. Fluorescent nucleotide base analogs as probes of nucleic acid structure, dynamics and interactions. *Curr Org Chem*. 2002; 6:775–793.
11. Wu P, Nordlund TM, Gildea B, McLaughlin LW. Base stacking and unstacking as determined from a DNA decamer containing a fluorescent base. *Biochemistry*. 1990; 29:6508–6514. [PubMed: 2207092]
12. Hawkins ME, Pfliegerer W, Mazumder A, Pommier YG, Balis FM. Incorporation of a fluorescent guanosine analog into oligonucleotides and its application to a real time assay for the HIV-1 integrase 3'-processing reaction. *Nucleic Acids Res*. 1995; 23:2872–2880. [PubMed: 7659509]
13. Ren RX, Chaudhuri NC, Paris PL, Rumney S, Kool ET. Naphthalene, phenanthrene, and pyrene as DNA base analogues: synthesis, structure, and fluorescence in DNA. *J Am Chem Soc*. 1996; 118:7671–7678. [PubMed: 20865136]
14. Godde F, Toulme JJ, Moreau S. Benzoquinazoline derivatives as substitutes for thymine in nucleic acid complexes. Use of fluorescence emission of benzo[g]quinazoline-2,4-(1*H*,3*H*)-dione in probing duplex and triplex formation. *Biochemistry*. 1998; 37:13765–13775. [PubMed: 9753465]
15. Okamoto A, Kanatani K, Saito I. Pyrene-labeled base-discriminating fluorescent DNA probes for homogeneous SNP typing. *J Am Chem Soc*. 2004; 126:4820–4827. [PubMed: 15080686]
16. Okamoto A, Tainaka K, Saito I. Clear distinction of purine bases on the complementary strand by a fluorescence change of a novel fluorescent nucleoside. *J Am Chem Soc*. 2003; 125:4972–4973. [PubMed: 12708835]
17. Okamoto A, Tanaka K, Fukuta T, Saito I. Design of base-discriminating fluorescent nucleoside and its application to T/C SNP typing. *J Am Chem Soc*. 2003; 125:9296–9297. [PubMed: 12889950]
18. Wilhelmsson LM, et al. Photophysical characterization of fluorescent DNA base analogue, tC. *J Phys Chem B*. 2003; 107:9094–9101.
19. Vrabel M, et al. Base-modified DNA labeled by [Ru(bpy)₃]²⁺ and [Os(bpy)₃]²⁺ complexes: construction by polymerase incorporation of modified nucleoside triphosphates, electrochemical and luminescent properties, and applications. *Chem Eur J*. 2009; 15:1144–1154. [PubMed: 19072947]
20. Miyata K, et al. Synthesis and properties of a new fluorescent bicyclic 4-*N*-carbamoyldeoxycytidine derivative. *Org Lett*. 2006; 8:1545–1548. [PubMed: 16597106]

