Enzymatic synthesis of chemically modified DNA for analytical and diagnostic applications

A thesis submitted to the Board of Mathematical, Physical and Life Sciences Division,

Department of Chemistry in partial fulfilment for the degree of

Doctor of Philosophy

At the

University of Oxford

Xiaomei Ren

St. Edmund Hall, Michaelmas Term 2015

Enzymatic synthesis of chemically modified DNA for analytical and diagnostic applications

Xiaomei Ren

(Doctor of Philosophy, St. Edmund Hall, Michaelmas Term, 2015)

Abstract

Functionalised nucleotides have found widespread adoption in applications ranging from genome labelling to disease diagnostics. A popular means of introducing these modifications into nucleic acids is *via* the copper-catalysed azide-alkyne cycloaddition reaction. However, the toxicity of copper limits its use *in vivo*. Therefore, the copper-free strain-promoted azide-alkyne cycloaddition (SPAAC) and inverse electron demand Diels-Alder (IEDDA) reactions were evaluated as a means of directly and indirectly labelling DNA and RNA. A variety of cyclooctyne, *trans*-cyclooctene and azide functionalised deoxyuridine triphosphates were synthesised and shown to be incorporated by primer extension, polymerase chain reaction (PCR) and reverse transcription. Bulky cyclooctyne and *trans*-cyclooctene triphosphates demonstrated moderate incorporation efficiency, while azide nucleotides with smaller functional groups were better substrates. PCR products of *ca.* 500 base pairs with 100% sequence fidelity can be formed, where all thymidine sites are replaced with azidohexanamidopropargyl deoxyuridine. These modified duplex products were efficiently labelled with fluorophores using the SPAAC and IEDDA reactions, and can be converted to fluorescent single-stranded probes through exonuclease digestion and streptavidin magnetic bead separation of the duplex followed by fluorophore-labelling. To expand the utility of this methodology, a new and simple single-tube dual labelling strategy was demonstrated to give a range of mixed dual

coloured probes. Nucleosides bearing azide, *trans*-cylcooctene modifications and a fluorescent base analogue were evaluated for direct cellular DNA labelling and detection. When cultured adenocarcinoma A549 cells were treated with these nucleosides, then chemically fixed and fluorescently labelled by the SPAAC or IEDDA reaction, cytoplasmic staining without nuclear DNA labelling was observed. To improve the efficiency of fluorescent labelling of genomic DNA in cultured cells, phenyl isopropyl alanine-protected nucleoside monophosphates were prepared in an attempt to directly deliver 5′-phosphorylated nucleosides into the cells. Preliminary fluorescence studies indicate bright foci in addition to cytoplasmic staining, the origins of which require further study.

Acknowledgements

For my amazing DPhil study, I appreciate the help and support I received from many people and organisations. Without them, this DPhil would not be achievable.

First of all, I need to say a big thank you to Tom for offering me a DPhil position in his group. He is a nice mentor and provides a fantastic working environment to study nucleic acids. All his guidance and support during this DPhil are extremely appreciated. Thanks to his move from Southampton to Oxford, I enjoyed working and living in two lovely cities.

I am also grateful for the help and encouragement I received from the current and previous members in the Brown group. Many thanks to Afaf for all her guidance and co-supervision during my DPhil work. She is the person who showed me how to deal with oligonucleotides and helped me throughout my project. Huge thanks to Arun, Sara, Adam, Yssy and Joanna for their help on revising this thesis. I also appreciate the support and company of Marta, Simon, Nittaya, Chengguang, Xiong, Agnes, Jieqiong, Anna, Pietro, Pui and Ellia. They are extremely good friends and we had a fantastic time together.

I would also like to acknowledge the funding I received from University of Southampton, University of Oxford (Department of Chemistry) and China Scholarship Council. Big thanks to Tom again for his help in applying for and providing funding.

A lot of thanks to the NMR and Mass Spectrometry service staff in the University of Southampton and University of Oxford. Thanks to all the PTCL building service staff for their kind help.

Finally, I want to say thank you to my lovely mom, dad and boyfriend who have supported me in what I like to do and have shared the happiness and difficulties during these four years.

Abbreviations

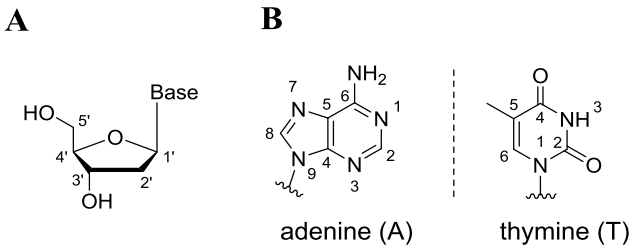
A	Absorbance
A/dA/dATP	Adenine/deoxyadenosine/deoxyadenosine triphosphate
AMV	Avian Myeloblastosis Virus
BCN	Bicyclo[6.1.0]nonyne
Boc	*t*-Butyloxycarbonyl
BSA	Bovine serum albumin
c	Molar concentration
C/dC/dCTP	Cytosine/deoxycytidine/deoxycytidine triphosphate
CuAAC	Copper-catalysed azide-alkyne cycloaddition
Cy3	Trimethine cyanine dye
Cy5	Pentamethine cyanine dye
DAPI	4′,6-Diamidino-2-phenylindole
DCM	Dichloromethane
DIBO	Dibenzocyclooctyne
DIPEA	*N*, *N*-Diisopropylethylamine
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DMT	4,4′-Dimethoxytrityl
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
dsDNA	Double-stranded DNA
EDTA	Ethylenediaminetetraacetic acid
eq	Equivalent

ESI	Electrospray ionisation
FAM	Carboxyfluorescein
FISH	Fluorescence *in situ* hybridisation
FRET	Fluorescence resonance energy transfer
G/dGTP	Guanine/deoxyguanosine triphosphate
HRMS	High resolution mass spectrometry
IEDDA	Inverse electron demand Diels-Alder
J	Coupling constant
LRMS	Low resolution mass spectrometry
M	Molar concentration
m/z	Mass/charge
M-MuLV	Moloney Murine Leukemia Virus
M_w	Molecular weight
NHS	*N*-Hydroxysuccinimide
NMR	Nuclear magnetic resonance
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PNP	*p*-Nitrophenyl
R_f	Retention factor
RNA	Ribonucleic acid
RNase	Ribonuclease
RP-HPLC	Reversed-phase high performance liquid chromatography
RT	Room temperature

SELEX	Systematic evolution of ligands by exponential enrichment
SNP	Single nucleotide polymorphism
SPAAC	Strain-promoted azide-alkyne cycloaddition
ssDNA	Single-stranded DNA
T/dT/dTTP	Thymine/deoxythymidine/deoxythymidine triphosphate
TBE	Tri-Borate-EDTA
TCO	*trans*-Cyclooctene
TEA	Triethylamine
TEAB	Triethylammonium bicarbonate
TFA	Trifluoroacetic acid/trifluoroacetyl
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TOF	Time-of-flight
U/UTP/dU/dUTP	Uracil/uridine triphosphate/deoxyuridine/deoxyuridine triphosphate
UV	Ultraviolet
δ	Chemical shift in parts per million
ε	Extinction coefficient
$\lambda_{ex}/\lambda_{em}$	Excitation/emission wavelength
Φ	Fluorescence quantum yield

Explanatory note

1. The following numbering system is used throughout this thesis for nucleosides (**A**) and purine/pyrimidine bases (**B**).



2. Oligonucleotides are written from the 5′-end to the 3′-end.

Contents

Chapter 1 Introduction

1 Introduction

1.1 Nucleic acids

1.1.1 DNA structure

Deoxyribonucleic acid (DNA) carries the genetic information in the majority of living systems. It is composed of monomers termed nucleotides. These nucleotides consist of phosphate groups, deoxyribose sugars and heterocyclic bases (Figure 1.1.A). Nucleotides without the 5′-phosphate groups are named nucleosides (Figure 1.1.B). Four canonical bases exist and can be divided into two groups – purines (adenine and guanine) and pyrimidines (thymine and cytosine) (Figure 1.1.C).

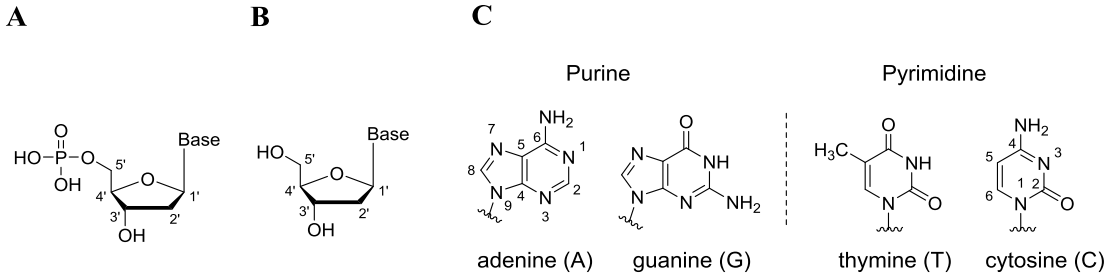


Figure 1.1. Chemical structures of deoxyribonucleotide (**A**), deoxyribonucleoside (**B**) and heterocyclic bases of DNA (**C**).

The heterocyclic base is linked to the 1′-position of deoxyribose by an *N*-glycosidic bond to give a deoxyribonucleoside (Figure 1.2). Rotation of this bond affords *syn*- and *anti*-conformations. Nucleosides are usually in conformations around the *anti* range due to restriction of bond rotation and steric interference from the sugar residue in the *syn*-conformation.

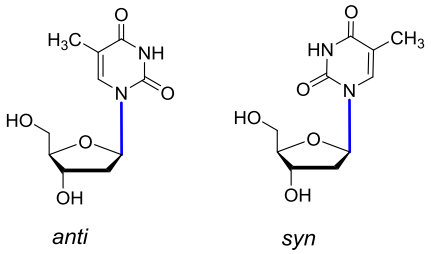


Figure 1.2. Extreme *anti*- and *syn*-conformations of pyrimidine nucleosides.

1.1.2 RNA structure

In some organisms, for example retroviruses, genetic information is stored in ribonucleic acids (RNA) as opposed to DNA. However, generally RNA is a transient copy of DNA, which is transported to ribosomes to be decoded into specific sequences of amino acids that form proteins. The processes of copying and decoding are called transcription and translation respectively, in which messenger RNA, ribosomal RNA and transfer RNA are involved. Another type of RNA is termed as small RNA, which is non-coding but regulates the translation of messenger RNA.[1] RNA is chemically similar to DNA. Its sugar moiety has a hydroxyl group on the 2′-position and contains the pyrimidine base uracil (U, Figure 1.3.A) instead of thymine in addition to A, G and C (Figure 1.1).

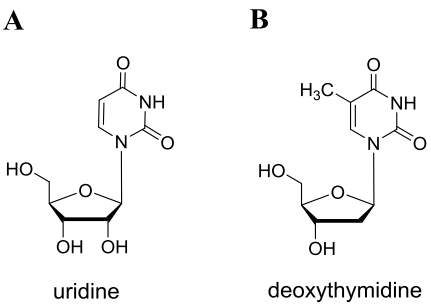


Figure 1.3. Chemical structures of uridine (**A**) and deoxythymidine (thymidine, **B**) nucleosides.

1.1.3 Nucleic acid secondary structure

The nucleotides introduced so far are incorporated within organisms into linear polymers called nucleic acids. The 5′-phosphate group of one nucleotide and the 3′-hydroxyl group of another are joined together by a phosphodiester bond. In physiological conditions (pH 6.4 – 7.4), all phosphodiester groups are deprotonated to give highly negatively charged nucleic acid strands (Figure 1.4.A). They have directionality with a 5′-hydroxyl group on one end of a strand and a 3′-hydroxyl group on the other end. By convention, the sequences of nucleic acids are written from the 5′-end to the 3′-end.

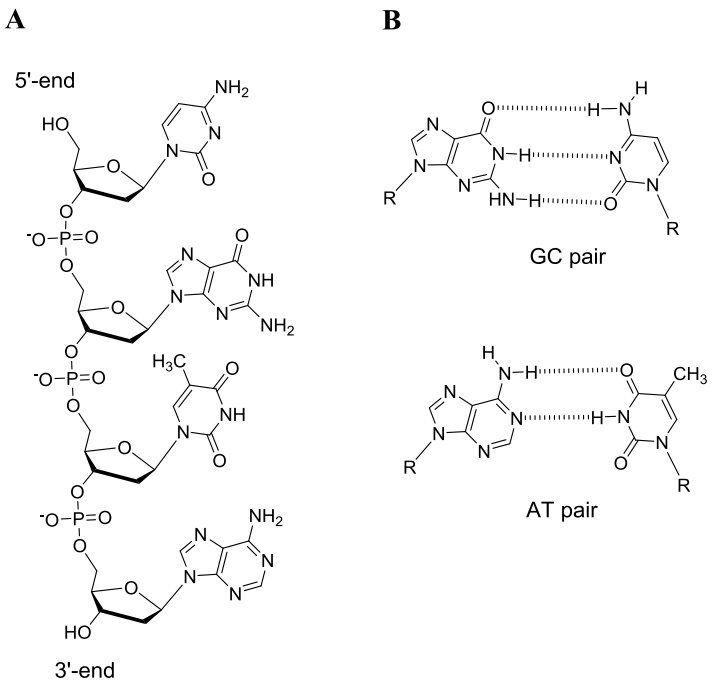


Figure 1.4. **A.** Chemical structure of oligonucleotide dCGTA. **B.** Watson-Crick base pairs (GC and AT pairs).

James Watson and Francis Crick found that nucleic acids can form a double-stranded helical structure.[2] This was based on the findings of an equal ratio between purine and pyrimidine bases by Chargaff, *i.e.* more specifically A = T and G = C, and X-ray fibre diffraction studies on DNA by Wilkins and Franklin.[3, 4] The two nucleic acid strands are

bound together *via* hydrogen bonds between nucleobases and are referred to as Watson-Crick base pairs. Adenine (A) pairs with thymidine (T) (or uridine in RNA) *via* two hydrogen bonds; guanine (G) pairs with cytosine (C) by three hydrogen bonds (Figure 1.4.B). The two strands are complementary and antiparallel, where the 5′-end of one strand is opposite to the 3′-end of the other.

Double-stranded DNA (dsDNA) coils to form a right-handed helix (B-form, Figure 1.5.A). The heterocyclic base pairs are positioned in the centre of the helix; the negatively charged sugar-phosphate chains form the external hydrophilic shield. The planar hydrophobic bases stack on top of each other. The distance between two continuous base pairs along the helix axis is around 3.4 Å (Figure 1.5.B). One full 360° turn of the DNA double helix contains 10.5 base pairs. Both hydrogen bonding and base stacking play important roles in stabilising the DNA duplex. The B-form helix has a spacious major groove and a small narrow minor groove, providing binding opportunities for external molecules. Large molecules interact with DNA preferentially in the major groove;[5, 6] only small molecules and a few proteins (*e.g.* integration host factor) can bind in the minor groove due to steric hindrance.[7, 8] When a RNA strand binds to a complementary DNA or RNA strand, they usually form a right-handed duplex with a different helical shape (A-form, Figure 1.5.A). Compared to the B-form, A-form helices have deeper major grooves and shallower minor grooves. The hydrophobic base pairs are pushed towards the outer surface of the A-form helix.

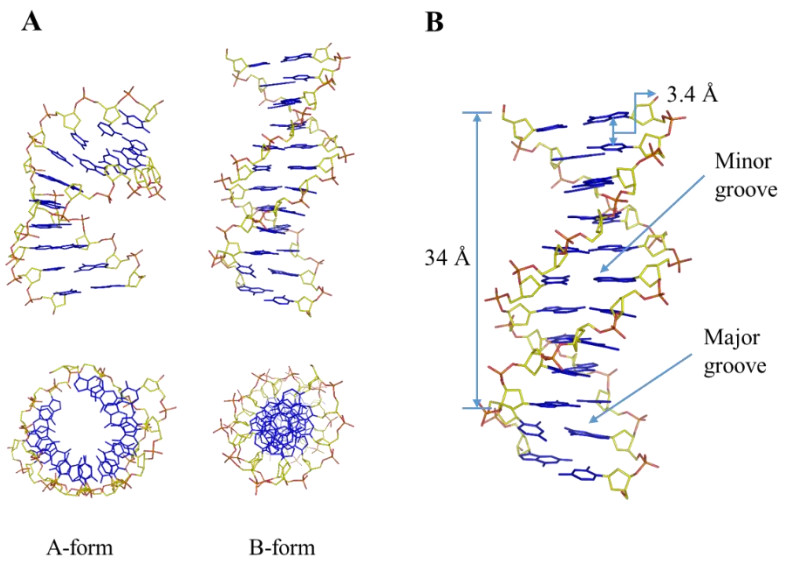


Figure 1.5. Structures of A- and B-form DNA duplexes (protein data base 440D and 1BNA, respectively)[9, 10] **A.** Top and side views. **B.** Structural parameters of B-form.

1.1.4 Modified nucleosides

There are various modified nucleosides in natural living systems which have specific biological roles, *e.g.* 5-methyldeoxycytidine (involved in epigenetic gene expression regulation, Figure 1.6.A).[11-13] Apart from these modifications, artificial chemical substitutions have been widely introduced to expand the applications of nucleic acids.[14, 15] Modifications can be introduced to the base, sugar and phosphate linker. For example, the arabinose sugar analogue of adenosine (*ara*-adenosine, Figure 1.6.B) is useful as an anti-viral drug;[16] base analogues such as 3-nitropyrrole (Figure 1.6.C) act as universal bases;[17-19] peptide nucleic acid oligomers (Figure 1.6.D) have been used for diagnostic and other applications since they do not have a polyanionic backbone.[20, 21] This enables them to bind to DNA and RNA with higher affinity compared to conventional oligonucleotides.

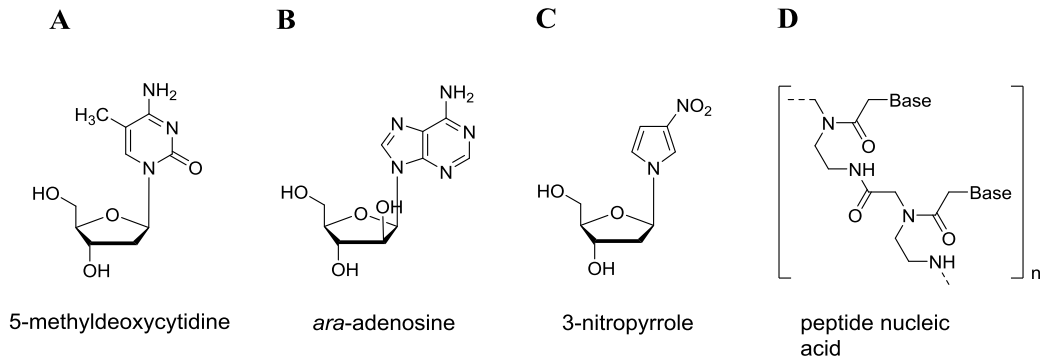


Figure 1.6. Chemical structures of modified nucleosides and DNA analogues. **A.** 5-Methyl deoxycytidine (dC);[11] **B.** *ara*-Adenosine;[16] **C.** 3-Nitropyrrole deoxyriboside;[19] **D.** Peptide nucleic acid.[20]

1.1.5 Thermal stability of DNA duplex

After the introduction of modifications into oligonucleotides, it is important to investigate their thermal stability. The melting temperature (T_m) of a nucleic acid duplex is the point at which half of the DNA duplex is dissociated into single strands, and can act as a measure of binding affinity to the complementary strand. This temperature is dependent on duplex length, base stacking and hydrogen bonding; longer or GC-rich oligonucleotides have higher T_m (*i.e.* higher stability) than shorter or AT-rich duplexes. Additionally, the stability of the DNA double helix is affected by the local solvent environment such as the nature and concentration of the counter-ions (cations) in solution.

DNA nucleotide bases are ultraviolet (UV) active. The UV signal of duplex DNA is lower than single strands due to its stronger π-stacking interaction (Figure 1.7.A). Therefore, UV detection is commonly used to acquire the T_m of a DNA duplex. However, at lower concentrations, the UV signal is too weak to be detected. In this case, more sensitive fluorescence melting can be used. The T_m of duplex DNA at nanomolar (nM) concentrations can be determined using the DNA intercalator SYBR Green, which

exhibits at least eleven times higher fluorescence after intercalation into DNA duplexes compared to the single-stranded form (Figure 1.7.B and C).[22] This intercalator has been widely used to monitor amplification in the polymerase chain reaction (PCR) and investigate the thermal stability of PCR products.[23, 24] The principle of fluorescence and its applications in nucleic acid research will be discussed in the following section.

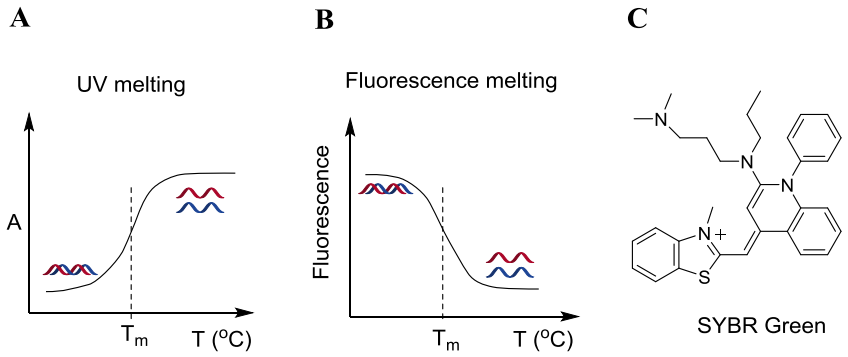


Figure 1.7. Schematic melting curves for UV melting (**A**) and fluorescence melting (**B**) using a DNA intercalator. T_m = melting temperature. **C.** Chemical structure of SYBR Green.

1.2 Fluorescence

1.2.1 Principles of fluorescence

Fluorescence is the emission of light by a molecule upon excitation by a photon of light (Figure 1.8). After absorbing the light, an electron in the fluorescent molecule is excited from its ground state S_0 to a higher vibrational level (S_1, S_2 or S_n). This unstable electron quickly releases some energy to the solvent *via* internal conversion and returns to its lowest vibrational level S_1. The electron then returns from S_1 back to its ground state with the emission of light (fluorescence). This process also can happen *via* non-radiative decay (*i.e.* releasing the energy by heat). Due to the energy loss in internal conversion and solvent reorganisation, the fluorescence emission wavelength (λ_{em}) is longer than the excitation wavelength (λ_{ex}). The difference is known as the Stokes shift. Fluorescence

quantum yield (Φ) is an important parameter to evaluate the fluorescent efficiency of a fluorophore; it is the ratio of the number of photons emitted to the number of photons absorbed. The fluorescence signal can be monitored with a lower limit of detection than UV-visible absorption and can be used to detect biomolecules at physiological concentrations *via* fluorescent labelling.

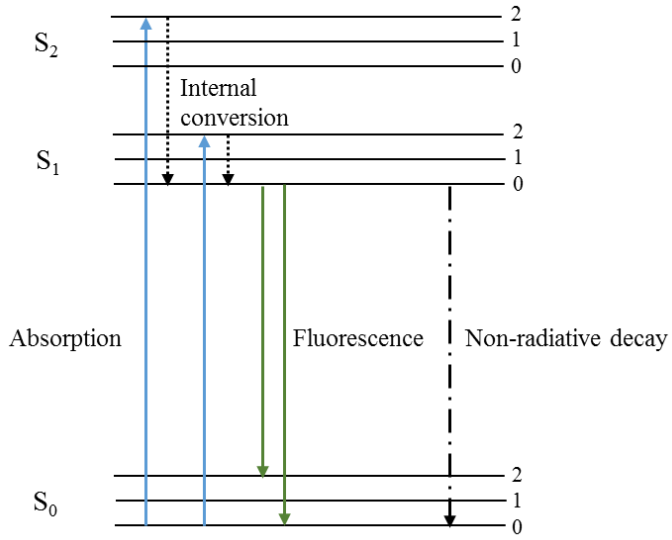


Figure 1.8. Jablonski diagram. S_0, S_1, S_2 are the singlet ground, first and second electronic states respectively. Each energy level has a number of vibrational energy levels shown as 0, 1 and 2.

1.2.2 Applications of fluorescence in nucleic acids

DNA can be labelled with strongly fluorescent dyes to facilitate its detection. This approach is safer than traditional radioisotopic labelling techniques, which are potentially harmful to the user. For *in vitro* studies, fluorescent labelling and detection are applied extensively in nucleic acid research, for example in DNA sequencing, forensic science and genetic analysis.[25-30] Fluorescence resonance energy transfer (FRET) based DNA probes, utilising the interaction between a donor and an acceptor fluorophore (or a quencher), are employed in DNA sequencing, PCR monitoring, mutation detection and

single nucleotide polymorphism (SNP) genotyping.[30, 31] Fluorescence *in situ* hybridisation (FISH), employing fluorescent DNA probes, has been applied to localise specific gene sequences of interest, to map genes in the chromosome and to visualise intracellular RNA.[32, 33] By using different fluorophores with distinct fluorescence properties, multiple genetic sequences can be detected simultaneously.[34] Furthermore, DNA intercalators (*e.g.* Hoechst) or modified DNA bases have been widely used to study the complex structure of DNA in the nuclei of living or fixed cells *via* fluorescence microscopy, which is an increasingly important technique in cell biology.[35-38]

1.3 Synthesis of fluorescent nucleic acids

The above applications require the preparation of fluorescently labelled oligonucleotides. This can be achieved using automated solid phase oligonucleotide synthesis or enzymatic synthesis using polymerases, often in combination with a post-synthetic labelling step.[39]

1.3.1 Solid phase synthesis of oligonucleotides

An oligonucleotide is a short strand of DNA or RNA, usually up to around 100 nucleotides in length. Automated solid-phase oligonucleotide synthesis is the most popular chemical method to assemble oligonucleotides using phosphoramidite monomers developed by Beaucage and Caruthers.[40-42] The oligonucleotides are synthesised on a solid support from the 3′-end to the 5′-end. Therefore the 3′-end must be bound to the solid support and one nucleotide is added in each synthesis cycle. Each cycle consists of a series of steps, namely activation, coupling, capping, oxidation and detritylation (Figure 1.9). After the synthesis is complete, oligonucleotides are cleaved from the solid support and deprotected.

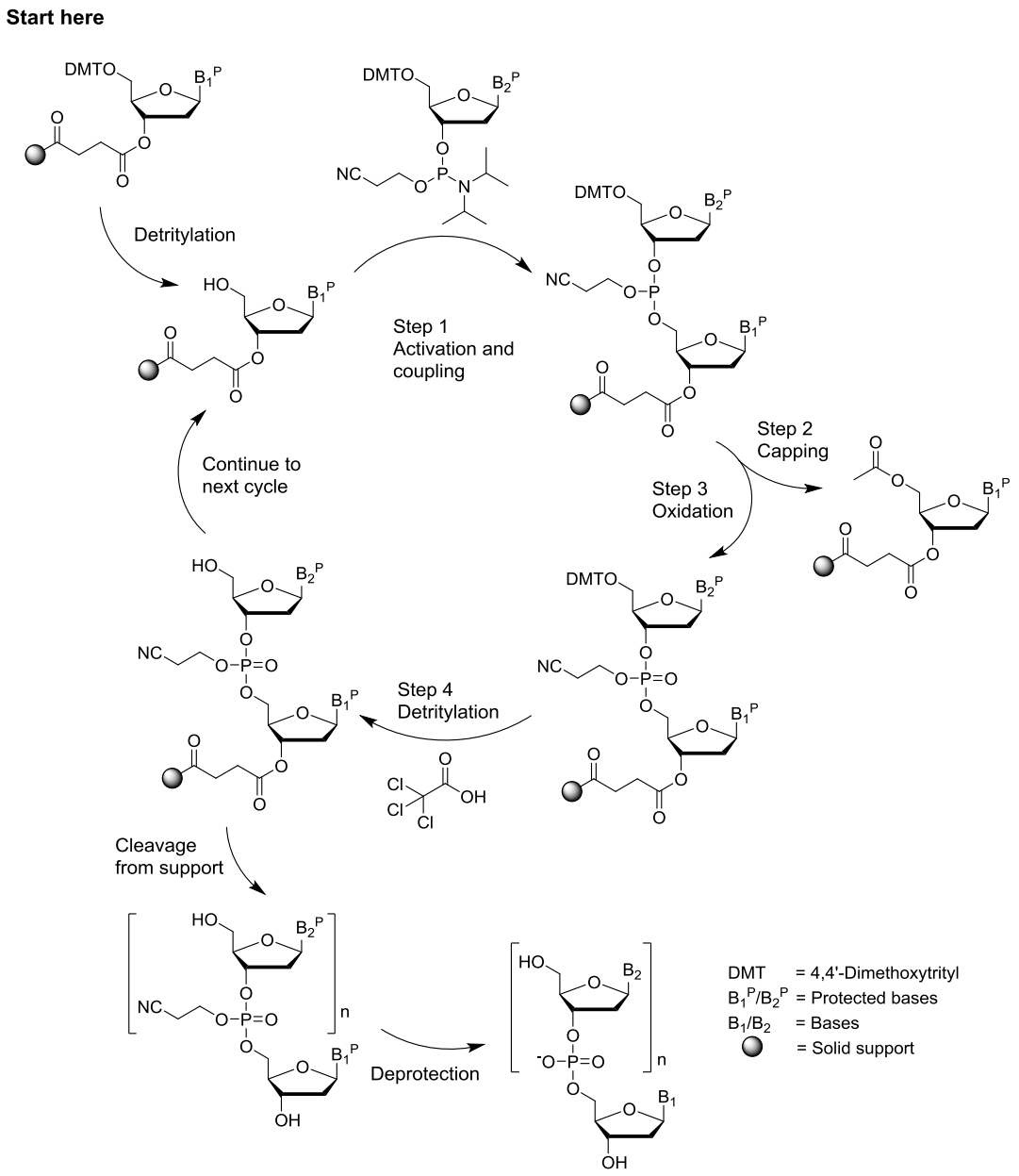


Figure 1.9. The synthesis cycle of the automated solid-phase oligonucleotide synthesis using phosphoramidite monomers.

1.3.1.1 Solid phase synthesis of fluorescent oligonucleotides

The advantages of solid phase synthesis lie in the rapid preparation of oligonucleotides; the ability to use large excesses of solution phase reagents to drive reactions to completion; ease of removal of excess reagents; and the ability to introduce a wide variety of modified

nucleotides at specific positions. A broad range of fluorescent tags have been introduced into nucleotides directly by solid phase synthesis.[39] These fluorescent phosphoramidites have to be stable during the harsh solid phase synthesis and deprotection (acidic, alkaline and oxidative) conditions. Common fluorophores, for example fluorescein and trimethine cyanine dye (Cy3), are sufficiently stable to survive solid phase synthesis and their phosphoramidites are commercially available (Figure 1.10.A). These fluorophores can be coupled to the 5′- or 3′-end of oligonucleotides. This allows the fluorophores to be placed relatively far from the nucleobases *via* covalent linkages thereby causing minimal disruption of base pairing. Some applications, such as SNP detection by Hybeacons, require the fluorophores to be located in the middle of nucleotide strands.[43, 44] Different fluorophores have been attached to the 5-position of the pyrimidines (T, U and C, Figure 1.10.B), which point into the major groove of DNA duplexes.[45] The 2′-position of the ribose, located around the minor groove of the DNA or RNA duplexes, has also been widely modified with fluorescent groups (Figure 1.10.C).[46] All these modification sites are distant from the base-pair hydrogen bonding sites, minimising distortions of the duplex. In addition to these modifications, fluorescent base analogues have been synthesised, which can pair with natural bases (Figure 1.10.D).[47-49] Fluorescence detection and analysis of oligonucleotides using fluorescent base analogues can give more precise structure information for DNA helix *via* accurate FRET measurements.[50, 51]

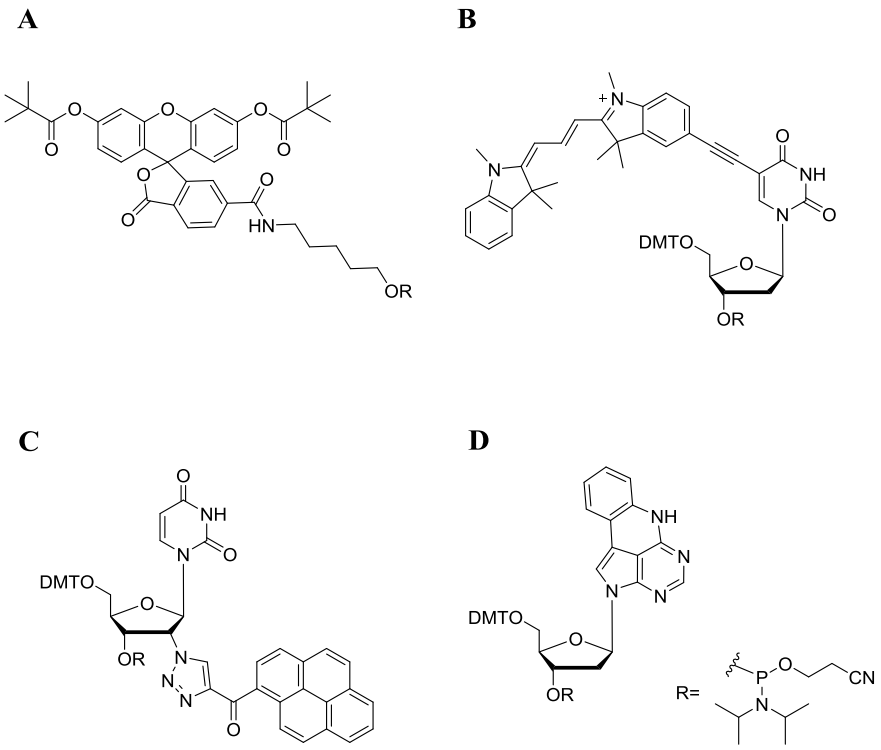


Figure 1.10. Chemical structures of fluorescent tags used in solid phase oligonucleotide synthesis. **A.** Commercially available fluorescein phosphoramidite; **B.** 5-Cy3 modified deoxythymidine (dT) phosphoramidite;[45] **C.** 2′-Pyrene modified deoxyuridine (dU) phosphoramidite;[46] **D.** Fluorescent adenine analogue.[48]

Failure sequences accumulate during solid-phase synthesis cycles, especially for modified monomers and restrict the length of synthetic oligonucleotides attainable to around 150 nucleotides. For example, if each step in the solid phase synthesis has an average coupling yield of 98.5%, the final yield of a 150-mer unmodified oligonucleotide is only 10.5%. The best achievable coupling yields are around 99.5% but even this makes it hard to make sequences longer than 200 bases, tiny by biological standards. For applications requiring oligonucleotides longer than this, such as gene synthesis and mapping,[52] alternative methods must be sought. Moreover, the incorporation of fluorescent labels with inherent sensitivity to base, oxidation or acid is obviously precluded in solid-phase oligonucleotide synthesis.[53] A better strategy is to synthesise the

fluorescent-modified DNA *via* enzymatic synthesis (Section 1.3.2) or to introduce sensitive fluorophores post-synthetically (Section 1.4).

1.3.2 Enzymatic synthesis of nucleic acids

DNA is assembled from nucleoside triphosphates by DNA polymerases in living systems. One parent DNA strand acts as a template for the production of a new strand that is complementary to this template. Enzymatic synthesis using polymerase enzymes is much faster and more efficient than chemical synthesis, but the synthesis scale is usually smaller.

1.3.2.1 DNA polymerase enzymes

DNA polymerases, RNA polymerases and reverse transcriptases are the key catalysts for enzymatic nucleic acid synthesis. Structural analyses of DNA polymerases show that they have three domains in common (the palm, thumb and finger domains, Figure 1.11).[54] The palm domain catalyses the transfer of phosphate groups. The finger domain plays a role in binding the nucleoside triphosphate and its paired template base together. The thumb domain is responsible for the processivity, translocation and positioning of the DNA duplex. Together they catalyse DNA replication *via* a mechanism involving two divalent metal ions (Mg^{2+} or Mn^{2+}).[55, 56]

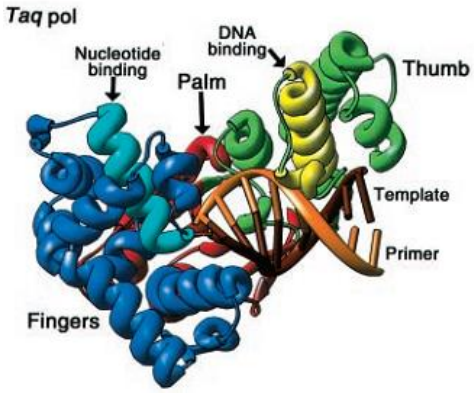


Figure 1.11. Overall structure of Taq polymerase with DNA template and primer.[54] Picture reproduced with permission of the American Society for Biochemistry and Molecular Biology (Appendix 8.1).

Polymerases can be divided into different families based on their amino acid sequences and structures; for general reviews see the references from T. A. Steitz.[54, 57] Family A polymerases are also referred to as the DNA polymerase I family, and include *Thermus aquaticus* (Taq) DNA polymerase and Pol-I from *Escherichia coli* (including its truncated Klenow fragment). Family B polymerase or the DNA polymerase α family includes all eukaryotic DNA replicating polymerases, T4 and RB69 phage polymerases. The reverse transcriptase family is a family of RNA dependent DNA polymerases using RNA strands as templates to produce complementary DNA products. The polymerases from these three families have different finger and thumb domains.

Studies on enzymatic incorporation of modified triphosphates show that family B polymerases are generally more efficient than family A polymerases.[58, 59] Some polymerases have been engineered to accept a broad range of modified triphosphates for non-natural oligonucleotide synthesis, for example KOD and Therminator™ II polymerases.[58, 60] Therefore, an alternative strategy to incorporate fluorescent bases into DNA is to employ these enzymes. DNA and single-stranded RNA can be synthesised *in vitro* using primer extension, PCR, transcription and reverse transcription.[39, 59, 61] These methods can yield densely modified nucleic acids up to thousands of base pairs long.

1.3.2.2 Primer extension

Primer extension requires an oligonucleotide primer which is complementary to a region near the 3′-end of a DNA template strand (Figure 1.12). The primer anneals to the template and forms a duplex. DNA polymerase can then be used to incorporate natural or modified deoxynucleoside 5′-triphosphates to the 3′-end of the primer until it reaches the 5′-end of the template. For ease of detection and analysis, the primer is commonly functionalised with a radioactive or fluorescent tag on the 5′-end. Primer extension is a

very efficient methodology for the introduction of multiple modifications into one strand of a DNA duplex.

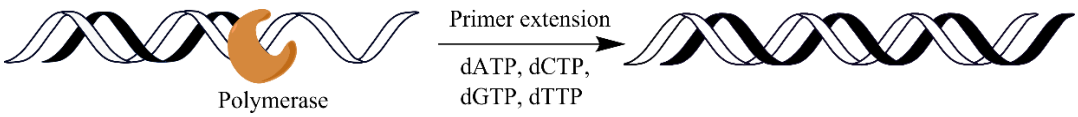


Figure 1.12. Primer extension using a DNA polymerase (yellow) and four canonical deoxynucleoside triphosphates, which can be replaced or supplemented with modified triphosphates.

1.3.2.3 Polymerase chain reaction

The primer extension approach is restricted to the incorporation of modified nucleotides in a linear reaction, where one template oligonucleotide gives rise to one product. In contrast, PCR, developed by Kary Mullis in the early 1980s, is an amplification procedure which can produce large amounts of densely modified DNA in both strands from a very small amount of template DNA (single strand or duplex).[62] PCR amplification requires a template and two primers which are complementary to the 3′-end of the sense- and anti-sense template strands respectively. A typical procedure consists of denaturation, annealing, and elongation (Figure 1.13). This process is carried out in multiple cycles at elevated temperatures and therefore requires thermostable DNA polymerases. The reaction is exponential, and typically around 30 cycles of PCR are carried out.

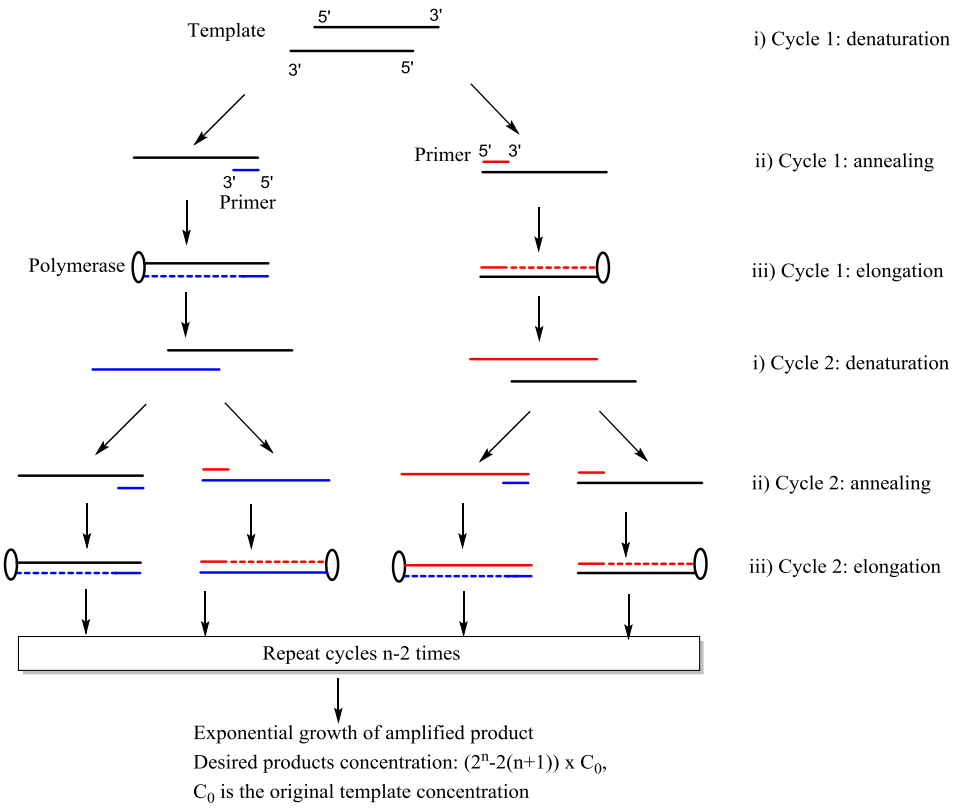


Figure 1.13. PCR procedure using a DNA duplex as a template. i) Denaturation at 94 to 96 ºC; ii) annealing at 50 to 60 ºC; iii) elongation at 72 ºC. The first two cycles are shown. The red and blue lines represent the amplified products and the circles stand for polymerases. In real cases the template can be much longer than the final PCR product.