21. Hurley DJ, Seaman SE, Mazura JC, Tor Y. Fluorescent 1,10-phenanthroline-containing oligonucleotides distinguish between perfect and mismatched base pairing. *Org Lett*. 2002; 4:2305–2308. [PubMed: 12098233]
22. Noe MS, Sinkeldam RW, Tor Y. Oligodeoxynucleotides containing multiple thiophene-modified isomorphous fluorescent nucleosides. *J Org Chem*. 2013; 78:8123–8128. [PubMed: 23859712]
23. Shin D, Sinkeldam RW, Tor Y. Emissive RNA alphabet. *J Am Chem Soc*. 2011; 133:14912–14915. [PubMed: 21866967]
24. Srivatsan SG, Greco NJ, Tor Y. A highly emissive fluorescent nucleoside that signals the activity of toxic ribosome-inactivating proteins. *Angew Chem Int Ed*. 2008; 47:6661–6665.
25. Wojciechowski F, Hudson RH. Fluorescence and hybridization properties of peptide nucleic acid containing a substituted phenylpyrrolocytosine designed to engage guanine with an additional H-bond. *J Am Chem Soc*. 2008; 130:12574–12575. [PubMed: 18761442]
26. Sinkeldam RW, Greco NJ, Tor Y. Fluorescent analogs of biomolecular building blocks: design, properties, and applications. *Chem Rev*. 2010; 110:2579–2619. [PubMed: 20205430]
27. Hawkins ME, Pfeleiderer W, Balis FM, Porter D, Knutson JR. Fluorescence properties of pteridine nucleoside analogs as monomers and incorporated into oligonucleotides. *Anal Biochem*. 1997; 244:86–95. [PubMed: 9025913]
28. Hawkins ME. Fluorescent pteridine nucleoside analogs: a window on DNA interactions. *Cell Biochem Biophys*. 2001; 34:257–281. [PubMed: 11898867]
29. Hawkins ME, Pfeleiderer W, Jungmann O, Balis FM. Synthesis and fluorescence characterization of pteridine adenosine nucleoside analogs for DNA incorporation. *Anal Biochem*. 2001; 298:231–240. [PubMed: 11700977]
30. Atkins, PW., De Paula, J. *Physical Chemistry: Thermodynamics, Structure, and Change*. 10. Freeman and Company; 2014.
31. Wojtuszewski Poulin K, Smirnov AV, Hawkins ME, Balis FM, Knutson JR. Conformational heterogeneity and quasi-static self-quenching in DNA containing a fluorescent guanine analogue, 3MI or 6MI. *Biochemistry*. 2009; 48:8861–8868. [PubMed: 19610668]
32. Seo YJ, Ryu JH, Kim BH. Quencher-free, end-stacking oligonucleotides for probing single-base mismatches in DNA. *Org Lett*. 2005; 7:4931–4933. [PubMed: 16235925]
33. Wagner C, Rist M, Mayer-Enthart E, Wagenknecht HA. 1-Ethynylpyrene-modified guanine and cytosine as optical labels for DNA hybridization. *Org Biomol Chem*. 2005; 3:2062–2063. [PubMed: 15917887]
34. Forster U, et al. 2-(1-Ethynylpyrene)-adenosine as a folding probe for RNA — pyrene in or out. *ChemBioChem*. 2010; 11:664–672. [PubMed: 20183842]
35. Vrabel M, Pohl R, Klepetarova B, Votruba I, Hocek M. Synthesis of 2'-deoxyadenosine nucleosides bearing bipyridine-type ligands and their Ru-complexes in position 8 through cross-coupling reactions. *Org Biomol Chem*. 2007; 5:2849–2857. [PubMed: 17700854]
36. Vrabel M, et al. Synthesis and photophysical properties of 7-deaza-2'-deoxyadenosines bearing bipyridine ligands and their Ru(II)-complexes in position 7. *Org Biomol Chem*. 2008; 6:2852–2860. [PubMed: 18688477]
37. Saito Y, Koda M, Shinohara Y, Saito I. Synthesis and photophysical properties of 8-arylbutadienyl 2'-deoxyguanosines. *Tetrahedron Lett*. 2011; 52:491–494.
38. Okamoto A, Saito Y, Saito I. Design of base-discriminating fluorescent nucleosides. *J Photochem Photobiol C*. 2005; 6:108–122.
39. de Silva AP, et al. Signaling recognition events with fluorescent sensors and switches. *Chem Rev*. 1997; 97:1515–1566. [PubMed: 11851458]
40. Wierzchowski J, Wielgus-Kutrowska B, Shugar D. Fluorescence emission properties of 8-azapurines and their nucleosides, and application to the kinetics of the reverse synthetic reaction of purine nucleoside phosphorylase. *Biochim Biophys Acta*. 1996; 1290:9–17. [PubMed: 8645713]
41. Tor Y, et al. Designing new isomorphous fluorescent nucleobase analogues: the thieno[3,2-d]pyrimidine core. *Tetrahedron*. 2007; 63:3608–3614.
42. Da Costa CP, Fedor MJ, Scott LG. 8-Azaguanine reporter of purine ionization states in structured RNAs. *J Am Chem Soc*. 2007; 129:3426–3432. [PubMed: 17326637]

43. Zhu Z, et al. Electronic band structure and sub-band-gap absorption of nitrogen hyperdoped silicon. *Sci Rep*. 2015; 5:10513. [PubMed: 26012369]
44. Ben Gaied N, et al. 8-Vinyl-deoxyadenosine, an alternative fluorescent nucleoside analog to 2'-deoxyribosyl-2-aminopurine with improved properties. *Nucleic Acids Res*. 2005; 33:1031–1039. [PubMed: 15718302]
45. Nadler A, Strohmeier J, Diederichsen U. 8-Vinyl-2'-deoxyguanosine as a fluorescent 2'-deoxyguanosine mimic for investigating DNA hybridization and topology. *Angew Chem Int Ed*. 2011; 50:5392–5396.
46. Mitsui T, Kimoto M, Kawai R, Yokoyama S, Hirao I. Characterization of fluorescent, unnatural base pairs. *Tetrahedron*. 2007; 63:3528–3537.
47. Greco NJ, Tor Y. Furan decorated nucleoside analogues as fluorescent probes: synthesis, photophysical evaluation and site-specific incorporation. *Tetrahedron*. 2007; 63:3515–3527. [PubMed: 18431439]
48. Sun KM, McLaughlin CK, Lantero DR, Manderville RA. Biomarkers for phenol carcinogen exposure act as pH-sensing fluorescent probes. *J Am Chem Soc*. 2007; 129:1894–1895. [PubMed: 17256942]
49. Schlitt KM, Millen AL, Wetmore SD, Manderville RA. An indolelinked C8-deoxyguanosine nucleoside acts as a fluorescent reporter of Watson–Crick versus Hoogsteen base pairing. *Org Biomol Chem*. 2011; 9:1565–1571. [PubMed: 21240404]
50. Fazakerley GV, Russell JC, Wolfe MA. Determination of the syn-anti equilibrium of some purine 3':5'-nucleotides by nuclear-magnetic-relaxation perturbation in the presence of a lanthanide-ion probe. *Eur J Biochem*. 1977; 76:601–605. [PubMed: 196852]
51. Leonard NJ, Sprecker MA, Morrice AG. Defined dimensional changes in enzyme substrates and cofactors. Synthesis of lin-benzoadenosine and enzymatic evaluation of derivatives of the benzopurines. *J Am Chem Soc*. 1976; 98:3987–3994. [PubMed: 819479]
52. Okamoto A, Tanaka K, Fukuta T, Saito I. Cytosine detection by a fluorescein-labeled probe containing base-discriminating fluorescent nucleobase. *ChemBioChem*. 2004; 5:958–963. [PubMed: 15239053]
53. Krueger AT, Kool ET. Fluorescence of size-expanded DNA bases: reporting on DNA sequence and structure with an unnatural genetic set. *J Am Chem Soc*. 2008; 130:3989–3999. [PubMed: 18311973]
54. Hernandez AR, Kool ET. The components of xRNA: synthesis and fluorescence of a full genetic set of size-expanded ribonucleosides. *Org Lett*. 2011; 13:676–679. [PubMed: 21214222]
55. Pesnot T, Wagner GK. Novel derivatives of UDP-glucose: concise synthesis and fluorescent properties. *Org Biomol Chem*. 2008; 6:2884–2891. [PubMed: 18688480]
56. Greco NJ, Tor Y. Simple fluorescent pyrimidine analogues detect the presence of DNA abasic sites. *J Am Chem Soc*. 2005; 127:10784–10785. [PubMed: 16076156]
57. Miyata K, et al. Synthesis and fluorescent properties of bi- and tricyclic 4-*N*-carbamoyldeoxycytidine derivatives. *J Org Chem*. 2007; 72:102–108. [PubMed: 17194087]
58. Flanagan WM, et al. A cytosine analog that confers enhanced potency to antisense oligonucleotides. *Proc Natl Acad Sci USA*. 1999; 96:3513–3518. [PubMed: 10097067]
59. Barhate N, Cekan P, Massey AP, Sigurdsson ST. A nucleoside that contains a rigid nitroxide spin label: a fluorophore in disguise. *Angew Chem Int Ed*. 2007; 46:2655–2658.
60. Mizuta M, Seio K, Miyata K, Sekine M. Fluorescent pyrimidopyrimidoindole nucleosides: control of photophysical characterizations by substituent effects. *J Org Chem*. 2007; 72:5046–5055. [PubMed: 17555352]
61. Korshun VA, et al. New pyrene derivatives for fluorescent labeling of oligonucleotides. *Nucleosides Nucleotides*. 1997; 16:1461–1464.
62. Ehrenschwender T, Wanninger-Weiss C, Wagenknecht HA. BODIPY-modified uridines as potential fluorescent probes for nucleic acids that are recognized by DNA-polymerases. *Nucleic Acids Symp Ser*. 2008; 52:349–350.
63. Dodd DW, et al. Blue fluorescent deoxycytidine analogues: convergent synthesis, solid-state and electronic structure, and solvatochromism. *Org Biomol Chem*. 2010; 8:663–666. [PubMed: 20090985]