One drawback of PCR and primer extension is that they produce duplex products, which cannot be used as single-stranded DNA (ssDNA) probes directly for FISH and other probe-based applications. An extra step such as exonuclease digestion or magnetic bead capture of one strand is required. To avoid this complication, asymmetric PCR (APCR) can also be used to produce mainly single-stranded products.[63] One primer in the APCR reaction is in large excess over the other primer (*e.g.*10:1 or 20:1 ratio). After consuming the limited primer, each reaction cycle can thereafter only produce one of the two strands. When the reaction is complete, the majority of products are single-stranded.

1.3.2.4 Enzymatic synthesis of fluorescent nucleic acids

In enzymatic fluorescent DNA synthesis, fluorophores are commonly attached to the 7-position of purine bases and the 5-position of pyrimidine bases.[64-67] Modifications on these positions are distant from the base-pair hydrogen bonding sites and can be accepted by different DNA and RNA polymerases. The Winter and Amacker groups synthesised a wide range of fluorophore-labelled triphosphates based on all four bases (A, C, G and T, Figure 1.14.A).[66, 67] Densely modified DNA products can be prepared by mixing different base-modified triphosphates, however, the enzymatic incorporation efficiency varied dramatically between different dyes. Deoxyuridine triphosphates (dUTPs) with fluorescein linked *via* rigid linkers (Figure 1.14.B) were studied by Burgess and co-workers.[64] The triphosphates with long conjugated linkers were incorporated better than those with shorter ethynyl linkers when using TaqFS polymerase.

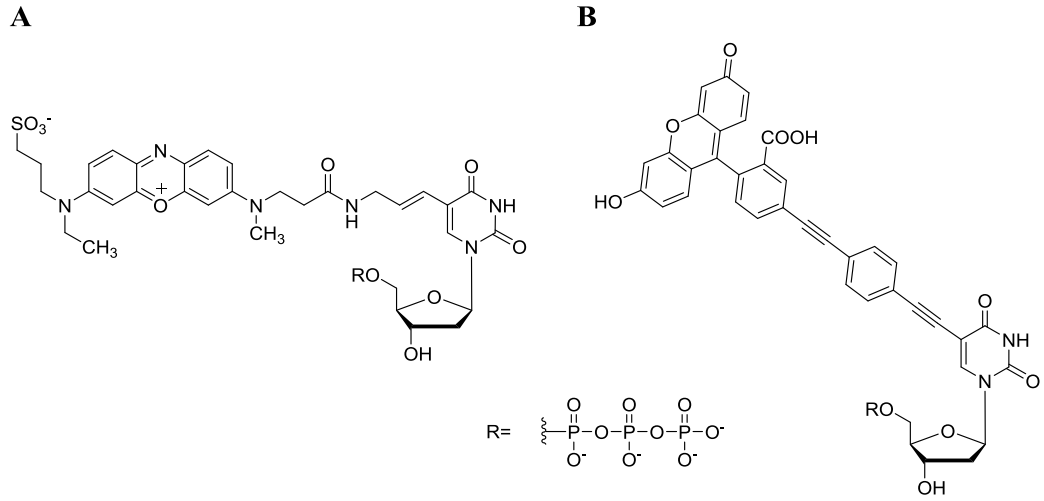


Figure 1.14. Chemical structures of fluorescent modified deoxynucleoside triphosphates for enzymatic DNA synthesis. **A.** Evoblue modified dUTP;[66] **B.** Fluorescein modified dUTP.[64]

In most cases, the modified triphosphates were mixed with the corresponding natural deoxynucleoside triphosphates (dNTPs) to prevent the premature termination of

enzymatic extension, especially for PCR, as the bulky and hydrophobic fluorescent modified nucleosides are not good substrates for the polymerases.[67] To prepare more densely labelled oligonucleotides, a post-labelling strategy is often used, which introduces simple, sterically undemanding functional groups into DNA followed by labelling with chemically activated fluorophores.[39, 68]

1.4 Nucleic acid labelling reactions

The preparation of highly modified DNA strands is desirable for a wide array of potential applications.[39, 68] Post-synthetic introduction of fluorescent labels allows for a higher degree of modularity, as changing the label does not require the synthesis of a novel nucleotide. Sensitive or reactive fluorophores can be introduced in this way instead of direct incorporation *via* solid phase or enzymatic synthesis. Here four of the most commonly used chemical labelling reactions are discussed, but there are other reactions that can be used for DNA functionalisation.[39, 68, 69]

1.4.1 Amino-active ester coupling

The most widely utilised post labelling methodology is to prepare DNA bearing primary amino groups, which can react with *N*-hydroxysuccinimide (NHS) esters of different fluorescent dyes (Figure 1.15). The reaction requires a buffer system between pH 7.2 and 9.0, and yields an amide bond. However, the NHS ester hydrolyses during the reaction and therefore a large excess of it is required for this labelling strategy.

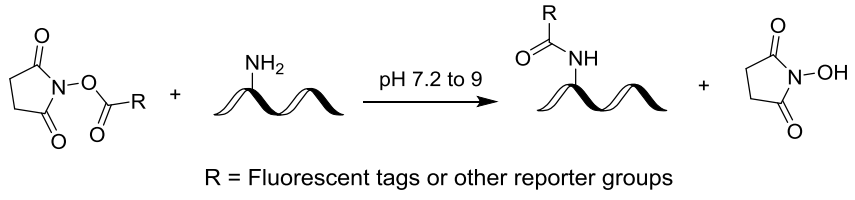


Figure 1.15. General protocol for the amino-DNA active ester coupling reaction.

Amino modified DNA and RNA have been synthesised by both solid phase synthesis and enzymatic incorporation (Figure 1.16).[65, 70-73] For solid phase synthesis, protecting groups (*i.e.* trifluoroacetyl group) are required to protect the amino group to prevent side-reactions (Figure 1.16.A).[65, 72] After the completion of oligonucleotide assembly, the protecting groups are removed to give the free primary amino group. The functionalised DNA strands can be further labelled with fluorescent reporter groups. In an enzymatic approach, Cox and Singer demonstrated the efficient incorporation of an amino modified dUTP (Figure 1.16.B) and its post-labelling with different fluorescent dyes.[73] A similar RNA uridine triphosphate was investigated by 't Hoen *et al.* to generate modified complementary RNA from RNA templates.[71] However, the hydrolysis of the NHS ester and long reaction times limit the use of this approach.

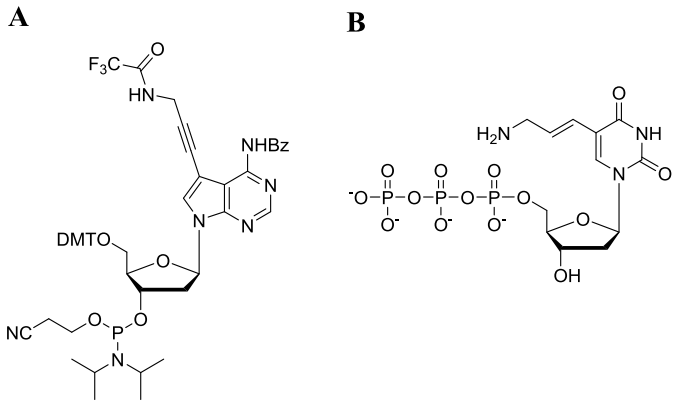


Figure 1.16. **A.** Amino modified 7-deazadeoxyadenine phosphoramidite used in solid phase DNA synthesis.[65] **B.** Amino modified dUTP used in enzymatic DNA synthesis.[73]

1.4.2 Copper-catalysed azide-alkyne cycloaddition reaction

Click chemistry is a concept invented by Sharpless *et al.* to ligate organic molecules selectively under mild conditions in the presence of various reactive groups.[74] Among them, the Cu^{I}-catalysed version of the Huisgen 1,3-dipolar cycloaddition reaction has been used extensively in many fields including drug discovery, bioconjugation and

materials science.[75-80] The reaction between the alkyne and azide is regioselective and produces 1,4-triazole. It is known as the copper-catalysed azide-alkyne cycloaddition (CuAAC) reaction (Figure 1.17.A).[75] The Matyjaszewski group demonstrated that azides with electron-withdrawing substituents and less steric congestion, or alkynes with carbonyl groups, have higher reactivity.[81]

Without the use of a catalyst, the Huisgen cycloaddition reaction leads to a mixture of triazole 1,4- and 1,5-regioisomers (Figure 1.17.B) and is around eight orders of magnitude slower.[75] Instead of using a copper catalyst, Fokin and Jia reported that a 1,5-disubstituted triazole is selectively formed from azides and terminal alkynes under catalysis by ruthenium cyclopentadienyl complexes (RuAAC reaction, Figure 1.17.C).[82, 83] This reaction has been used to modify nucleic acids.[84, 85]

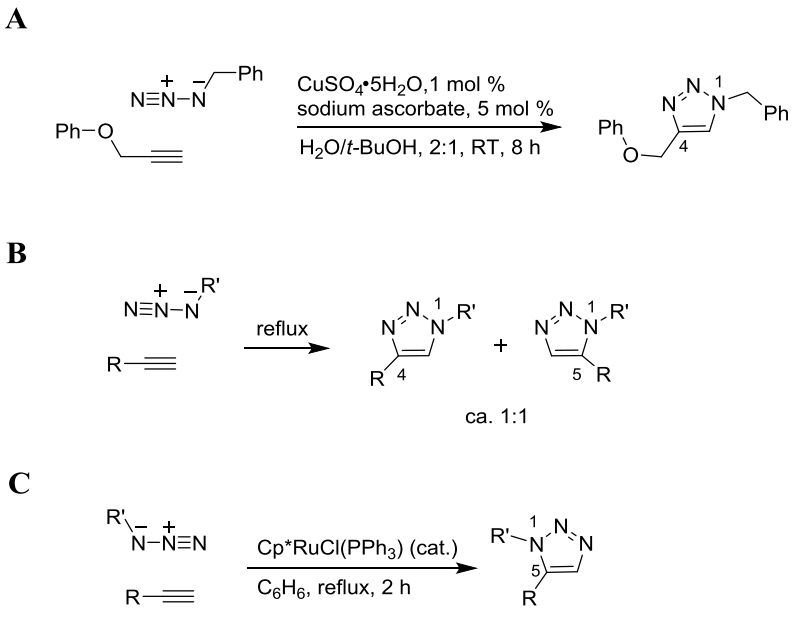


Figure 1.17. **A.** The copper-catalysed azide-alkyne cycloaddition reaction.[75] **B.** The uncatalysed azide-alkyne cycloaddition reaction.[75] **C.** The ruthenium-catalysed azide-alkyne cycloaddition.[82]

Several years after the first report of the CuAAC reaction, Fokin and co-workers published details on the mechanism (Figure 1.18).[86] There are two copper ions involved

in the catalytic cycle. One copper is linked to the acetylide by a σ-bond, improving the nucleophilicity of the intermediate; the second copper forms a weak π-complex with the alkyne group to afford intermediate **II**, which then coordinates with an azide group to form intermediate **III**. After the nucleophilic attack and ring closure, intermediate **V** is formed *via* the intermediate **IV**. In the last step, the copper catalyst leaves intermediate **V** and the final 1,2,3-triazole product **VI** is produced. The copper catalyst then proceeds to the next cycle.

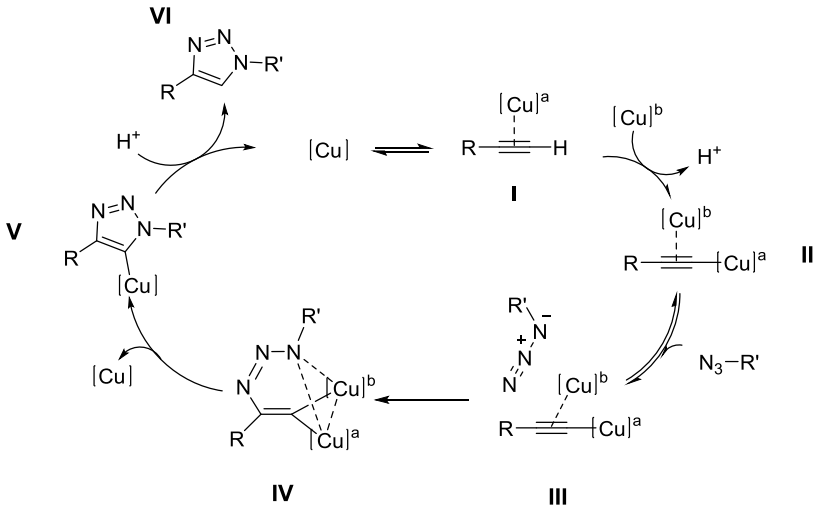


Figure 1.18. Mechanism of the CuAAC reaction.[86]

1.4.2.1 Nucleic acid functionalisation *via* the CuAAC labelling

The CuAAC reaction has been widely used in nucleic acid research.[53, 69] The azide and alkyne groups individually are inert to nucleic acid functional groups and react with each other very efficiently in aqueous media in the presence of a copper catalyst. Here, the application of this reaction in the fluorescent labelling of nucleic acids is described in detail. The CuAAC reaction provides a flexible way to prepare various fluorescent DNA and RNA. Alkyne/azide-modified nucleic acids can be labelled with azido/alkynyl fluorophores.

1.4.2.1.1 Labelling of alkyne-modified nucleic acids

Alkyne modifications have been introduced into oligonucleotides *via* solid phase phosphoramidite synthesis, using procedures developed by the Carell and Seela groups amongst others (Figure 1.19).[87-92] The base and sugar modifications usually carry long linkers to reduce steric hindrance in the labelling reaction (Figure 1.19.A).[88, 89] Carell and co-workers reported that by introducing different protecting groups on the alkyne modification, single to triple labelling with different reporter groups can be achieved efficiently (click-click-click).[91] By choosing a bi-functional group, two modifications were introduced at the same time by the Seela group (Figure 1.19.B).[93] The alkyne groups were also attached to the 5′-end or 3′-end of oligonucleotides without inhibiting DNA binding affinity (Figure 1.19.C).[94, 95] All these alkyne functionalised DNA/RNA strands can be subsequently labelled with different fluorescent or other reporter groups.

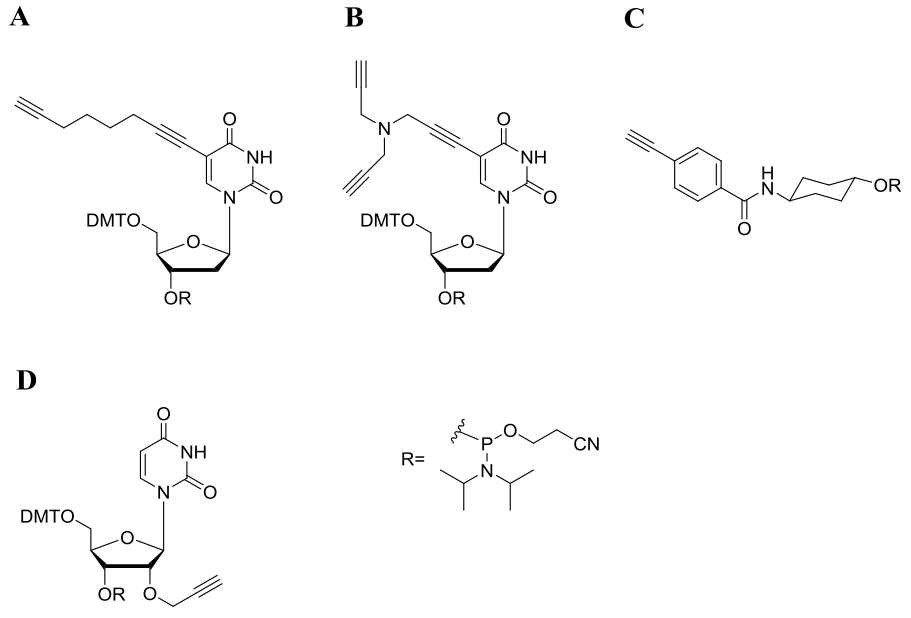


Figure 1.19. Alkyne-modified phosphoramidites used in solid phase oligonucleotide synthesis. **A.** Octadiynyl dU;[88, 89] **B.** Tripropargylamino dU;[93] **C.** Monomer for terminal alkyne modification;[94] **D.** 2′-Propargyl uridine.[96]

Alkyne-bearing nucleoside triphosphates have also been enzymatically incorporated into DNA and RNA using different polymerases (Figure 1.20).[97-103] Carell and co-workers attached alkyne groups to the 5-position of dUTPs. These modified triphosphates demonstrated good incorporation compatibility with different DNA polymerases and the resultant DNA was efficiently labelled with various reporter groups *via* the CuAAC reaction (Figure 1.20.A).[100-102, 104] The Berdis group prepared a non-natural nucleotide, which was selectively incorporated opposite an abasic site, and used this as a probe for non-instructional lesion monitoring (Figure 1.20.B).[98, 105] Site-specific enzymatic incorporation has also been achieved by using the third base-pairing system devised by the Hirao group (Figure 1.20.C).[103] X-ray studies of rigid modified deoxycytidine triphosphates (dCTPs) and KlenTaq polymerase complexes demonstrated that the triphosphate binding site of the polymerase is flexible enough to enable primer elongation when using modified triphosphates (Figure 1.20.D).[106]

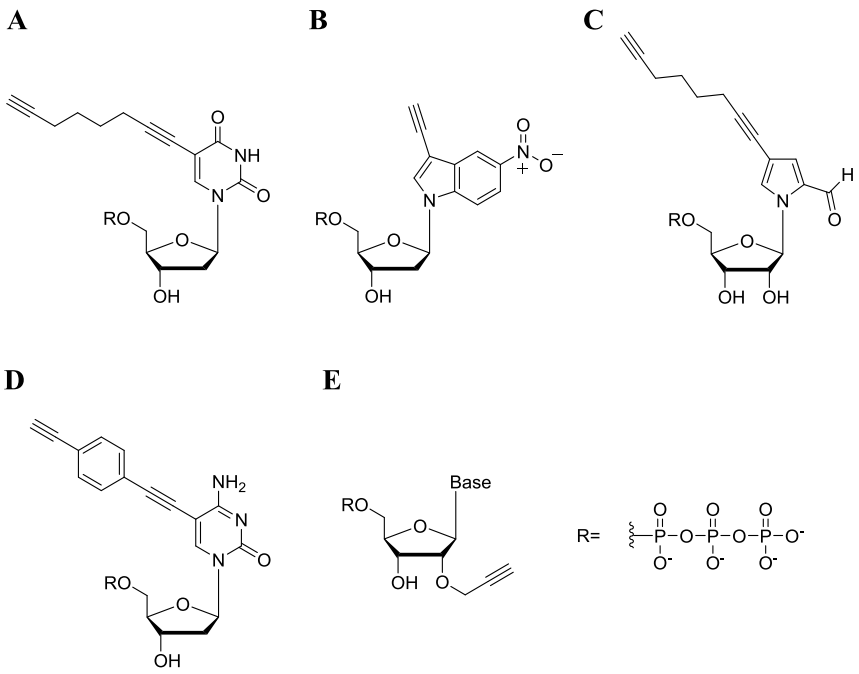


Figure 1.20. Modified nucleoside triphosphates used in enzymatic incorporation reactions. **A.** Octadiynyl dUTP;[100] **B.** Non-natural nucleoside triphosphate;[98, 105] **C.** Nucleotide with a pyrrole base analogue;[103] **D.** Rigid and aromatic modified dCTP;[106] **E.** 2′-Propargyl nucleoside triphosphate.[99]

1.4.2.1.2 Labelling of azide-modified nucleic acids

The introduction of azide modifications into DNA and RNA generally requires enzymatic incorporation, because of the reactivity of azides to phosphorus(III) during the solid phase synthesis (the Staudinger reaction).[107] Several azido nucleotides have been synthesised using H-phosphonate or 2-chlorophenyl phosphodiester derivatives in solid phase synthesis to avoid this problem (Figure 1.21.A),[108, 109] but this method deviates from the commonly accepted best strategy for oligonucleotide synthesis (the phosphoramidite method). In 2007 the Carell group reported the synthesis of an azide modified triphosphate with an alkyl linkage (Figure 1.21.B).[101] However, this modified triphosphate was found not to be a good substrate for DNA polymerases due to the flexible linkage. The Srivatsan group used a similar alkyl linker for azido uridine triphosphates (UTPs) and demonstrated the production of efficient azido-modified RNA transcripts (Figure 1.21.C).[110] Marx and co-workers prepared an azido triphosphate with an alkynyl linkage and avoided the internal cycloaddition problem resulting from the CuAAC reaction by labelling the azide-modified primer extension products *via* the Staudinger reaction (Figure 1.21.D).[111] However, the reaction is sensitive to oxygen, limiting its applications. A range of sugar and base modified azido triphosphates (Figure 1.21.E) were attached to the termini of RNA and DNA strands by the Jäschke and Kjems groups using different nucleotidyl transferases.[112-114] However, this strategy can only introduce one or multiple modifications to the 3′-end of DNA strands, and cannot be used for general internal sequence modification. Overall, there is a need for a new type of azide triphosphate that could be efficiently incorporated using enzymes and labelled with fluorescent tags.