64. Jarchow-Choy SK, Krueger AT, Liu H, Gao J, Kool ET. Fluorescent xDNA nucleotides as efficient substrates for a template-independent polymerase. *Nucleic Acids Res.* 2011; 39:1586–1594. [PubMed: 20947563]
65. Hardman SJ, Botchway SW, Thompson KC. Evidence for a nonbase stacking effect for the environment-sensitive fluorescent base pyrrolocytosine — comparison with 2-aminopurine. *Photochem Photobiol.* 2008; 84:1473–1479. [PubMed: 18513237]
66. Noe MS, Rios AC, Tor Y. Design, synthesis, and spectroscopic properties of extended and fused pyrrolo-dC and pyrrolo-C analogs. *Org Lett.* 2012; 14:3150–3153. [PubMed: 22646728]
67. Schmidt OP, Mata G, Luedtke NW. Fluorescent base analog reveals T-Hg^{II}-T base pairs have high kinetic stabilities that perturb DNA metabolism. *J Am Chem Soc.* 2016; 138:14733–14739. [PubMed: 27709934]
68. Paris PL, Langenhan JM, Kool ET. Probing DNA sequences in solution with a monomer-excimer fluorescence color change. *Nucleic Acids Res.* 1998; 26:3789–3793. [PubMed: 9685497]
69. Guckian KM, et al. Factors contributing to aromatic stacking in water: evaluation in the context of DNA. *J Am Chem Soc.* 2000; 122:2213–2222. [PubMed: 20865137]
70. Kool ET. Active site tightness and substrate fit in DNA replication. *Annu Rev Biochem.* 2002; 71:191–219. [PubMed: 12045095]
71. Guckian KM, et al. Experimental measurement of aromatic stacking affinities in the context of duplex DNA. *J Am Chem Soc.* 1996; 118:8182–8183. [PubMed: 20882117]
72. Morales JC, Kool ET. Functional hydrogen-bonding map of the minor groove binding tracks of six DNA polymerases. *Biochemistry.* 2000; 39:12979–12988. [PubMed: 11041863]
73. Cho Y, Kool ET. Enzymatic synthesis of fluorescent oligomers assembled on a DNA backbone. *ChemBioChem.* 2006; 7:669–672. [PubMed: 16502476]
74. Hollenstein M, Wojciechowski F, Leumann CJ. Polymerase incorporation of pyrene-nucleoside triphosphates. *Bioorg Med Chem Lett.* 2012; 22:4428–4430. [PubMed: 22633686]
75. Kool ET. Replacing the nucleobases in DNA with designer molecules. *Acc Chem Res.* 2002; 35:936–943. [PubMed: 12437318]
76. Spadafora M, Mehiri M, Burger A, Benhida R. Friedel–Crafts and modified Vorbrüggen ribosylation. A short synthesis of aryl and heteroaryl-C-nucleosides. *Tetrahedron Lett.* 2008; 49:3967–3971.
77. Sun L, Wang M, Kool ET, Taylor JS. Pyrene Nucleotide as a mechanistic probe: evidence for a transient abasic site-like intermediate in the bypass of dipyrimidine photoproducts by T7 DNA polymerase. *Biochemistry.* 2000; 39:14603–14610. [PubMed: 11087416]
78. Dai N, Kool ET. Fluorescent DNA-based enzyme sensors. *Chem Soc Rev.* 2011; 40:5756–5770. [PubMed: 21290032]
79. Grigorenko NA, Leumann CJ. 2-Phenanthrenyl-DNA: synthesis, pairing, and fluorescence properties. *Chem Eur J.* 2009; 15:639–645. [PubMed: 19053088]
80. Wojciechowski F, Lietard J, Leumann CJ. 2-Pyrenyl-DNA: synthesis, pairing, and fluorescence properties. *Org Lett.* 2012; 14:5176–5179. [PubMed: 23025305]
81. Gao J, Strassler C, Tahmassebi D, Kool ET. Libraries of composite polyfluors built from fluorescent deoxyribosides. *J Am Chem Soc.* 2002; 124:11590–11591. [PubMed: 12296712]
82. Gao J, Watanabe S, Kool ET. Modified DNA analogues that sense light exposure with color changes. *J Am Chem Soc.* 2004; 126:12748–12749. [PubMed: 15469249]
83. Teo YN, Kool ET. DNA-multichromophore systems. *Chem Rev.* 2012; 112:4221–4245. [PubMed: 22424059]
84. Chan KM, Kolmel DK, Wang S, Kool ET. Color-change photoswitching of an alkynylpyrene excimer dye. *Angew Chem Int Ed.* 2017; 56:6497–6501.
85. Coleman RS, Madaras ML. Synthesis of a novel coumarin C-riboside as a photophysical probe of oligonucleotide dynamics. *J Org Chem.* 1998; 63:5700–5703.
86. Morales-Rojas H, Kool ET. A porphyrin C-nucleoside incorporated into DNA. *Org Lett.* 2002; 4:4377–4380. [PubMed: 12465891]