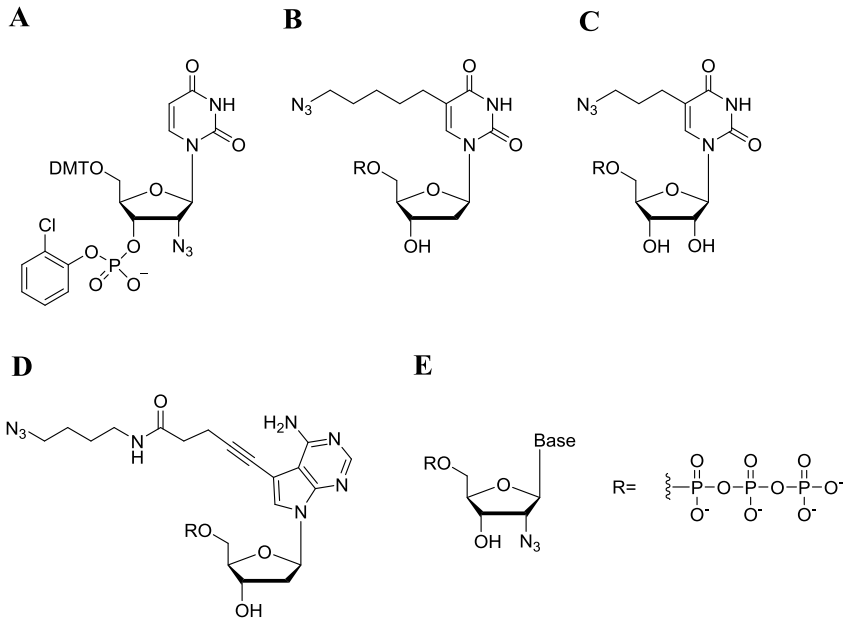


Figure 1.21. **A.** Monomer used in solid phase triester DNA synthesis.[108] **B** to **E.** Monomers used for enzymatic azide incorporation. **B.** Azido dUTP;[101] **C.** Azido UTP;[110] **D.** Azido deazadeoxyadenine triphosphate;[111] **E.** 2′-Azido nucleoside triphosphate.[112]

The CuAAC reaction has also been utilised to make base analogues, to cross-link DNA strands and to make DNA nanostructures.[53, 69, 80] Although it has found widespread use in the nucleic acid field, it has limitations for *in vivo* applications due to the cytotoxicity of copper and the requirement for anaerobic reaction conditions.

1.4.3 Strain-promoted azide-alkyne cycloaddition reaction

To overcome the limitations of CuAAC, the strain-promoted azide-alkyne cycloaddition (SPAAC) reaction was introduced by Bertozzi and co-workers in 2004.[115] This reaction involves the uncatalysed reaction between an azide and a ring strained alkyne (Figure 1.22). The 8-membered ring contains a triple bond, which is more reactive than linear alkynes due to the distortion of its sp-orbitals. Due to ring strain, it adopts a geometry between that of an sp and sp^3 hybridised carbon atom. The reaction to form an sp^2 hybridised triazole product greatly reduces ring strain. Two regioisomeric products are

formed in the SPAAC reaction, similar to the uncatalysed azide and alkyne cycloaddition reaction. The most important advantage of the SPAAC reaction is that it can be used for both *in vitro* and *in vivo* application due to the absence of toxic metal catalysts compared to the CuAAC reaction.[116-120]

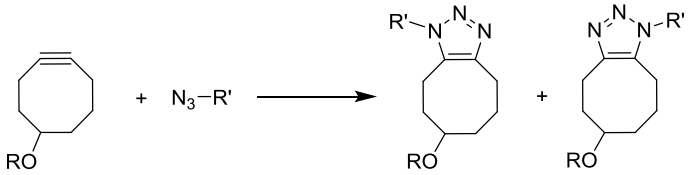


Figure 1.22. The strain-promoted azide-alkyne cycloaddition reaction.

Various analogues of cyclooctyne have been synthesised to increase its reactivity.[116, 117] Dibenzocyclooctyne (DIBO, Figure 1.23.A), carrying two aromatic rings fused onto the cyclooctyne, was synthesised by the Boons group.[121] Its reactivity with azides is increased due to the electron withdrawing effect and the extra ring strain from the aromatic rings. The aqueous solubility of DIBO was, however, lower than the unsubstituted cyclooctyne, which is not ideal for nucleic acid applications. In 2010 an equally reactive alkyne bicyclo[6.1.0]nonyne (BCN, *endo* form, Figure 1.23.B) with a 3-membered ring fused onto the cyclooctyne ring was reported by van Delft *et al.*[122] Furthermore this molecule is symmetric and does not possess any chiral centres, simplifying the stereochemistry of triazole products formed from it. For applications in which diastereoisomers of the SPAAC reaction products might be undesirable, achiral BCN is the preferred alkyne of choice.[122, 123] The second-order rate constant between BCN and benzyl azide is 0.29 $M^{-1}s^{-1}$ in acetonitrile/water (1:2) solution, *i.e.* the reaction is faster than that of the plain cyclooctyne ($\approx 2 \times 10^{-3}$ $M^{-1}s^{-1}$).[122]

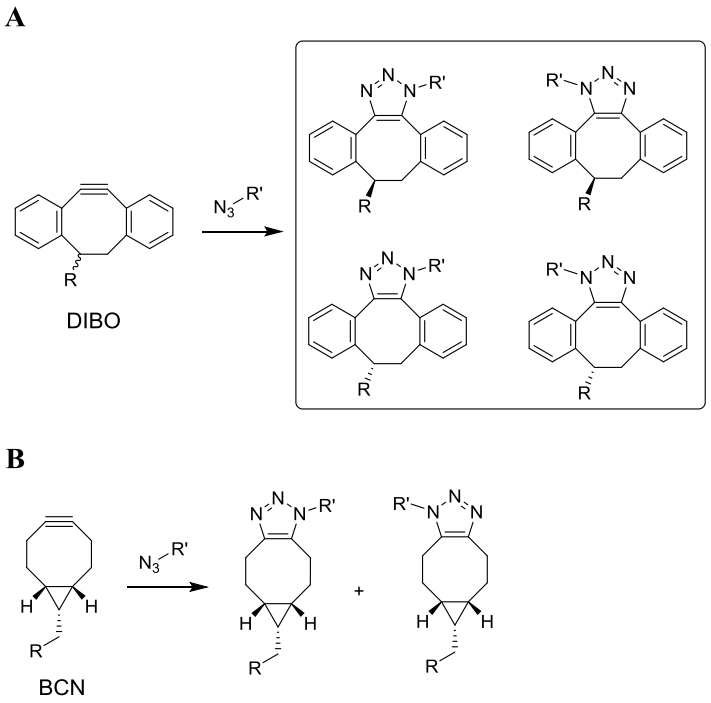


Figure 1.23. **A.** Reaction between DIBO and azide *via* the SPAAC reaction. **B.** Reaction between BCN and azide *via* the SPAAC reaction.

1.4.3.1 Labelling of cyclooctyne-modified nucleic acids

In 2010, the SPAAC reaction was first used in oligonucleotide labelling and copper-free DNA strand ligation.[124, 125] DIBO and BCN cyclooctyne phosphoramidite monomers have been synthesised for terminal labelling of DNA and RNA (Figure 1.24.A and B).[123, 124, 126-128] In the Brown group, DIBO and BCN were incorporated into oligonucleotides by attachment to the 5-position of the uracil base or the 2′-position of the pentose sugar (Figure 1.24.C).[129] The above cyclooctyne modified DNA strands were further labelled with fluorescent tags, conjugated to proteins or ligated to DNA strands *via* the SPAAC reaction.

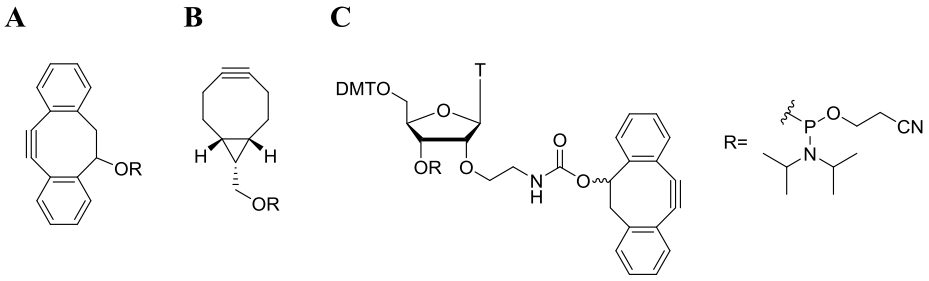


Figure 1.24. Chemical structures of cyclooctyne phosphoramidites used in DNA functionalisation. **A** and **B.** Cyclooctyne phosphoramidites for terminal labelling;[123, 126] **C.** 2′-Sugar modified phosphoramidite.[129]

The DIBO group (Figure 1.24.A) was attached to the 5′-end of oligonucleotides during solid-phase synthesis by the Taton group.[126] The functionalised product was used as a modified primer for PCR reactions. The DIBO modification was found to be stable to PCR thermal cycling and was subsequently labelled with azide-bearing Texas Red to afford fluorescent PCR products. However, other enzymatic studies of cyclooctyne-modified triphosphates are limited. Therefore, it would be useful to investigate their biocompatibility for enzymatic nucleic acid synthesis and subsequent functionalisation.

It should be noted that cyclooctynes can react with nucleophiles such as thiols which are prevalent in cells, thereby sequestering them. For protein labelling, Boelens and co-workers partially solved this problem by first introducing iodoacetamide, which alkylates peptidylcysteine thiols, and performing the SPAAC labelling.[130] For other *in vivo* biomolecule-labelling applications, alternative copper-free and efficient reactions are required to overcome this problem.

1.4.4 Inverse electron demand Diels-Alder reaction

An inverse electron demand Diels-Alder (IEDDA) reaction between tetrazine and olefins is another fast metal free click reaction developed by Fox and co-workers (Figure 1.25.A).[131] This reaction has been widely used in biomolecule conjugation and functionalisation.[132-134] The reaction mechanism is shown in Figure 1.25.B.[131, 135] The first cycloaddition step between the diene tetrazine and dienophile olefin is the rate-limiting step. The intermediate **II** is formed after the elimination of nitrogen from the highly strained cyclic intermediate **I**. Further isomerisation intermediate **II** affords the final product **III**. Cyclised dienes have higher reactivity than the linear version due to ring distortion in the transition state of intermediate **I**. The *trans*-cyclooctene-tetrazine reaction is one of the fastest known bioorthogonal click reactions with a second-order rate constant up to 30 000 $M^{-1}s^{-1}$ in phosphate buffered saline (PBS) at 37 °C.[136] This is several orders of magnitude faster than the SPAAC reaction (which in turn is much faster than the CuAAC reaction).[135]

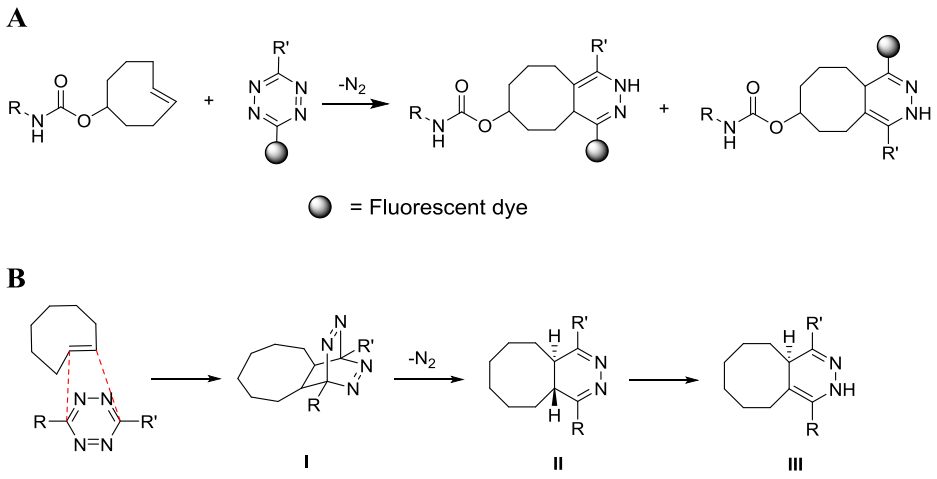


Figure 1.25. **A.** General reaction scheme of *trans*-cyclooctene and a tetrazine-derivatised fluorescent dye. **B.** Reaction mechanism of inverse electron demand Diels-Alder reaction.[131]

1.4.4.1 Labelling of *trans*-cyclooctene-modified nucleic acids

The tetrazine group is very reactive and bulky, therefore the dienophiles (norbornenes and *trans*-cyclooctenes) are usually attached to the oligonucleotides. They have been attached to the terminal and internal positions of oligonucleotides using their phosphoramidites (Figure 1.26).[132, 133, 137, 138] The Jäschke group has pioneered the use of IEDDA on nucleic acids. Norbornene and *trans*-cyclooctene functionalised DNA and RNA were efficiently labelled with tetrazine reporter groups.[133, 137, 138] Cyclopropene and tetrazine were also introduced to the terminal position of oligonucleotides by the Devaraj group through coupling their NHS active esters to the oligonucleotides containing amino groups.[132] When the tetrazine is located close to a fluorophore, it behaves as a quencher. This effect was alleviated after tetrazine reacts with *trans*-cyclooctene affording an intense fluorescence signal.

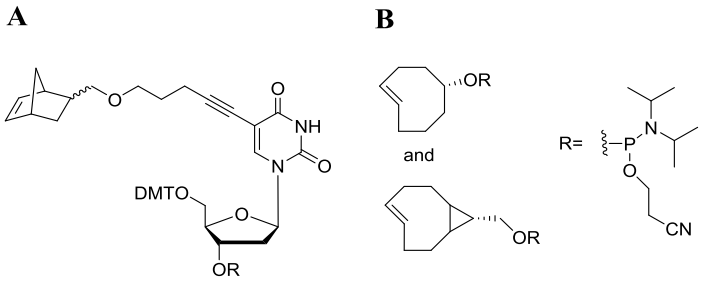


Figure 1.26. Cyclo-dienophile modified phosphoramidites used in DNA functionalisation. **A.** Monomer used for internal modification of DNA;[133] **B.** Monomers used for terminal modification.[138]

The *trans*-cyclooctene and norbornene modified nucleic acids have also been prepared from their modified triphosphates using primer extension and PCR (Figure 1.27.A and B).[139, 140] The Royzen group has reported the enzymatic incorporation of a *trans*-cyclooctene modified cytidine triphosphate (Figure 1.27.C).[134] In the Jäschke group, various cyclo-dienophile-modified RNA monophosphates were successfully attached to

the termini of RNA strands using RNA polymerases (Figure 1.27.D).[141] These modified products were further conjugated to the tetrazine-modified peptides or fluorophores.

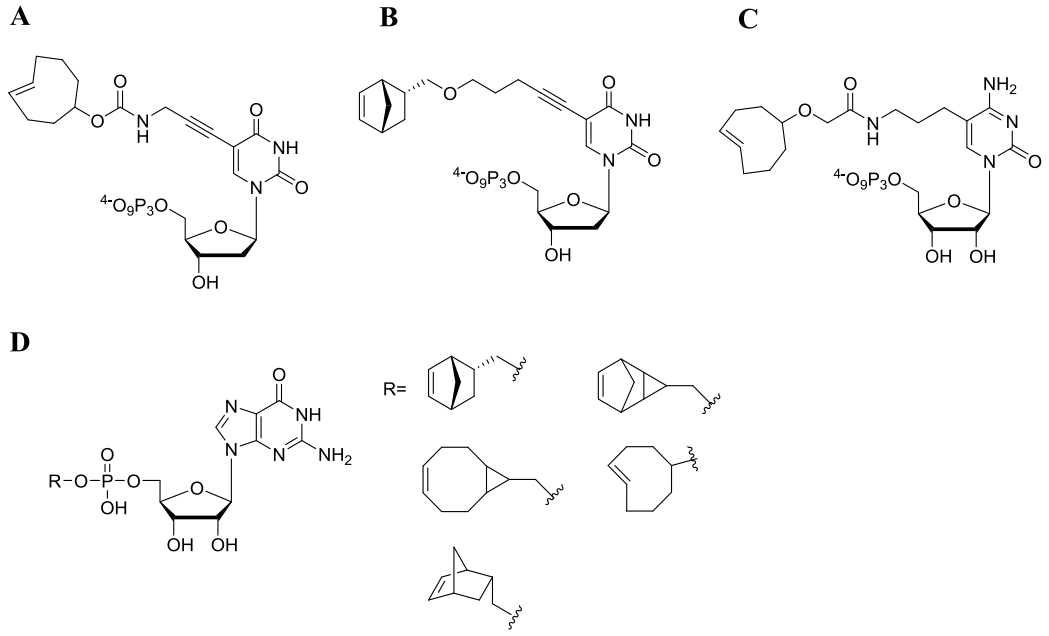


Figure 1.27. Cyclo-dienophile-modified triphosphates (**A** to **C**)[134, 139, 140] and protected monophosphate (**D**)[141] used in DNA and RNA functionalisation.

1.4.5 Dual/multiple fluorescent labelling

Fluorescent oligonucleotides carrying different combinations of fluorophores have wide applications in the detection of nucleic acid, and usually utilise FRET in the detection method. Direct incorporation of modified phosphoramidites during oligonucleotide synthesis or nucleoside triphosphates during enzymatic synthesis requires the laborious and time-consuming synthesis of different monomers/nucleotides. In comparison, modular post-labelling strategies are potentially simpler. As discussed in Section 1.4.2, the Carell group demonstrated a multiple labelling strategy by using different alkyne-protecting groups within a DNA strand which can be selectively deprotected and progressively labelled.[91] To avoid these selective deprotection strategies, which could be

difficult to optimise and are not truly orthogonal, the CuAAC and IEDDA reactions have been combined by the Jäschke group to label oligonucleotides site-specifically in a single-tube reaction (Figure 1.28).[138] The linear alkynes do not cross react with tetrazine, which makes them orthogonal to the IEDDA reaction. Houk *et al.* reported that IEDDA can also be used in combination with SPAAC to conduct multiple orthogonal labelling reactions.[142] Hilderbrand and co-workers applied this orthogonal pair to develop dual-coloured cell-surface imaging.[143] Among the cyclooctyne modifications mentioned in Section 1.4.3, BCN can react both with azides and tetrazines, while DIBO can only react with azides *via* the SPAAC reaction. In addition to above labelling methods, there are also other compatible dual labelling combinations that have been studied for the functionalisation of DNA.[102]

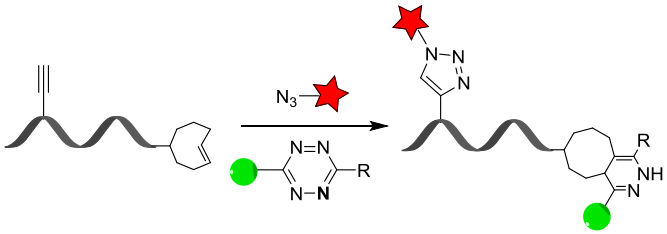


Figure 1.28. Orthogonal labelling reaction using the combination of CuAAC and IEDDA.[138]

1.5 Fluorescent cellular DNA imaging using modified nucleosides

With the development of fluorescence microscopy, fluorophores have been widely used to monitor biological processes.[36, 144] Carbohydrates, proteins and nucleic acids are key elements in living organisms.[145] Fluorescent DNA imaging in cells has been used as an important tool to monitor the DNA replication, transcription and other metabolic processes.[146] This section focuses on cellular DNA imaging using modified nucleosides.