87. Okamoto A, Tainaka K, Fujiwara Y. Nile Red nucleoside: design of a solvatofluorochromic nucleoside as an indicator of micropolarity around DNA. *J Org Chem.* 2006; 71:3592–3598. [PubMed: 16626146]
88. Dzantiev L, Alekseyev YO, Morales JC, Kool ET, Romano LJ. Significance of nucleobase shape complementarity and hydrogen bonding in the formation and stability of the closed polymerase-DNA complex. *Biochemistry.* 2001; 40:3215–3221. [PubMed: 11258938]
89. Hirao I, Kimoto M, Yamashige R. Natural versus artificial creation of base pairs in DNA: origin of nucleobases from the perspectives of unnatural base pair studies. *Acc Chem Res.* 2012; 45:2055–2065. [PubMed: 22263525]
90. Singh V, Wang S, Kool ET. Genetically encoded multispectral labeling of proteins with polyfluorophores on a DNA backbone. *J Am Chem Soc.* 2013; 135:6184–6191. [PubMed: 23590213]
91. Teo YN, Wilson JN, Kool ET. Polyfluorophores on a DNA backbone: a multicolor set of labels excited at one wavelength. *J Am Chem Soc.* 2009; 131:3923–3933. [PubMed: 19254023]
92. Matray TJ, Kool ET. Selective and stable DNA base pairing without hydrogen bonds. *J Am Chem Soc.* 1998; 120:6191–6192. [PubMed: 20852721]
93. Matray TJ, Kool ET. A specific partner for abasic damage in DNA. *Nature.* 1999; 399:704–708. [PubMed: 10385125]
94. Kosuri S, Church GM. Large-scale *de novo* DNA synthesis: technologies and applications. *Nat Methods.* 2014; 11:499–507. [PubMed: 24781323]
95. Verma S, Eckstein F. Modified oligonucleotides: synthesis and strategy for users. *Annu Rev Biochem.* 1998; 67:99–134. [PubMed: 9759484]
96. Omumi A, Beach DG, Baker M, Gabryelski W, Manderville RA. Postsynthetic guanine arylation of DNA by Suzuki-Miyaura cross-coupling. *J Am Chem Soc.* 2011; 133:42–50. [PubMed: 21067186]
97. Lercher L, McGouran JF, Kessler BM, Schofield CJ, Davis BG. DNA modification under mild conditions by Suzuki-Miyaura cross-coupling for the generation of functional probes. *Angew Chem Int Ed.* 2013; 52:10553–10558.
98. Wicke L, Engels JW. Postsynthetic on column RNA labeling via Stille coupling. *Bioconjug Chem.* 2012; 23:627–642. [PubMed: 22321023]
99. Krause A, Hertl A, Muttach F, Jaschke A. Phosphine-free Stille-Migita chemistry for the mild and orthogonal modification of DNA and RNA. *Chem Eur J.* 2014; 20:16613–16619. [PubMed: 25322724]
100. Bosakova A, Perlikova P, Tichy M, Pohl R, Hocek M. 6-Aryl-4-aminopyrimido[4,5-*b*]indole 2'-deoxyribonucleoside triphosphates (benzo-fused 7-deaza-dATP analogues): synthesis, fluorescent properties, enzymatic incorporation into DNA and DNA-protein binding study. *Bioorg Med Chem.* 2016; 24:4528–4535. [PubMed: 27498304]
101. Srivatsan SG, Tor Y. Enzymatic incorporation of emissive pyrimidine ribonucleotides. *Chem Asian J.* 2009; 4:419–427. [PubMed: 19072942]
102. Kent T, et al. DNA polymerase theta specializes in incorporating synthetic expanded-size (xDNA) nucleotides. *Nucleic Acids Res.* 2016; 44:9381–9392. [PubMed: 27591252]
103. Kimoto M, et al. A new unnatural base pair system between fluorophore and quencher base analogues for nucleic acid-based imaging technology. *J Am Chem Soc.* 2010; 132:15418–15426. [PubMed: 20939572]
104. Someya T, Ando A, Kimoto M, Hirao I. Site-specific labeling of RNA by combining genetic alphabet expansion transcription and copper-free click chemistry. *Nucleic Acids Res.* 2015; 43:6665–6676. [PubMed: 26130718]
105. Cekan P, Sigurdsson ST. Single base interrogation by a fluorescent nucleotide: each of the four DNA bases identified by fluorescence spectroscopy. *Chem Commun.* 2008:3393–3395.
106. Mizuta M, et al. A pyrimidopyrimidoindole nucleoside (dC PPI): photophysical properties and thermal stability of the modified DNA duplexes. *Nucleosides Nucleotides Nucleic Acids.* 2007; 26:1335–1338. [PubMed: 18066779]
107. Okamoto A, Tainaka K, Saito I. Synthesis and properties of a novel fluorescent nucleobase, naphthopyridopyrimidine. *Tetrahedron Lett.* 2003; 44:6871–6874.

108. Saito Y, Miyauchi Y, Okamoto A, Saito I. Synthesis and properties of novel base-discriminating fluorescent (BDF) nucleosides: a highly polarity-sensitive fluorophore for SNP typing. *Tetrahedron Lett.* 2004; 45:7827–7831.
109. Brauns EB, Madaras ML, Coleman RS, Murphy CJ, Berg MA. Measurement of local DNA reorganization on the picosecond and nanosecond time scales. *J Am Chem Soc.* 1999; 121:11644–11649.
110. Jeong HS, Kang S, Lee JY, Kim BH. Probing specific RNA bulge conformations by modified fluorescent nucleosides. *Org Biomol Chem.* 2009; 7:921–925. [PubMed: 19225675]
111. Borjesson K, et al. Nucleic acid base analog FRET-pair facilitating detailed structural measurements in nucleic acid containing systems. *J Am Chem Soc.* 2009; 131:4288–4293. [PubMed: 19317504]
112. Godde F, Toulme JJ, Moreau S. 4-amino-1H-benzo[g]quinazoline-2-one: a fluorescent analog of cytosine to probe protonation sites in triplex forming oligonucleotides. *Nucleic Acids Res.* 2000; 28:2977–2985. [PubMed: 10908362]
113. Hwang GT, Seo YJ, Kim BH. Pyrene-labeled deoxyuridine and deoxyadenosine: fluorescent discriminating phenomena in their oligonucleotides. *Tetrahedron Lett.* 2005; 46:1475–1477.
114. Marti AA, Jockusch S, Li Z, Ju J, Turro NJ. Molecular beacons with intrinsically fluorescent nucleotides. *Nucleic Acids Res.* 2006; 34:e50. [PubMed: 16595796]
115. Wojtuszewski K, Hawkins ME, Cole JL, Mukerji I. HU binding to DNA: evidence for multiple complex formation and DNA bending. *Biochemistry.* 2001; 40:2588–2598. [PubMed: 11327882]
116. Kirk SR, Luedtke NW, Tor Y. 2-Aminopurine as a real-time probe of enzymatic cleavage and inhibition of hammerhead ribozymes. *Bioorg Med Chem.* 2001; 9:2295–2301. [PubMed: 11553468]
117. Raney KD, Sowers LC, Millar DP, Benkovic SJ. A fluorescence-based assay for monitoring helicase activity. *Proc Natl Acad Sci USA.* 1994; 91:6644–6648. [PubMed: 8022830]
118. Sandin P, et al. Highly efficient incorporation of the fluorescent nucleotide analogs tC and tCO by Klenow fragment. *Nucleic Acids Res.* 2009; 37:3924–3933. [PubMed: 19401439]
119. Beharry AA, Lacoste S, O'Connor TR, Kool ET. Fluorescence monitoring of the oxidative repair of DNA alkylation damage by ALKBH3, a prostate cancer marker. *J Am Chem Soc.* 2016; 138:3647–3650. [PubMed: 26967262]
120. Mizrahi RA, et al. A fluorescent adenosine analogue as a substrate for an A-to-I RNA editing enzyme. *Angew Chem Int Ed.* 2015; 54:8713–8716.
121. Claudio-Montero A, et al. Fluorescent nucleoside derivatives as a tool for the detection of concentrative nucleoside transporter activity using confocal microscopy and flow cytometry. *Mol Pharm.* 2015; 12:2158–2166. [PubMed: 25923048]
122. Lukinavicius G, et al. A near-infrared fluorophore for live-cell superresolution microscopy of cellular proteins. *Nat Chem.* 2013; 5:132–139. [PubMed: 23344448]
123. Dempsey GT, Vaughan JC, Chen KH, Bates M, Zhuang X. Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nat Methods.* 2011; 8:1027–1036. [PubMed: 22056676]
124. Yuen LH, et al. Pattern-based detection of toxic metals in surface water with DNA polyfluorophores. *Angew Chem Int Ed.* 2014; 53:5361–5365.
125. Burns DD, Teppang KL, Lee RW, Lokensgard ME, Purse BW. Fluorescence turn-on sensing of DNA duplex formation by a tricyclic cytidine analogue. *J Am Chem Soc.* 2017; 139:1372–1375. [PubMed: 28080035]
126. Li Y, Fin A, McCoy L, Tor Y. Polymerase-mediated site-specific incorporation of a synthetic fluorescent isomorphous G surrogate into RNA. *Angew Chem Int Ed.* 2017; 56:1303–1307.
127. Merkel M, Dehmel L, Ernsting NP, Wagenknecht HA. An isosteric and fluorescent DNA base pair consisting of 4-aminophthalimide and 2,4-diaminopyrimidine as C-nucleosides. *Angew Chem Int Ed.* 2017; 56:384–388.
128. Xu W, Zeng Z, Jiang JH, Chang YT, Yuan L. Discerning the chemistry in individual organelles with small-molecule fluorescent probes. *Angew Chem Int Ed.* 2016; 55:13658–13699.