Nucleotides do not readily enter cells due to their anionic nature. However, free nucleosides can be taken up by cells and converted to monophosphates, diphosphates and triphosphates by cellular kinases (Figure 1.29).[147-149] The first and rate-limiting step is the conversion of the nucleoside into a monophosphate.[150] The final reactive triphosphates can then be incorporated to the DNA or RNA strands during replication and transcription. Therefore, in principle, modified nucleosides can be utilised to functionalise nucleic acids *in vivo*.

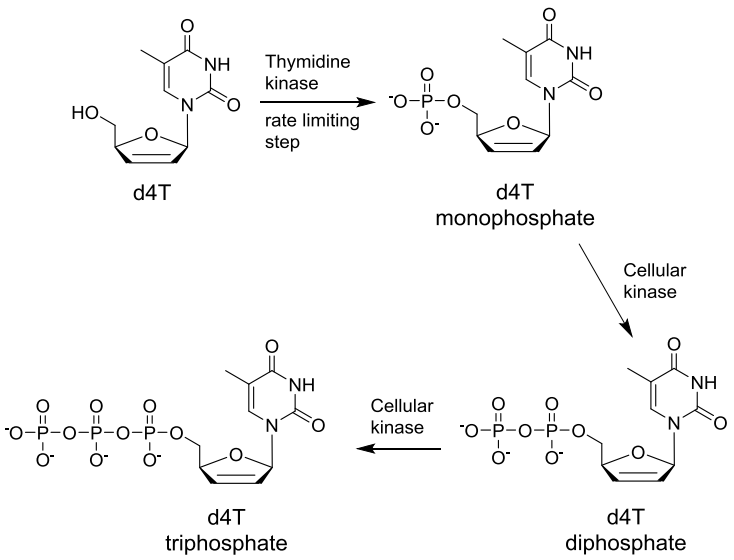


Figure 1.29. The transformation of 2′,3′-dideoxy-2′,3′-didehydrothymidine (d4T) into its active triphosphate.[147, 151]

Gratzner *et al.* monitored nascent DNA in proliferating cells by treatment with 5-bromo-2′-deoxyuridine (BrdU, Figure 1.30.A) and subsequent staining with fluorescently-labelled antibodies specific for BrdU.[152] However, the antibody has poor cell membrane permeability due to its large size. The Salic and Greenberg groups found that 5-ethynyl-2′-deoxyuridine (EdU, Figure 1.30.B) and 5-ethynyl-2′-deoxycytidine (EdC, Figure 1.30.C) were converted to their triphosphates by cellular kinases and incorporated into genomic DNA by endogenous polymerases.[38, 97] Subsequent

cell-fixation and derivatisation of the alkyne-modified DNA with a fluorescent azide *via* the CuAAC reaction yielded DNA which can be visualised by fluorescence microscopy. The Salic group also adapted this approach for the detection of RNA synthesis in cells by using the ribonucleoside 5-ethynyluridine instead.[153] The methodology is compatible with fixed cells, but the toxicity of copper and EdU precludes its application in live cells. The addition of EdU to live cells causes DNA damage, cell cycle arrest, and apoptosis.[154] More recently, the less metabolically toxic (2′S)-2′-deoxy-2′-fluoro-5-ethynyluridine (F-*ara*-EdU, Figure 1.30.D) was used by Luedtke and Neef for cellular imaging.[155]

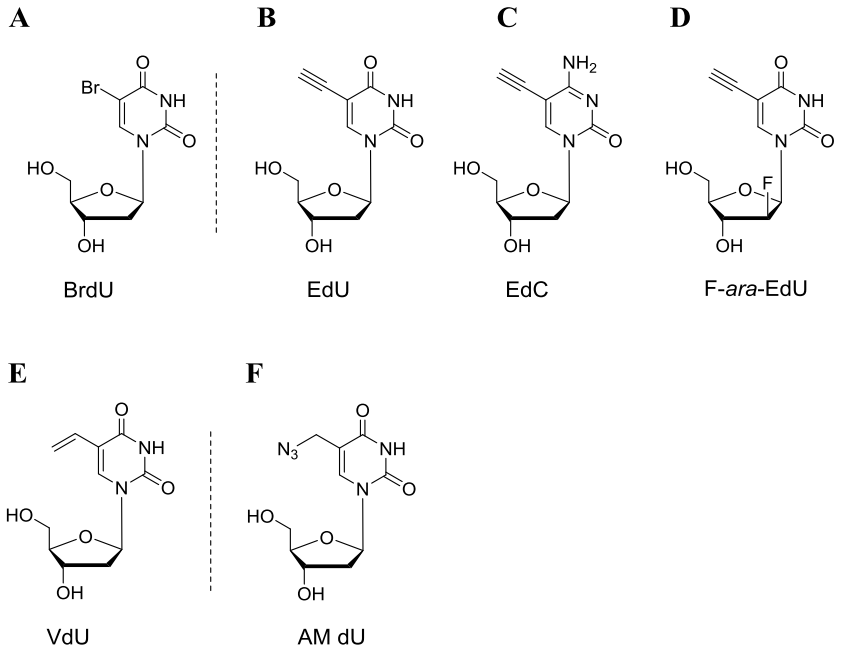


Figure 1.30. Modified nucleosides for cell imaging. **A.** BrdU used in antibody staining.[152] **B** to **D.** monomers used in the CuAAC labelling. **B.** EdU;[38] **C.** EdC;[97] **D.** F-*ara*-EdU.[155] **E.** VdU used in the IEDDA labelling.[156] **F.** AM dU used in the SPAAC labelling.[157]

Moreover, efficient copper-free chemical methodologies have been developed to overcome the problem of CuAAC staining for *in cellulo* applications. In the Luedtke group, 5-vinyl-2′-deoxyuridine (VdU, Figure 1.30.E) was used to label cells with tetrazine fluorophores *via* the IEDDA reaction.[156] VdU had less cytotoxicity and higher

stability than EdU. It was found to be compatible with BrdU and F-*ara*-EdU labelling, which have been used to label cellular DNA sequentially. In the same year, this research group also demonstrated that toxic metal-ion based catalysis can be negated if 5-azidomethyl-2′-deoxyuridine (AM dU, Figure 1.30.F) is incorporated into genomic DNA.[157] Subsequent treatment with a cyclooctyne (BCN) dye derivative yielded the fluorescently labelled DNA *via* the SPAAC reaction.

Unlike bulky dye-labelled nucleosides, the modified nucleosides described above are sufficiently similar to their natural analogues to be converted to triphosphates by endogenous cellular enzymes. As mentioned previously, this conversion of nucleoside to triphosphate is essential, since triphosphates themselves are unable to pass the cell membrane (and enter cells) due to their negative charge.

1.6 Objectives of the research

Investigating new technologies for use in DNA detection for biological applications is the main objective of this DPhil project. Enzymatic synthesis of modified DNA is an efficient strategy to introduce different numbers of functional groups. The employment of copper-free click reactions to densely label the modified DNA is more efficient than the direct enzymatic incorporation of fluorescent nucleoside triphosphates, providing a simple way to increase the signal intensity of fluorescent probes compared to the single-fluorophore labelled probes. The aims of the work in Chapter 2 to 4 are to investigate the enzymatic incorporation compatibility of various (cyclooctyne, *trans*-cyclooctene and azide) modified triphosphates and the feasibility of preparing fluorescent probes by labelling the functionalised enzymatic products using the SPAAC and IEDDA reactions.

Monitoring nucleic acids in cells is very important for biological and pharmacological studies. The aim of the work described in Chapter 5 is to expand the tool-box for cellular DNA imaging. Synthesis of different modified nucleosides and their applications in cellular DNA imaging are investigated.

Chapter 2 Synthesis and applications of cyclooctyne-modified deoxyuridine triphosphates

2 Synthesis and applications of cyclooctyne-modified deoxyuridine triphosphates

2.1 Background and aims

Nucleoside 5′-triphosphates are utilised in many important biological reactions, including DNA replication, reverse transcription and RNA transcription.[158] They are potentially useful tools for introducing densely packed chemical modifications, such as fluorophores, into DNA and RNA.[59, 101, 159-161] Oligonucleotides labelled with multiple fluorophores give rise to far stronger signal intensity than conventional single fluorophore probes.[73] This is of particular importance for the fluorescent detection of DNA which are present at very low concentrations (*e.g.* specific loci of genomic DNA in fixed cells). Fluorescently modified nucleotides have been incorporated directly into DNA using enzymatic methods such as PCR (Chapter 1, Section 1.3.2).[67, 162] However, a high ratio of the natural unmodified dNTP to the fluorescently modified triphosphate (~9:1) must be used to achieve efficient enzymatic DNA synthesis.[71] This limits the number of fluorophores (labelling density) that can be introduced. Alternatively, DNA or RNA containing reactive functional groups can be post-synthetically labelled with fluorophores (Chapter 1, Section 1.4).[39, 68, 69] These reactive groups can be introduced *via* nucleotides that can subsequently be labelled using the appropriate chemical reaction. A variety of different functional groups can be introduced into DNA.

In this context, the strain-promoted azide-alkyne cycloaddition reaction (SPAAC, Section 1.4.3) would be of great use since it does not require a toxic metal catalyst (unlike the copper-catalysed azide-alkyne cycloaddition reaction) and it is an efficient click chemistry reaction.[115] Although different modified cyclooctynes have been introduced

into DNA strands *via* solid phase chemical synthesis,[68] there is only one example of enzymatic incorporation of cyclooctyne nucleotides. In 2015 Jäschke *et al.* described an aza-dibenzocyclooctyne modified deoxyuridine triphosphate (dUTP) and its successful 3′-terminal incorporation to DNA using nucelotidyl transferase enzymes.[114]

The aim of the work in this chapter is to investigate the feasibility of enzymatically incorporating reactive cyclooctyne derivatives into DNA and to examine their use in the synthesis of fluorescent probes (Figure 2.1). To this end, dibenzocyclooctyne (DIBO) and bicyclo[6.1.0]nonyne (BCN) modified triphosphates are explored in detail.

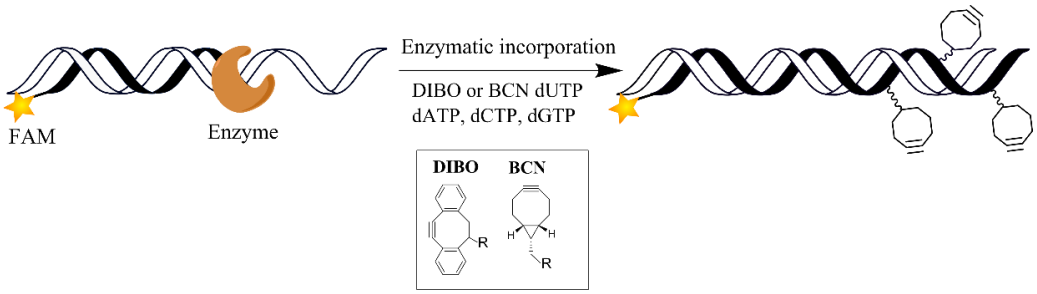


Figure 2.1. Enzymatic (polymerase) DNA synthesis using cyclooctyne-modified triphosphates.

2.2 Preparation of cyclooctyne-modified dUTPs

Nucleoside triphosphates are commonly prepared from the corresponding nucleosides using the Yoshikawa or the Ludwig-Eckstein method (Figure 2.2).[158, 163, 164] In the Yoshikawa procedure (Figure 2.2.A), phosphorus oxychloride ($POCl_3$) is employed as the phosphorylating agent using trimethylphosphate as solvent to generate the 5′-dichlorophosphate (an activated monophosphate, intermediate **II**).[163] The reaction mixture is subsequently treated with pyrophosphate and quenched with water to afford the nucleoside 5′-triphosphate. This methodology is simple, and the employment of trimethylphosphate as solvent favours the formation of the 5′-triphosphate. However, the

reaction generates many by-products including 3′-phosphate and 3′,5′-diphosphate, which are very difficult to remove. For the Ludwig-Eckstein reaction (Figure 2.2.B), salicyl chlorophosphite, which is less reactive than phosphorus oxychloride and more selective, reacts with the 5′-OH to generate the intermediate **V**.[164] After treatment with pyrophosphate and iodine solution (to oxidise the P^{III} to P^{V}), the desired nucleoside 5′-triphosphate is obtained. Compared to the Yoshikawa procedure, this method is more regioselective but requires an additional oxidation step.

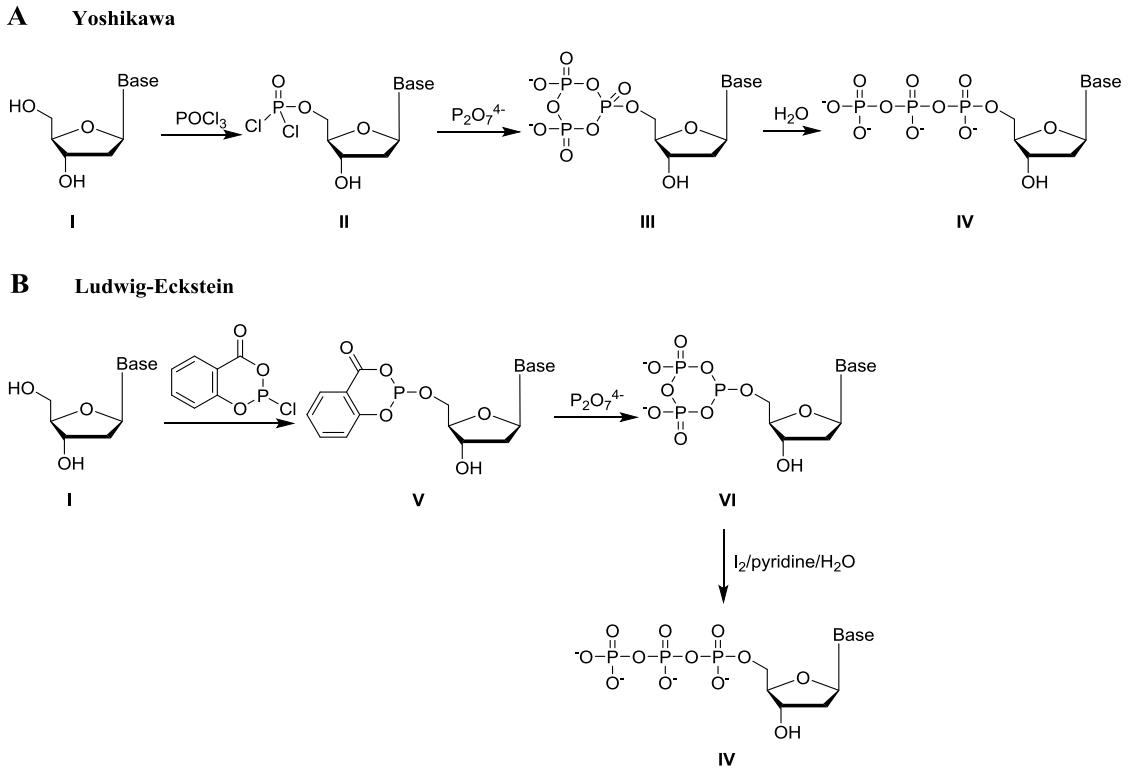


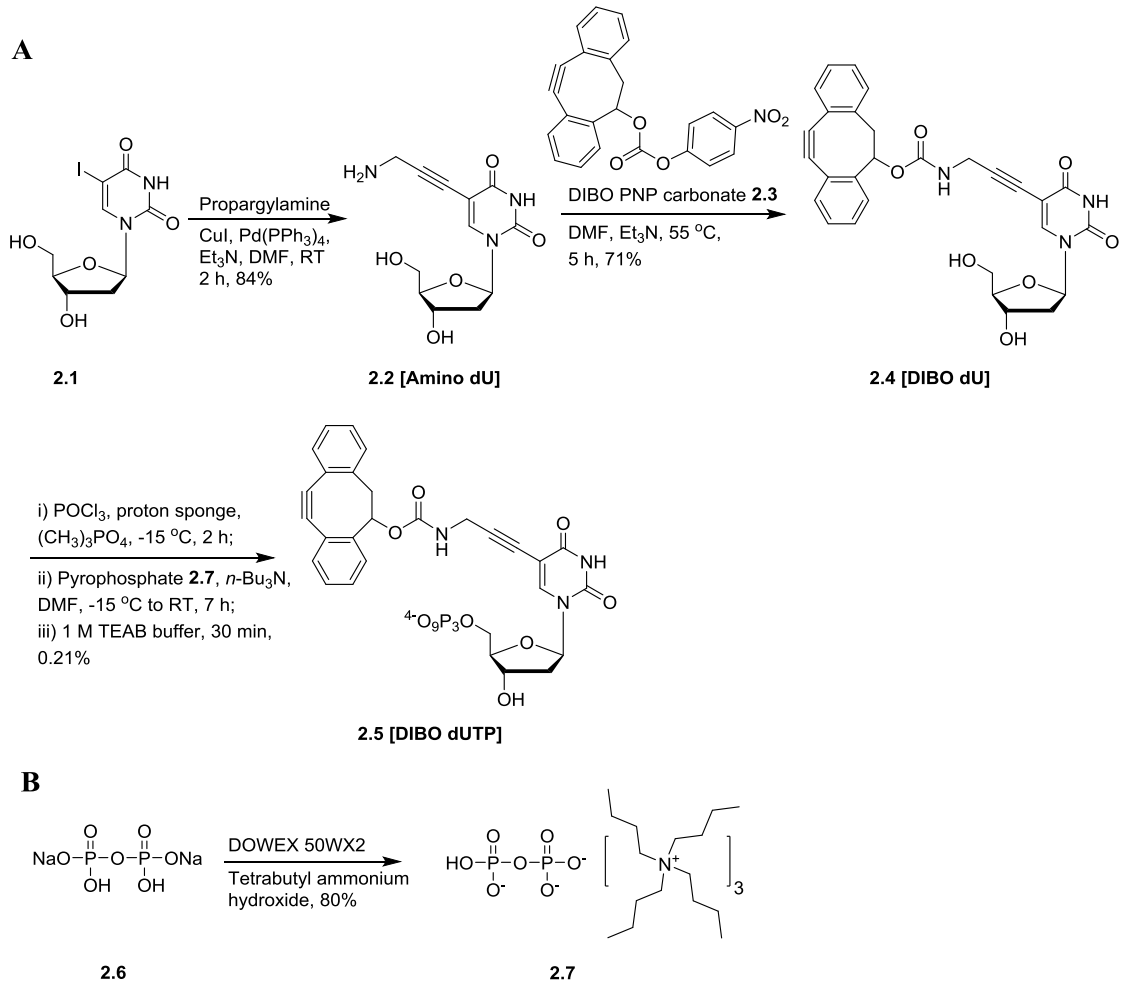
Figure 2.2. Synthesis of deoxynucleoside triphosphates by the Yoshikawa (**A**) or the Ludwig-Eckstein method (**B**).[158, 164]

2′-Deoxyuridine is one of the easiest natural nucleosides to chemically modify and has been extensively used to introduce functional groups into nucleic acids for various applications (*e.g.* fluorescent detection, antiviral drug development).[38, 84, 160, 165-167] Importantly modifications attached at the 5-position of the uracil base protrude into the

major groove of the DNA double helix and are remote from the Watson-Crick base-pairing hydrogen bonding groups. Therefore these modifications cause minimal disruption to the normal B-DNA helical conformation and do not significantly destabilise the duplex.

With this in mind, the DIBO cyclooctyne was joined to the 5-position of uracil base by an aminopropargyl linkage (Scheme 2.1.A). Modified deoxyuridine triphosphates with alkynyl linkages on the 5-position have been incorporated enzymatically by different DNA polymerases.[59, 168] The alkynyl linkage can actually stabilise DNA duplex by enhanced π-stacking, which is useful.[169, 170] Therefore, this linkage has been widely applied to introduce modifications into nucleic acids.[88, 171] The DIBO-modified triphosphate was synthesised as shown in the Scheme 2.1.A. The aminopropargyl group was attached to 5-iodo-2′-deoxyuridine **2.1** by a Sonogashira cross coupling reaction.[172] The DIBO *p*-nitrophenyl (PNP) carbonate **2.3** (synthesised by Dr Montserrat Shelbourne in the Brown group) was then coupled to the amino group of 5-aminopropargyl-2′-deoxyuridine **2.2** (amino dU) to produce the 5-DIBO-2′-deoxyuridine nucleoside **2.4** (DIBO dU).[128] Triphosphorylation of the 5′-OH group of this compound afforded the triphosphate **2.5** using the Yoshikawa reaction as described in Figure 2.2.A.[163] In order to minimise the production of 3′-phosphate and other by-products, the Yoshikawa reaction was carried out at -15 ºC. The 5-DIBO-2′-deoxyuridine-5′-triphosphate (DIBO dUTP) **2.5** was obtained in very low yield; only 0.21% yield was obtained for the triphosphorylation reaction. The tris(tetrabutylammonium) hydrogen pyrophosphate **2.7** used in the Yoshikawa reaction was prepared according to the reported method (Scheme 2.1.B).[173] This compound is very hygroscopic and had to be aliquoted inside a glovebox under inert gas due to the water sensitivity of the Yoshikawa reaction. In 2015, Hodgson and co-workers reported the use

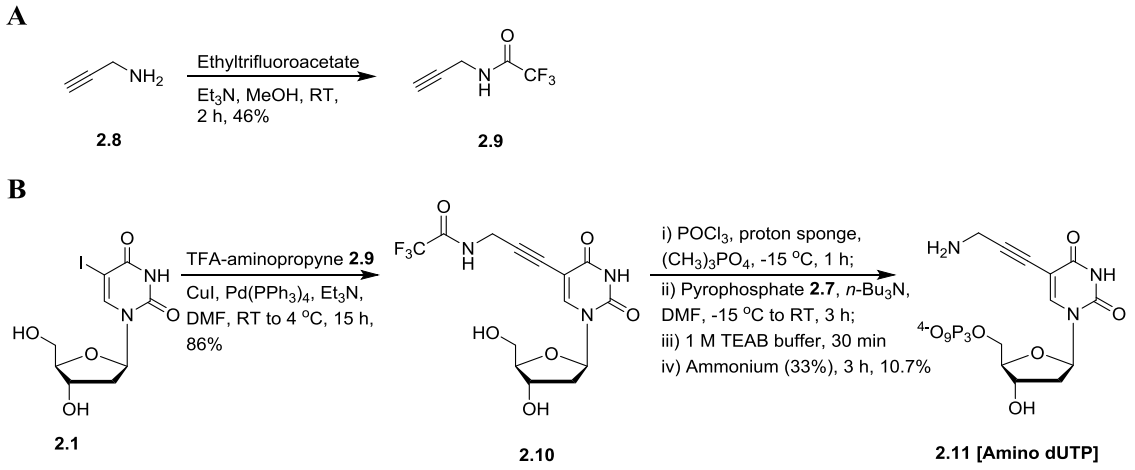
of a tris[bis(triphenylphosphoranylidene)ammonium] pyrophosphate, which is easier to be dry and is much less hygroscopic than the normal tri- and tetra-alkyl ammonium pyrophosphates.[174] It is a good alternative pyrophosphate salt for triphosphate synthesis. This could be used in the future nucleoside triphosphate syntheses.



Scheme 2.1. **A.** Synthesis of DIBO dUTP. **B.** Synthesis of tris(tetrabutylammonium) hydrogen pyrophosphate.

Taking into consideration the very poor yield of DIBO dUTP and the complex steps to prepare DIBO PNP carbonate, an alternative synthetic route was designed (Scheme 2.2).[121, 128] The primary amine of amino dU **2.2** is a suitable attachment point for *N*-hydroxysuccinimide (NHS) derivatives of reporter or other functional groups. Its

triphosphate (aminopropargyl dUTP) is accepted efficiently as a substrate by various DNA polymerase enzymes and has been used for subsequent fluorescent label attachment.[73, 175] However, its moderate reactivity and unwanted hydrolysis of the NHS active ester during labelling reactions limit its use. Therefore, 5-aminopropargyl-2′-deoxyuridine-5′-triphosphate **2.11** (amino dUTP) was synthesised as a common intermediate to introduce the far more reactive cyclooctyne group, enabling the preparation of a broad range of functionalised dUTPs (Scheme 2.2). In this approach, the primary amine must be protected before the Yoshikawa reaction. Trifluoroacetyl (TFA) protected aminopropyne **2.9** (TFA-aminopropyne) was synthesised and then reacted with 5-iodo-2′-deoxyuridine **2.1** *via* the Sonogashira reaction. Compound **2.10** was then converted to its 5′-triphosphate by the Yoshikawa procedure. The TFA protected amino dUTP was deprotected in concentrated aqueous ammonia to afford amino dUTP **2.11** in 9.5% yield for the Yoshikawa and deprotection reactions.



Scheme 2.2. Synthesis of TFA-aminopropyne (**A**) and amino dUTP (**B**).

Three cyclooctyne-modified deoxyuridine triphosphates were prepared by labelling amino dUTP **2.11** individually with an active ester or carbonate of DIBO, BCN-I and BCN-II (Figure 2.3). DIBO dUTP **2.5** and two different BCN dUTP derivatives with

either short or long linkers (BCN-I dUTP **2.12** and BCN-II dUTP **2.13**, respectively) were obtained. These modified triphosphates were synthesised to study the effect of the steric bulk and linkage length on their enzymatic incorporation efficiency, in addition to subsequent cyclooctyne reactivity and potential side reactions.

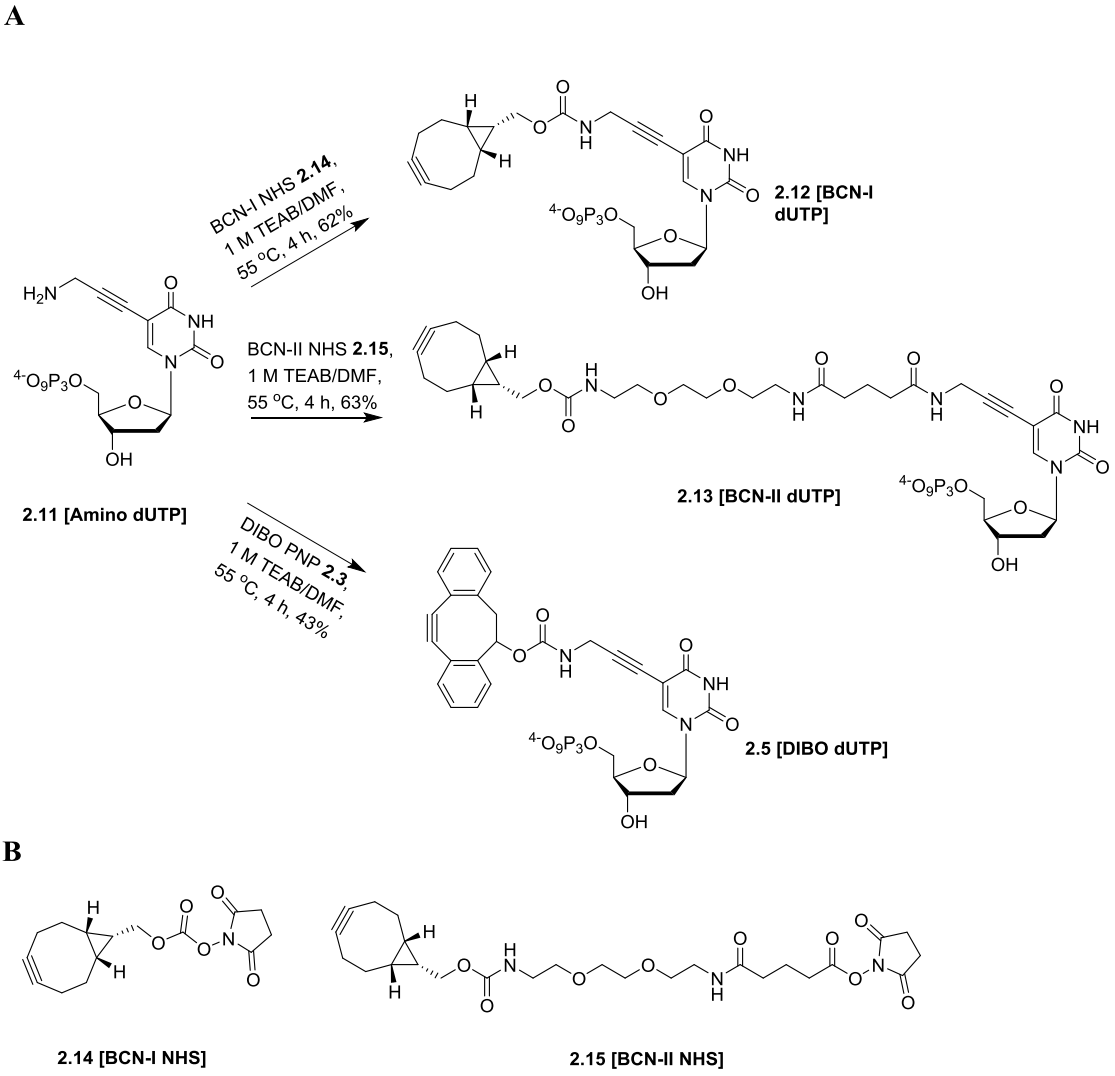


Figure 2.3. **A.** Labelling of amino dUTP to produce cyclooctyne-modified dUTPs. **B.** Chemical structures of BCN-I NHS carbonate and BCN-II NHS ester.

2.3 Primer extension using cyclooctyne dUTPs

Primer extension (Chapter 1, Section 1.3.2) is an efficient biochemical protocol to prepare chemically modified DNA of defined sequence and composition by replacing one of the four canonical nucleotides (dATP, dTTP, dCTP or dGTP) with a modified triphosphate (Figure 2.4). It is a linear copying process that introduces modifications into only one strand of the DNA duplex. Primer extension reactions using DIBO, BCN-I and BCN-II dUTPs were carried out with Therminator™ II, KOD and Gotaq polymerases as well as DNA polymerase I large (Klenow) fragment. The 5′-terminal of the primer was labelled with 6-carboxyfluorescein (FAM, Table 2.1, P3) for ease of visualisation on analytical denaturing polyacrylamide gel electrophoresis (PAGE). Two DNA templates contain the same primer region but have distinct arrangements of adenine bases in the extension region (Table 2.1, T1 and T2). The primers can be extended using only the modified dUTP to give a 19-mer product for template T1 and a 22-mer product (four consecutive incorporation sites) for template T2. Alternatively, primer extension can be performed in the presence of modified dUTP with dATP, dCTP and dGTP to give a 29-mer full-length product for both templates. The natural dNTPs (dATP, dCTP, dGTP, dTTP) and amino dUTP **2.11** were used in positive control experiments for comparison.[73, 175]

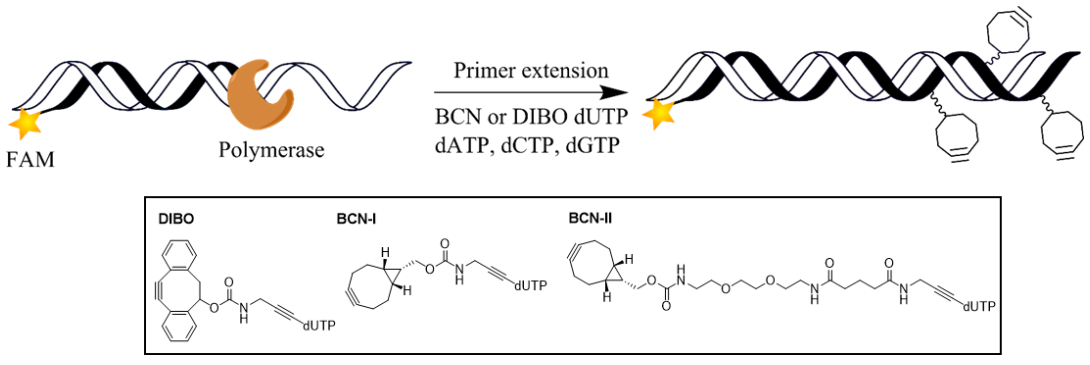


Figure 2.4. Primer extension using cyclooctyne (BCN or DIBO) dUTPs and a DNA template.

Table 2.1. Oligonucleotide sequences used in the primer extension, reverse transcription and PCR.

Code	Sequences (5′-3′)
T1	C**A**GTC**A**CTGT**A**CTGCCGACACACATAACC (DNA template)
T2	C**A**GTC**A**C**AAAA**CTGCCGACACACATAACC (DNA template)
P3	FAM–GGTTATGTGTGTCGGCAG (primer)
T4	C**A**GUC**A**CUGU**A**CUGCCGACACACAUAACC (RNA template)
T5	C**A**GUC**A**C**AAAA**CUGCCGACACACAUAACC (RNA template)
P6	GCATTCGAGCAACGTAAG (PCR primer)
P7	GGTTATGTGTGTCGGCAG (PCR primer)
T8	GG**TTAT**G**T**G**T**G**T**CGGC**A**G**TATT**G**T**C**A**G**T**G**T**G**AATT**CC**A**G**A**G**T**G**T**G**A** G**ATT**G**T**G**T**GC**T**GGCG**AT**C**TTA**CG**TT**GC**T**CG**AAT**GC (PCR template)

FAM is 6-carboxyfluorescein. The bold A (A/T for PCR) shows sites where modified triphosphates are incorporated.

2.3.1 Primer extension using Therminator™ II polymerase

Therminator™ II polymerase is a genetically engineered 9°N^{TM} DNA polymerase, which is more tolerant of modified triphosphates. A wide range of sugar and base modified nucleotides have been incorporated efficiently into nucleic acid by this polymerase.[99, 160, 176] The enzymatic incorporation efficiency of three cyclooctyne modified triphosphates (BCN-I, BCN-II and DBIO dUTPs) was initially studied using Therminator™ II polymerase. To ensure the primer anneals with the template to form a DNA duplex, primer extension was carried out at 60 °C, which is lower than the polymerase's optimum functioning temperature (72 °C). Denaturing PAGE analysis showed that Therminator™ II polymerase gave good incorporation of all cyclooctyne dUTPs using template T1 (Figure 2.5). For the reactions with dTTP or modified dUTPs only, one or two additional nucleotides were incorporated to the 3′-end of partially extended primers (Figure 2.5, lanes 1, 1^{a}, 1^{b}, 1^{c} and 1^{d}). This is presumably caused by the un-templated incorporation, which is a well-known phenomenon.[134, 159, 177] By using

modified dUTP, dATP, dCTP and dGTP, full-length products were observed for all modified triphosphates (Figure 2.5, lanes 2, 2^a, 2^b, 2^c and 2^d). All the copies of the template containing BCN-II dU exhibited slower electrophoretic mobility compared to the natural counterparts. This is due to the extra steric bulk of the BCN-II dU, which carries a longer linker than the BCN-I dU.

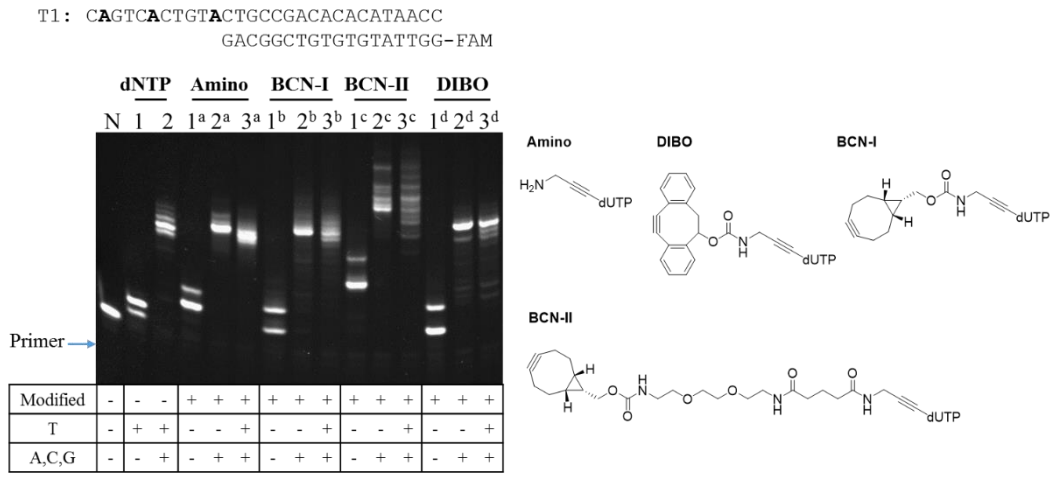


Figure 2.5. Primer extension using modified dUTPs, primer P3 and template T1 for Therminator™ II polymerase (1 unit, 60 °C). 3.2 nmol of each triphosphate was used unless otherwise stated. Lane N, P3 + T1 without triphosphates; lane $1/1^a/1^b/1^c/1^d$, dTTP or modified dUTP; lane $2/2^a/2^b/2^c/2^d$, dTTP or modified dUTP + dATP + dCTP + dGTP; lane $3^a/3^b/3^c/3^d$, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP. (20% denaturing PAGE)

In some applications it is advantageous to reduce the density of reactive groups and ultimately the density of fluorophores introduced *via* the modified dUTP to avoid fluorescence quenching in the labelled DNA. This can be done by using a mixture of modified dUTP and dTTP in the same extension reaction. This will only be successful if the modified dUTP competes with dTTP for incorporation. Therefore primer extension with modified dUTP and dTTP at a 1:1 ratio was explored. The denaturing PAGE analysis showed multiple or broad bands of the fully extended products containing mixed modified dU and dT sites opposite to the dA sites in the template (Figure 2.5, lanes 3^a, 3^b, 3^c and 3^d).

This shows that the amino and the cyclooctyne dUTPs were all able to compete with natural dTTP.

In the case of the demanding template T2, carrying four consecutive adenine bases, amino and BCN dUTPs were incorporated more efficiently than DIBO dUTP by Therminator™ II polymerase. Full-length products were observed on the denaturing PAGE for amino and BCN dUTPs with dATP, dCTP and dGTP reactions (Figure 2.6.A, lanes 2^a, 2^b and 2^d); whereas multiple short bands from the reaction with DIBO dUTP were observed (lane 2^c). To investigate the effect of mixing modified dUTP with dTTP on the primer extension with the template T2, the ratios of modified dUTP to dTTP were varied from 1:1 to 1:0 (Figure 2.6.B). The BCN-II dUTP reactions produced multiple bands in the presence of dTTP, as previously observed with template T1 (Figure 2.5 and Figure 2.6.B). For the DIBO dUTP (1:1) reaction, further weak extension bands were observed compared to the DIBO dUTP (1:0) products (Figure 2.6.B, lanes 1^d and 3^d). However, the main product bands were six and four base extended primers (confirmed by mass spectrometry). The polymerase terminated extension after the modification sites, presumably as a result of DNA duplex distortion caused by the clustered bulky DIBO groups.

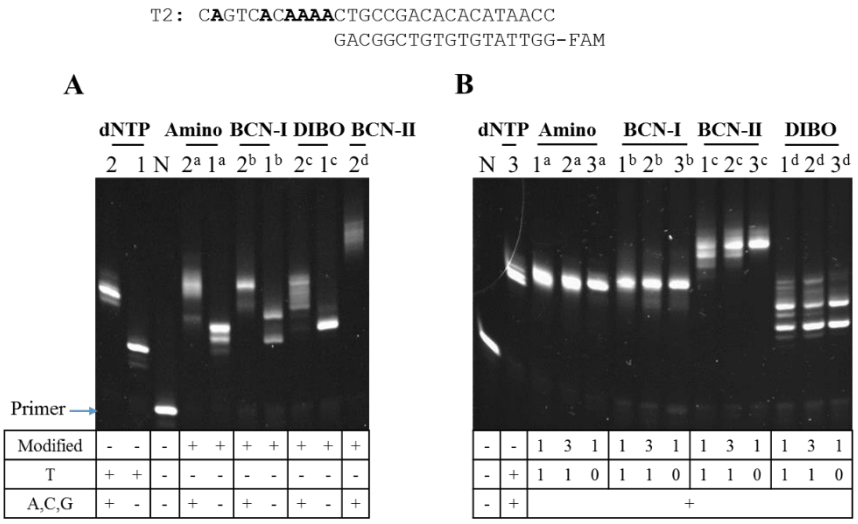


Figure 2.6. Primer extension using modified dUTPs, primer P3 and template T2 for Therminator™ II polymerase (60 °C). 1.6 nmol of each triphosphate was used unless otherwise stated. **A.** Reactions with Therminator™ II (2 units). Lane N, P3 + T1 without triphosphates; lane $1/1^a/1^b/1^c$, dTTP or modified dUTP; lane $2/2^a/2^b/2^c/2^d$, dTTP or modified dUTP + dATP + dCTP + dGTP. **B.** Reactions with Therminator™ II (1 unit) and different ratios of modified dUTP to dTTP + dATP + dCTP + dGTP. Lane $1^a/1^b/1^c/1^d$, dTTP + modified dUTP; lane $2^a/2^b/2^c/2^d$, dTTP + modified dUTP (4.8 nmol); lane $3/3^a/3^b/3^c/3^d$, dTTP or modified dUTP. (20% denaturing PAGE)

DIBO dUTP reaction conditions were further optimised based on the above observations. Mixing the modified triphosphate with a larger proportion of its canonical triphosphate (dTTP) ought to prevent the premature termination of primer extension. This strategy has been widely used to incorporate bulky fluorophore-functionalised triphosphates.[67] The DIBO dUTP to dTTP ratio was therefore lowered to 1:2 and indeed further extended product bands were observed (Figure 2.7.A, lane 4). Moreover, increasing the amount of polymerase from 0.5 unit to 2 units also produced longer elongated products (Figure 2.7.A, lanes 7 to 9; Figure 2.7.B, lanes 1 to 3). However, increasing the amount of DIBO dUTP from 3.2 nmol to 6.4 nmol did not improve the primer extension efficiency (Figure 2.7.A, lane 10). Since Therminator™ II functions more efficiently at 75 °C than 60 °C (according to the supplier), the primer extension reactions were examined at 72 °C. As

expected they proceeded more rapidly than at 60 °C as judged by the amount of the remaining primers in the reactions after 10 min (Figure 2.7.B, lanes 4 and 7). Ultimately, reaction with 2 units of polymerase at 72 °C (close to the optimum temperature) afforded the best extension result, however, the main product was only six bases extended (Figure 2.7.B, lane 3).