Figure 1. Structures of natural nucleobases and fluorescent analogues

a, Nucleobases carry hereditary information through specific hydrogen bonding and steric interactions; that is, adenine with thymine/uracil and guanine with cytosine. Adjacent nucleobases interact with each other via π -stacking, thus forming the double-helix conformation. **b**, Potential chemical modification sites of canonical pyrimidines. Examples cover ring substitution, conjugated linker extension and ring fusion. Yellow shading highlights modified structures and the purple bonds and atoms are base-pairing moieties. **c**, Potential chemical modification sites of canonical purines. Examples include ring structure modification, substitution and ring fusion. **d**, Examples of non-canonical fluorescent nucleobases, showing larger sizes that allow coverage of the redder end of the spectrum. Colours indicate approximate emission hue: ~370–390 nm, violet; 410–430 nm, blue; 440–470 nm, cyan; 480–510 nm, green; 520–540 nm, yellow; 550–600 nm, orange; >600 nm, red.

Figure 2. A chronicle of fluorescent nucleobase development

In 1969, Stryer reported the first canonical fluorescent nucleobases, featuring 2AP and formycin; in 1972, Leonard synthesized the first non-canonical fluorescent nucleobase, etheno-dA. The fluorescent nucleobase family gradually expanded during the following two decades, including m⁵K reported in 1990 by McLaughlin and pteridines by Hawkins in 1995. C-glycosidic nucleobases with hydrocarbon fluorophores directly attached to the sugar were first reported by Kool in 1996. Since 2000, the number of fluorescent nucleobases has increased markedly, featuring notable examples such as BPP by Saito and Okamoto, tC by and Wilhelmsson, dU^{Phen} (Hocek) and thieno-appended analogues of Tor. Dozens of non-canonical nucleobases have been introduced since 2010 (examples by Kool shown here). Hundreds of fluorescent nucleobases have been reported thus far and their numbers and applications are still rapidly expanding.

Figure 3. Molecular strategies for design of fluorescent purine analogues

a, Numbering of the purine skeleton. **b**, Potential modification sites of adenine and guanine. Expansion and modification mostly occur on positions 2, 5 and 8 of adenine and positions 5, 6, 7 and 8 of guanine, as indicated by the arrows. **c**, Examples of purine ring fusion modifications. **d**, Examples of extending the purine scaffold through conjugated linkers. **e**, Examples of purine substituent modifications. **f**, Examples of purine ring fusions. Native base-pairing groups are rendered in pink, whereas fluorescence modifications are shown in blue.

Figure 4. Molecular strategies for design of fluorescent pyrimidine analogues

a, Numbering of the pyrimidine scaffold. **b**, Potential modification sites of thymine and cytosine. Expansion and modification mostly occur on positions 5 and 6 of thymine and positions 4, 5 and 6 of cytosine, as indicated by the arrows. **c**, Modification of the pyrimidine substituents at position 5. This category includes changes at only one especially versatile chemical position. **d**, Modification of the pyrimidine substituents at positions 5 and 6. The extensions form a new six-membered ring containing heteroatoms, which can be further expanded for fluorescent modification. **e**, Extension of pyrimidines via conjugated linkers. **f**, Pyrimidine ring fusion.

Figure 5. Molecular features and examples of non-canonical fluorescent nucleobases

a, Nucleobases composed of aromatic hydrocarbons. **b**, Nucleobases composed of planar heterocyclic fluorophores. The lack of hydrogen bonding and weaker π -stacking are compensated by versatile energy states brought from the heteroatoms. These fluorophores contribute to a broader spectrum of emission wavelengths. Functional groups can be added to expand functionality, such as metal binding. **c**, Examples of nucleobases based on hydrocarbons. **d**, Photoreaction of adjacent phenethynylpyrene nucleobases yields a colour change in emission. The left image shows phenylalkynylpyrene excimer emission whereas the right image shows pyrene monomer emission, both excited at 360 nm. **e**, C-glycosidic nucleobases based on known fluorophores. **f**, Simple heterocyclic nucleobases used in the detection of DNA repair activity. **g**, Nucleobase pairs based on shape complementarity. Although they lack hydrogen bonding, the conformation of these bases counterpart each other, thus forming unnatural base pairing.

Figure 6. Methods for incorporating fluorescent nucleobases into DNA or RNA

a, Direct oligonucleotide synthesis via synthesizer and phosphoramidite chemistry. The main steps of DNA synthesis are: (1) removal of 4,4'-dimethoxytrityl (DMT) under acidic conditions; (2) coupling of the nucleoside phosphoramidite with the growing chain; and (3) oxidation of the phosphorus linkage. **b**, Post-synthesis modification using mild coupling methods or gene-editing methods. Fluorescent nucleobases equipped with organoboron or organostannane groups are coupled to halogen-labelled nucleobases in DNA strands. Gene-editing and ligation methods enzymatically join smaller labelled strands to make longer ones. **c**, Direct enzymatic incorporation using fluorescent nucleoside triphosphate derivatives. When the fluorescent nucleobases are labelled in the primers or supplied as free nucleobases in the pool, polymerases that recognize them can incorporate the fluorescent nucleobases into DNA sequences. TP, triphosphate; dNTP, deoxynucleoside triphosphate.

Figure 7. Examples of applications of fluorescent nucleobases

a, Fluorogenic sensing of a demethylation enzyme, ALKBH3. The emission of pyrene is initially quenched by the positive charge of 1-methylated adenine (m1A). When ALKBH3 demethylates m1A, the quenching effect is removed and a signal is generated. **b**, Fluorogenic analysis of adenine-to-inosine RNA editing enzyme. The emission maxima of the thiolated adenine (thA) and inosine (thI) are different. Hence by measuring the intensity of thA and thI at their respective maximal wavelengths, the activity of the A-to-I enzyme can also be measured. dsRNA, double-stranded RNA; ssDNA, single-stranded DNA. **c**, Kinetic and thermodynamic investigation of the effects of mercury on DNA metabolism. The fluorescent thymine can chelate mercury with another thymine ring and link DNA strands. This can be used to probe mercury metabolism *in vivo* and to study the effects of mercury on DNA status. k_1 , strand displacement rate constant; k_{-1} , reverse reaction rate constant. **d**, Visualization and analysis of human concentrative nucleoside transporters (hCNTs). The fluorescent nucleoside can enter the plasma membrane through the transporters, thus allowing the measurement of the transport activity. TGF- β 1, transforming growth factor β 1.

Table 1

Comparison of three labelling methods for incorporating fluorescent nucleobases into DNA or RNA.

Labelling methods	Advantages	Disadvantages
Direct chemical synthesis	Site-specific incorporation at any position Little or no constraint on fluorophore structure	High cost on preparative synthesis scales Requires access to DNA synthesizer ~100 nt or less in length
Post-synthesis modification	Site-specific incorporation at any position Less expensive than direct synthesis	Limited structural diversity available May require challenging purification
Enzymatic incorporation	Low cost Access to labelled DNAs/RNAs ~100–1,000 nt in length	Some constraints on positional labelling Fluorophore structure limited by enzyme constraints Base-pair choices limited