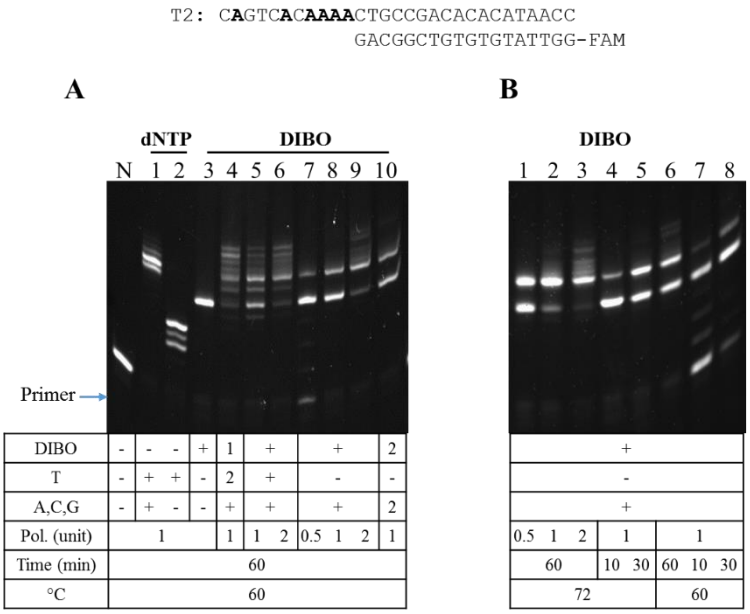


Figure 2.7. Primer extension using DIBO dUTP, primer P3 and template T2 for Therminator™ II polymerase. 1 unit of polymerase and 3.2 nmol of each triphosphate were used unless otherwise stated. **A.** Reactions at 60 °C for 1 h. Lane N, P3 + T1 without triphosphates; lane 1, dNTPs; lane 2, dTTP; lane 3, DIBO dUTP; lane 4, dTTP (6.4 nmol) + DIBO dUTP + dATP + dCTP + dGTP; lanes 5 and 6, dTTP + DIBO dUTP + dATP + dCTP + dGTP (Pol. 1 and 2 units); lanes 7 to 9, DIBO dUTP + dATP + dCTP + dGTP (Pol. 0.5, 1 and 2 units); lane 10, DIBO dUTP + dATP + dCTP + dGTP (6.4 nmol of each). **B.** Reactions with DIBO dUTP + dATP + dCTP + dGTP at 72 °C unless otherwise stated. Lanes 1 to 3, reactions with 0.5, 1 and 2 units of polymerase (1 h); lanes 4 to 5, reactions for 10 and 30 min; lanes 6 to 8, reactions for 60, 10 and 30 min at 60 ºC. (20% denaturing PAGE)

2.3.2 Primer extension using KOD polymerase

Another thermostable DNA polymerase – KOD – was also used in primer extension. KOD is a recombinant form of *Thermococcus kodakaraensis* KOD1 DNA polymerase,

which exhibits efficient and high fidelity amplification. Different versions of KOD polymerase have been used to introduce modified nucleotides into DNA.[102, 160, 178-181] Primer extension reactions with cyclooctyne dUTPs were carried out at 72 ºC close to its optimum temperature (75 ºC). In the presence of dTTP or modified dUTP only, one extra base was incorporated for template T1 (the same as for Therminator™ II, Figure 2.5 and Figure 2.8.A). BCN-I dU fully extended products of both template T1 and T2 were observed by denaturing PAGE and mass spectrometry (Figure 2.8, lane 2^b and Table 2.2). Efficient BCN-II dUTP incorporation was also observed on the denaturing PAGE for template T1 and T2 reactions (Figure 2.8, lane 2^d). However, only a ten-base extended product peak was detected by mass spectrometry analysis for the T2 product (Table 2.2). For the reaction with DIBO dUTP, dATP, dCTP and dGTP, the primer was extended to the end of template T1 efficiently (Figure 2.8.A, lane 2^c). With template T2, nine to ten bases elongated products were obtained (Figure 2.8.B, lanes 2^c and 4^c; Table 2.2). Compared to the Therminator™ II results, which gave mainly six-base extension products for DIBO dUTP using template T2 (Figure 2.7), KOD incorporated DIBO dUTP more efficiently. Multiple bands were observed when cyclooctyne dUTPs were mixed in equal ratios with natural dTTP (Figure 2.8, lanes 3^a, 3^b, 3^c and 3^d). The main bands migrated more close to the fully cyclooctyne-modified product bands than those of amino dUTP (Figure 2.8). This suggests that the cyclooctyne dUTPs competed efficiently with natural dTTP. Degraded DNA strands were also detected by denaturing PAGE and mass spectrometry due to the 3′-5′ exonuclease activity of the polymerase.

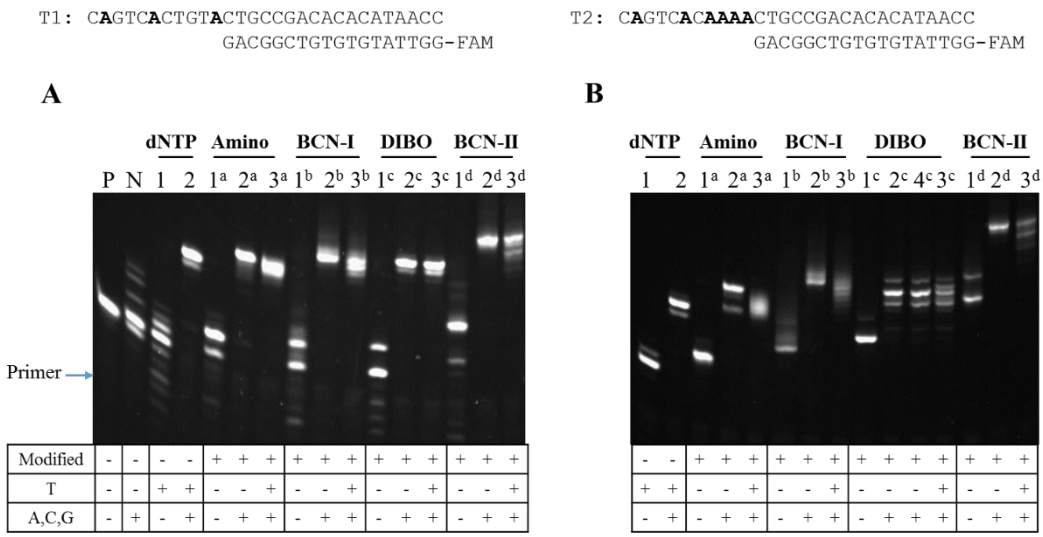


Figure 2.8. Primer extension using modified dUTPs, primer P3, template T1 (**A**) or T2 (**B**) for KOD polymerase (1 unit, 0.4 M $MgCl_2$, 72 °C). 3.2 nmol of each triphosphate was used unless otherwise stated. Lane P, P3; Lane N, P3 + T1 with dATP + dCTP + dGTP; lane $1/1^a/1^b/1^c/1^d$, dTTP or modified dUTP; lane $2/2^a/2^b/2^c/2^d$, dTTP or modified dUTP + dATP + dCTP + dGTP; lane $3^a/3^b/3^c/3^d$, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP; lane 4^c, DIBO dUTP + dATP + dCTP + dGTP reaction with 1 M $MgCl_2$. (20% denaturing PAGE)

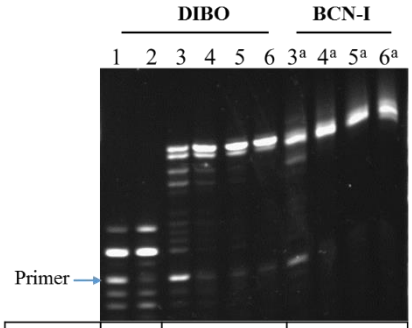
Table 2.2. Mass spectrometry analysis of primer extension using different modified dUTPs and polymerases. A, C, G = dATP, dCTP, dGTP. For representative mass spectra (HPLC-MS) see Appendix 8.2.

	Template T1		**Template T2**	
	dXTP	dXTP+A+C+G	dXTP	dXTP+A+C+G
	Found (Calc.)	Found (Calc.)	Found (Calc.)	Found (Calc.)
$dTTP^{all}$	6442(6442)	9557(9556)	7355(7355)	9554(9554)
BCN-I dUTP	$6657(6657)^{G,K}$	$10202(10201)^{G,K}$	$8215(8216)^{D,K,T}$	$10845(10845)^{D,K}$
BCN-II dUTP	$6902(6901)^{D,G,K}$	$10934(10935)^{G,K}$	$9193(9193)^{D,G,T}$	11981(11982)(10 bases)^{D}
DIBO dUTP	$6728(6727)^{D,G,K}$	$10412(10412)^{D,G,T,K}$	$8496(8496)^{D,K,G,T}$	10936(10936)(10 bases)^{D,K}

Note: The superscripts indicate the specific polymerase(s) for which the mass was confirmed. All = all four polymerases. T = Therminator™ II, G = Gotaq, K = Klenow, D = KOD.

A time course study for the template T1 reaction was also conducted using DIBO and BCN-I dUTPs with KOD polymerase. For single DIBO dUTP incorporation, a higher

conversion was observed when the reaction time was increased from 30 min to 60 min (Figure 2.9, lanes 1 and 2). The reaction containing BCN-I dUTP, dATP, dCTP and dGTP went to completion in 30 min, while DIBO dUTP required more than 60 min to fully extend to the end of the template (Figure 2.9). This indicates that the BCN-I dUTP was incorporated more efficiently than the DIBO dUTP, as observed for the template T2 reactions (Figure 2.8.B).



Modified	+	+	+
A,C,G	-	+	+
Time (min)	30 60	10 30 60 90	10 30 60 90
Template	T1		

Figure 2.9. Primer extension time course for DIBO and BCN-I dUTPs using KOD polymerase (1 unit, 1 M $MgCl_2$). Lanes 1 to 2, DIBO dUTP reactions for 30 and 60 min; lanes 3/3ª to 6/6ª, modified dUTP + dATP + dCTP + dGTP reactions for 10, 30, 60 and 90 min, respectively. (20% denaturing PAGE)

2.3.3 Primer extension using Gotaq polymerase

Both Therminator™ II and KOD are engineered family B DNA polymerases which accept modified triphosphates more efficiently than natural polymerases. To investigate the compatibility of these cyclooctyne modified triphosphates with other polymerases, Gotaq, which belongs to the A-family of DNA polymerases, was also examined in the primer extension assay. It is a full-length form of *Taq* DNA polymerase and exhibits 5′-3′ exonuclease activity. The optimum temperature for Gotaq activity is 75 ºC, whilst primer extension was performed at a slightly lower temperature of 72 ºC. The BCN and DIBO

dUTPs were incorporated efficiently by Gotaq using dispersed-A template T1 as fully extended products were observed by denaturing PAGE and mass spectrometry (Figure 2.10.A and Table 2.2). For template T2 primer extension experiments, only a trace of fully elongated product was obtained for BCN-II dUTP (Figure 2.10.B, lane 2^{d}). For BCN-I and DIBO dUTPs, the primers were extended less than six bases with intense unextended primer bands (Figure 2.10.B). The reason for this may be due to the longer linker of BCN-II which reduces steric hindrance and consequently might make it a better substrate for Gotaq polymerase. Premature termination was reduced by mixing modified dUTP with dTTP for template T2 (Figure 2.10, lanes 3^{a}, 3^{b}, 3^{c} and 3^{d}); several bands, which migrated quicker than the fully modified products and comparably to the unmodified product, were observed for both template T1 and T2 reactions. As expected, compared to Therminator II™ and KOD polymerases (Figure 2.6.B and Figure 2.8), Gotaq is more selective for natural dTTP than modified dUTPs.

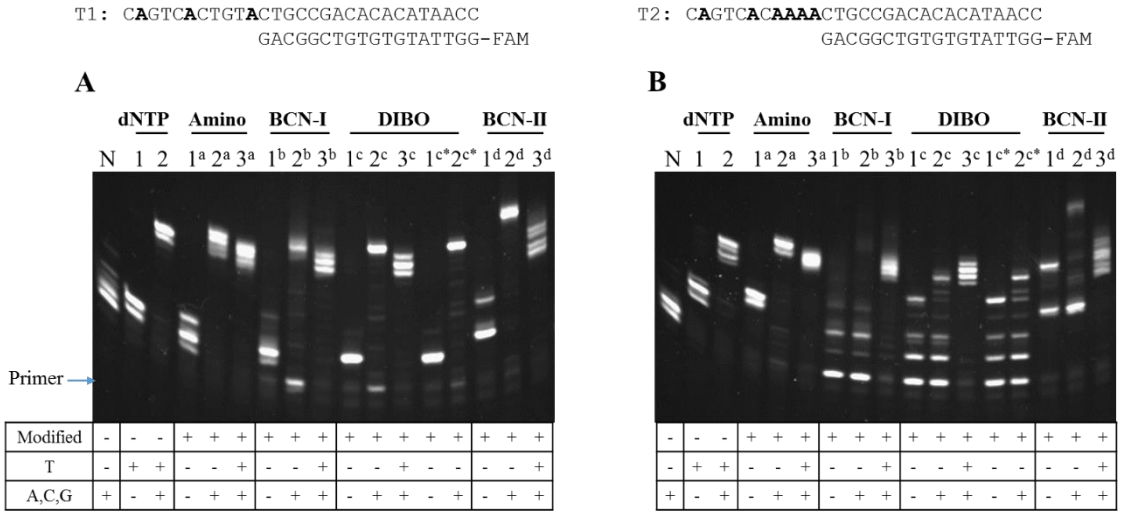


Figure 2.10. Primer extension using modified dUTPs, primer P3, template T1 (**A**) or T2 (**B**) for Gotaq polymerase (1 unit, 72 °C). 3.2 nmol of each triphosphate was used unless otherwise stated. Lane N, P3 + T1 with dATP + dCTP + dGTP; lane $1/1^{a}/1^{b}/1^{c}/1^{d}$, dTTP or modified dUTP; lane $2/2^{a}/2^{b}/2^{c}/2^{d}$, dTTP or modified dUTP + dATP + dCTP + dGTP; lane $3^{a}/3^{b}/3^{c}/3^{d}$, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP; lanes 1^{c*} and 2^{c*}, repeated reactions for DIBO dUTP. (20% denaturing PAGE)

The reaction conditions for cyclooctyne dUTPs using Gotaq polymerase were further optimised for template T2. Increasing the amount of polymerase from 1 unit to 3 units made the reaction proceed further; intense longer extended products were observed for DIBO and BCN-II dUTP reactions by denaturing PAGE analysis (Figure 2.11, lanes $2/2^a/2^b$ to $4/4^a/4^b$). Doubling the concentration of modified triphosphates (Figure 2.11, lanes 5, 5^d and 5^c) or increasing the reaction volume thereby reducing the concentration (Figure 2.11, lanes 6, 6^d and 6^c), however, did not facilitate the reaction.

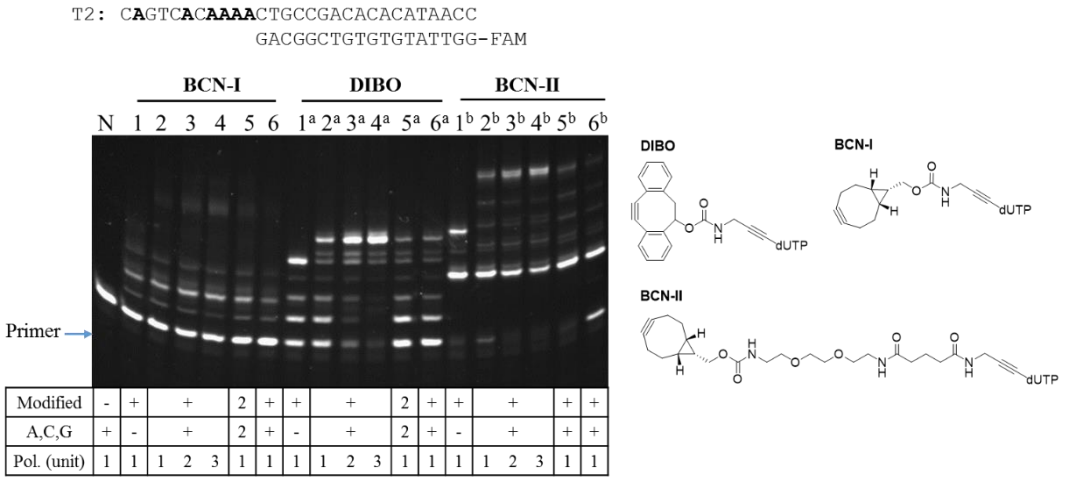


Figure 2.11. Primer extension optimisation using cyclooctyne modified dUTPs, primer P3 and template T2 for Gotaq polymerase (72 °C). Reactions with 1 unit of polymerase in 20 µL were carried out unless otherwise stated. Lane N, P3 + T1 with dATP + dCTP + dGTP; lane $1/1^a/1^b$, modified dUTP; lanes $2/2^a/2^b$ to $4/4^a/4^b$, modified dUTP + dATP + dCTP + dGTP (Pol. 1, 2 and 3 units); lane $5/5^a/5^b$, modified dUTP + dATP + dCTP + dGTP (6.4 nmol of each); lane $6/6^a/6^b$, modified dUTP + dATP + dCTP + dGTP reaction in 40 µL. (20% denaturing PAGE)

2.3.4 Primer extension using Klenow polymerase

Another natural polymerase, DNA polymerase I large (Klenow) fragment was tested for primer extension. It is a family A polymerase and for ease of description is subsequently referred to as Klenow polymerase. It is a proteolytic product of *E. coli* DNA polymerase I which has 3′-5′ exonuclease activity and functions efficiently at 37 ºC.[55, 182] A range of

modified nucleotides have been shown to be accepted as substrates by Klenow to prepare modified DNA.[183-185] Primer extension using cyclooctyne dUTPs showed that partially and fully elongated products were obtained consistently for template T1 as expected (Figure 2.12.A and Table 2.2). Mixing BCN or DIBO dUTP with dTTP in a 1:1 ratio afforded three bands on denaturing PAGE, similar to the previous Gotaq results, indicating that dTTP is its preferred substrate (Figure 2.10.A and Figure 2.12.A, lanes 3^a, 3^b, 3^c and 3^d). For template T2, fully extended products from Therminator™ II polymerase were used as positive controls (Figure 2.12.B). Only amino and BCN-I dUTP with dATP, dCTP and dGTP were extended to full length (Figure 2.12.B, lanes 2^a and 2^b). The DIBO and BCN-II extensions terminated early and no full-length product was obtained (Figure 2.12.B, lanes 2^c and 2^d).

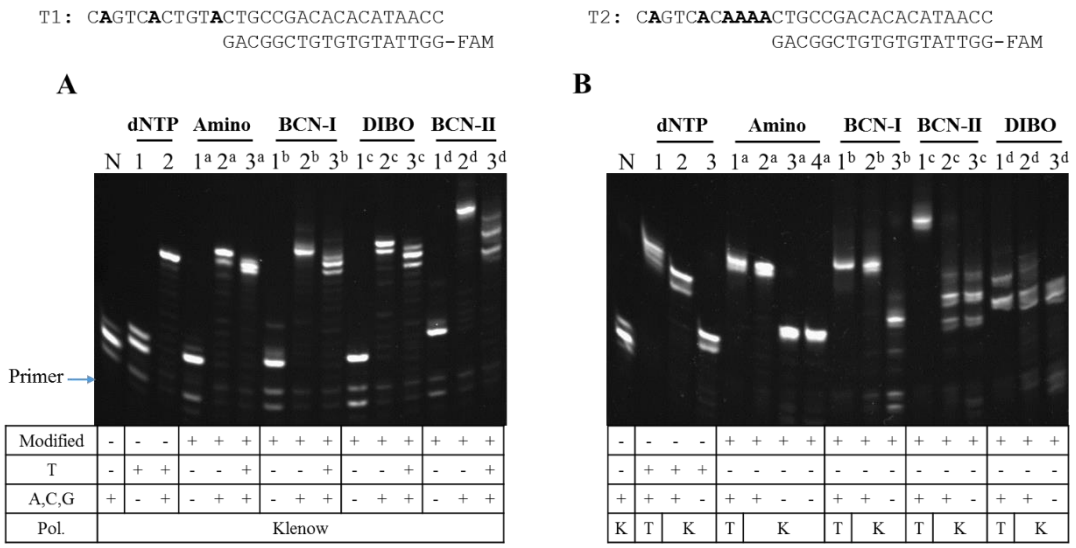


Figure 2.12. Primer extension using modified dUTPs for Klenow polymerase (37 °C). 3.2 nmol of each triphosphate was used unless otherwise stated. **A.** Reactions with primer P3 and template T1 using Klenow (1 unit). Lane N, P3 + T1 with dATP + dCTP + dGTP; lane $1/1^a/1^b/1^c/1^d$, dTTP or modified dUTP; lane $2/2^a/2^b/2^c/2^d$, dTTP or modified dUTP + dATP + dCTP + dGTP; lane $3^a/3^b/3^c/3^d$, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP. **B.** Reactions using Klenow (K) and Therminator™ II (T) with primer P3 and template T2. Lane N, P3 + T1 with dATP + dCTP + dGTP (K, 2 units); lane $1/1^a/1^b/1^c/1^d$, dTTP or modified dUTP + dATP + dCTP + dGTP (T, 1 unit); lane $2/2^a/2^b/2^c/2^d$, dTTP or modified dUTP + dATP + dCTP + dGTP (K, 2 units); lane $3^a/3^b/3^c/3^d$, dTTP or modified dUTP (K, 2 units); lane 4^a, amino dUTP (K, 1 unit). (20% denaturing PAGE)

The reaction conditions for DIBO and BCN-II dUTPs were optimised for template T2 (Figure 2.13.A). Mixing them with dTTP enabled further primer extension (Figure 2.13.A, lanes 3, 3^c, 4 and 4^c). Ten base extended products containing DIBO and BCN-II dU modifications were obtained individually by increasing the amount of polymerase to 2 units (Figure 2.13.A, lanes 7 and 7^a, confirmed by mass spectrometry). The effect of reaction time on Klenow-mediated incorporation was also studied for both template T1 and T2 using amino, BCN-II and DIBO dUTPs. The reactions for T1 were complete in 60 min (Figure 2.13.B, lanes 4^a, 4^b and 4^c), the same as for the amino dUTP T2-templated reaction (Figure 2.13.C, lane 4^a).

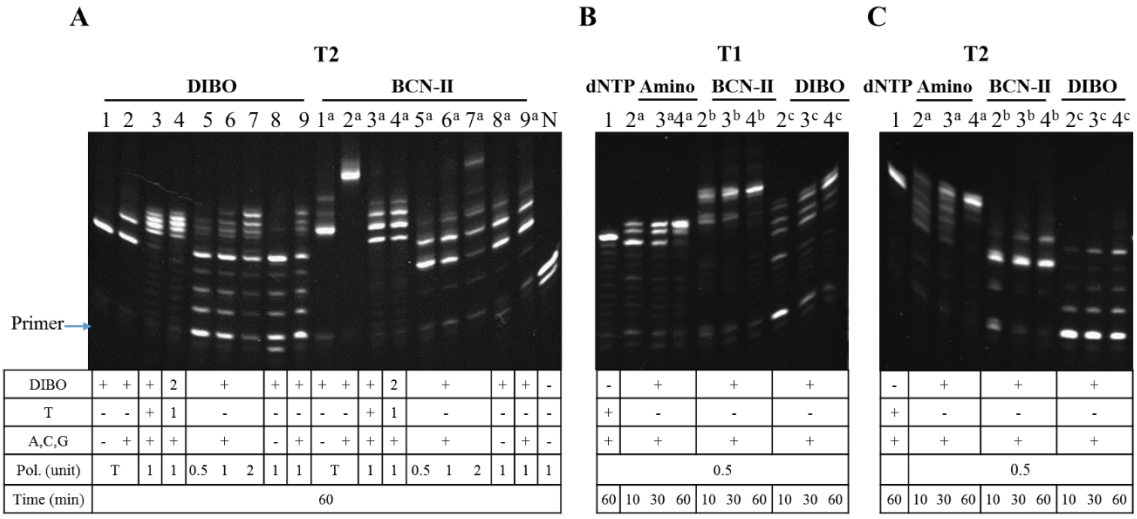


Figure 2.13. **A.** Primer extension optimisation using DIBO and BCN-II dUTPs, primer P3 and template T2 for Klenow polymerase (37 °C). 1 unit of each polymerase and 3.2 nmol of each triphosphate were used unless otherwise stated. Lane N, P3 + T1 without triphosphates; lane $1/1^a$, modified dUTP (Therminator™ II, 60 °C); lane $2/2^a$, modified dUTP + dATP + dCTP + dGTP (Therminator™ II, 60 °C); lane $3/3^a$, dTTP + modified dUTP + dATP + dCTP + dGTP; lane $4/4^a$, dTTP + modified dUTP (6.4 nmol) + dATP + dCTP + dGTP; lanes $5/5^a$ to $7/7^a$; modified dUTP + dATP + dCTP + dGTP (Pol. 0.5, 1 and 2 units); lane $8/8^a$, modified dUTP; lane $9/9^a$, modified dUTP + dATP + dCTP + dGTP reaction in 40 µL. Primer extension time course using template T1 (**B**) or T2 (**C**) for Klenow polymerase (0.5 unit, 37 °C). Lane 1, dNTPs (60 min); lanes $2^a/2^b/2^c$ to $4^a/4^b/4^c$, modified dUTP + dATP + dCTP + dGTP (10, 30 and 60 min). (20% denaturing PAGE)

In summary, for primer extension assays using amino and cyclooctyne dUTPs with four different polymerases, fully extended products with four modified dUTPs without natural dTTP were obtained efficiently for template T1. In the presence of the template T2, BCN dUTPs, which are less bulky than DIBO dUTP, were generally incorporated more efficiently. The DIBO-modified template T2 full-length product was not observed for all polymerases (Table 2.3).

Table 2.3. Summary of the enzymatic compatibility of modified triphosphates (amino, BCN-I, BCN-II, DIBO dUTPs) for various DNA polymerases based on the template T2 primer extension results.

Polymerase	Compatibility of dUTPs
Therminator™ II	Amino ≈ BCN-I ≈ BCN-II > **DIBO**
KOD	Amino ≈ BCN-I > BCN-II > **DIBO**
Gotaq	Amino > BCN-II > **DIBO** > **BCN-I**
Klenow	Amino ≈ BCN-I > BCN-II > **DIBO**

Note: The bold text means that no fully extended product was observed for this modified dUTP.

2.3.5 Steady state kinetics measurements using Gotaq polymerase

Of the four polymerases tested, Gotaq polymerase displayed the least efficient incorporation of the modified triphosphates. To rationalise these observations more in depth, kinetic studies were performed by single nucleotide incorporation for Gotaq polymerase (Figure 2.14.A).[186, 187] Assuming that the incorporation reaction follows the simplified scheme shown in Figure 2.14.B, the reaction kinetics can be described by the Michaelis-Menten equation (Figure 2.14.C). In this equation, K_M is the Michaelis constant and is the substrate concentration when the reaction rate is half of maximal; v_0 and V_{max} are the initial and maximal rate for the enzymatic reaction respectively; k_{cat} is defined as a catalytic constant and is the unimolecular rate constant for productive reaction.

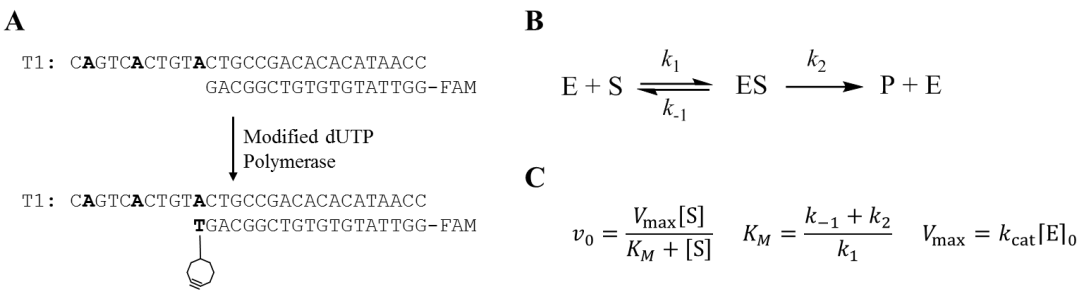


Figure 2.14. **A.** Single nucleotide incorporation in primer extension. **B.** Simplified enzymatic reaction scheme. S is the substrate; E is the polymerase; ES is the substrate and polymerase complex intermediate; P is the product; k_1, k_{-1}, k_2 are the reaction constants for each reaction respectively. **C.** Michaelis-Menten kinetic equation for an enzymatic reaction. [S] is the concentration of substrate; $[E]_0$ is the concentration of polymerase; v_0 and V_{max} are the initial and the maximal reaction rate respectively; K_M is the Michaelis constant; k_{cat} is the catalytic constant.

To determine these parameters, the reactions were carried out with different dTTP or modified dUTP concentrations at 60 ºC rather than 72 ºC. The lower temperature ensured the reaction rate remained within the linear range for subsequent data analysis. One-base extended product yields were determined using gene-snap imaging software according to their relative fluorescence intensities and were used to calculate the V_0. The data were fitted to the Michaelis-Menten equation using Origin to obtain V_{max} and K_M (Figure 2.15). For these measurements, the supplier-defined unit of enzyme concentration was used to calculate k_{cat}.

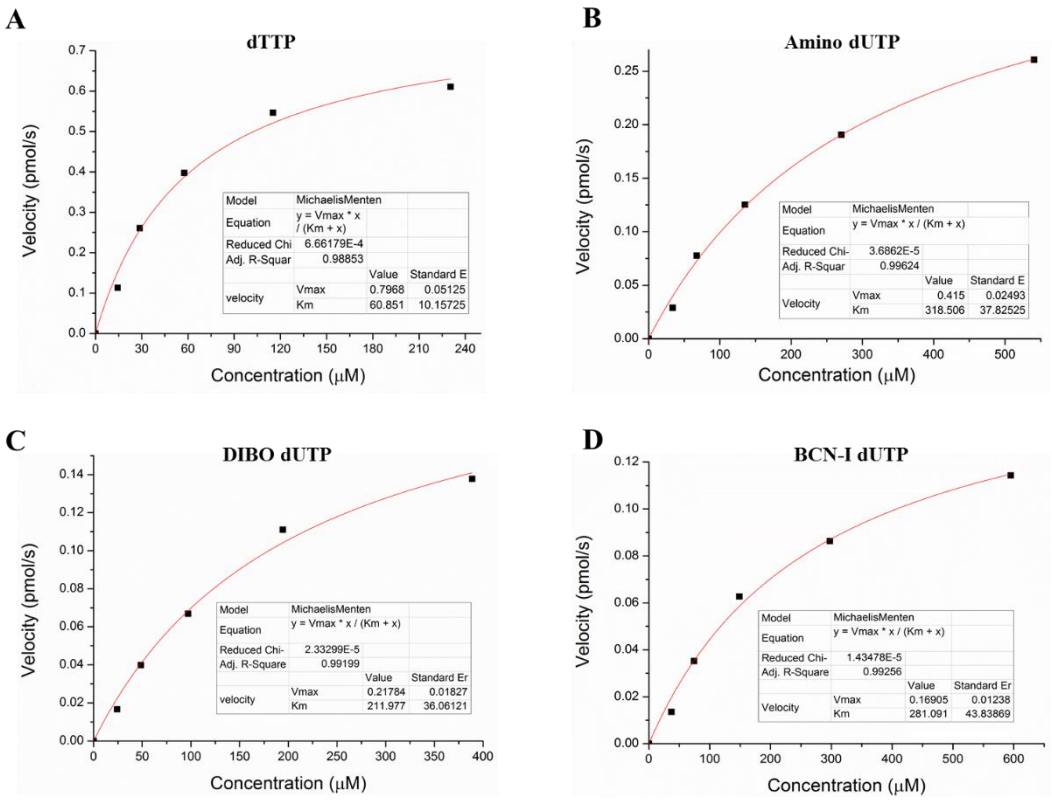


Figure 2.15. Michaelis-Menten kinetic analyses for single nucleotide incorporation using Gotaq polymerase. **A.** dTTP; **B.** Amino dUTP; **C.** DIBO dUTP; **D.** BCN-I dUTP. Representative initial reaction rate calculation was shown in Appendix 8.6.1.

The catalytic efficiency k_{cat}/K_M value is an indication for the incorporation efficiency of a triphosphate by a polymerase. The k_{cat}/K_M value of natural dTTP was set as 1.00 (Table 2.4). Amino dUTP (k_{cat}/K_M = 0.10) was incorporated more efficiently than DIBO dUTP (0.08) and BCN-I dUTP (0.05). However, they were all significantly less efficient than natural dTTP. This clearly demonstrates Gotaq (a family A polymerase) has high selectivity for natural dTTP and is consistent with the previous denaturing PAGE analysis of Gotaq reactions (Figure 2.10 and Table 2.3).

Table 2.4. The kinetics of single triphosphate incorporation using Gotaq polymerase.

Polymerase	dXTP	V_{max} (pmol s^{-1})	K_M (μM)	k_{cat} (pmol s^{-1} $unit^{-1}$)	k_{cat}/K_M Relative value
Gotaq	dTTP	0.8 ± 0.05	61 ± 10	1.6 ± 0.10	1.00
	Amino dUTP	0.42 ± 0.02	318 ± 36	0.84 ± 0.04	0.10
	DIBO dUTP	0.22 ± 0.02	212 ± 34	0.44 ± 0.04	0.08
	BCN-I dUTP	0.17 ± 0.01	281 ± 44	0.34 ± 0.02	0.05

2.3.6 Fluorescent labelling of cyclooctyne-modified primer extension products

After the success in enzymatically introducing cyclooctyne modifications into DNA strands, their labelling efficiencies were investigated by reacting with an azido functionalised fluorophore *via* the SPAAC reaction (Figure 2.16). This strategy can be used to make probes for fluorescent detection of DNA targets and other types of DNA functionalisation. Cy3 is a bright fluorophore and has been widely used in the preparation of fluorescent DNA probes and to study nucleic acid structure and dynamics.[45, 188] Therefore, Cy3-azide **2.16** (synthesised by Dr Marta Gerowska in the Brown group) functionalised with an azide modification on a hexyl linker was used as the model fluorophore.[189]

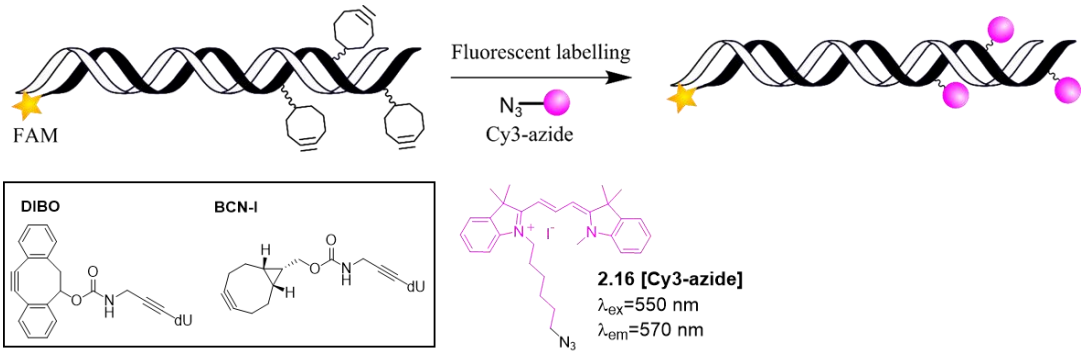


Figure 2.16. Fluorescent labelling of cyclooctyne-modified primer extension products.

The extension products from single cyclooctyne nucleotide incorporation were fluorescently labelled to study the efficiency of 3′-terminal base labelling (Figure 2.17). Denaturing PAGE analysis showed that the BCN-I- and DIBO-modified oligonucleotides were labelled efficiently with Cy3-azide by the SPAAC reaction; there were no unlabelled products detected after 1 h of labelling (Figure 2.17.A). The Cy3 labelled products had much slower electrophoretic mobility compared to the unlabelled controls (Figure 2.17.A) and the Cy3-labelled products can be observed visually as pink bands (Figure 2.17.B). Expected masses were observed by mass spectrometry for all the single terminal Cy3-labelled BCN-I and DIBO products (Figure 2.17.C). Intense fluorescence were obtained (as measured by fluorimetry, Figure 2.17.D). These results confirm that the BCN and DIBO reactive groups are stable to the enzymatic incorporation conditions and can subsequently be fluorescently labelled.

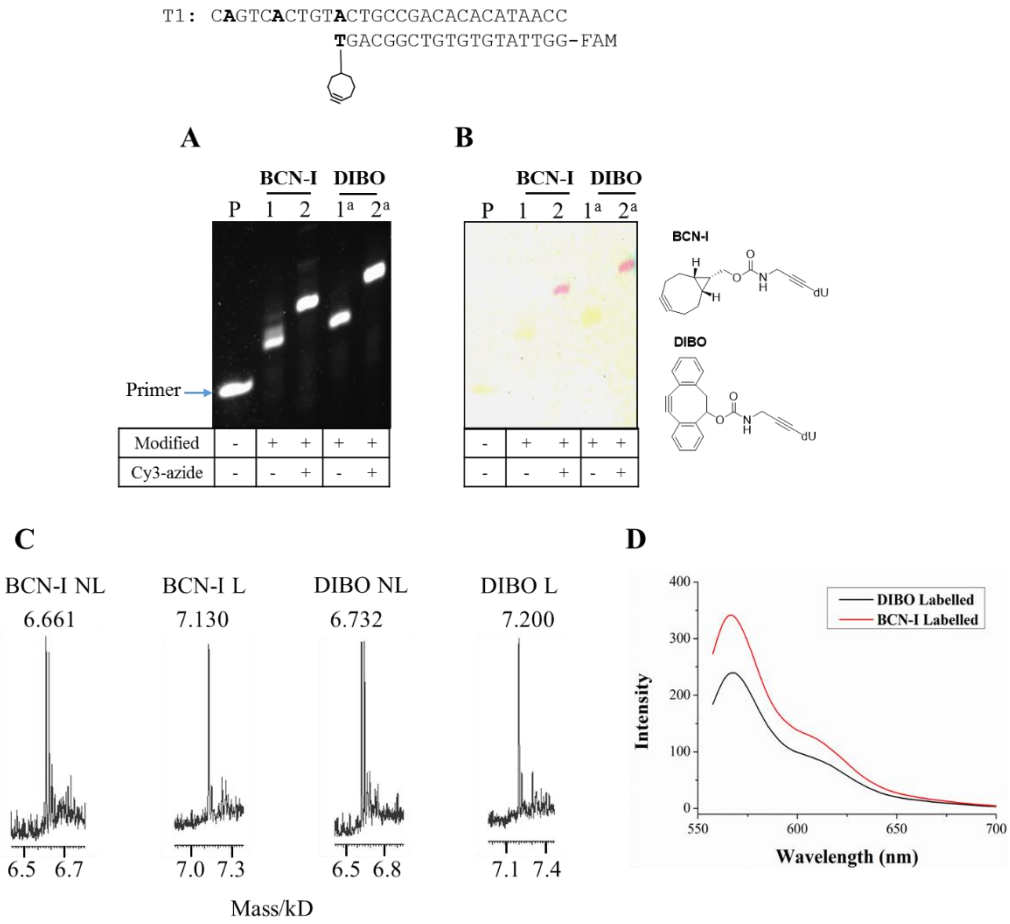


Figure 2.17. Fluorescent labelling of single BCN-I- and DIBO-modified products from Gotaq polymerase. **A.** Gel was visualised on a transilluminator (302 nm). **B.** Gel picture was taken under normal light. Lane P, P3; lane 1/1ª, unlabelled products; lane 2/2ª, cyclooctyne-modified products labelled with Cy3-azide at RT for 1 h. (20% denaturing PAGE) **C.** Mass spectra of unlabelled (NL) and Cy3-azide labelled products (L). BCN-I NL: calc.6657 [M], found 6661 [M] and 6683 [M+Na]; BCN-I L: calc. 7125 [M], found 7130 [M]; DIBO NL, calc. 6727 [M], found 6732 [M] and 6754 [M+Na]; DIBO L, calc.7195 [M], found 7200 [M]. **D.** Fluorescence emission spectra of Cy3-azide labelled products, excited at 540 nm.

Multiple labelling reactions for the fully extended products from both template T1 and T2 were also carried out (Figure 2.18). The labelled products exhibited a similar slow migration rate to that observed above, and produced more intense fluorescent bands compared to the unlabelled products (Figure 2.18.A and B). A pink band was observed for the Cy3 labelled BCN-I dU product from template T2 (Figure 2.18.C). The labelling

reactions between DIBO dUTP **2.5** or BCN-I dUTP **2.12** monomer and Cy3-azide were also carried out to show that the fastest migration bands were due to the labelled cyclooctyne triphosphates which had not been removed (Figure 2.18.A, lane 7/7ª). The Cy3-azide labelled BCN-I dUTP product gave a single band, whereas DIBO dUTP, which is racemic, afforded multiple stereoisomer bands (Chapter 1, Section 1.4.3).

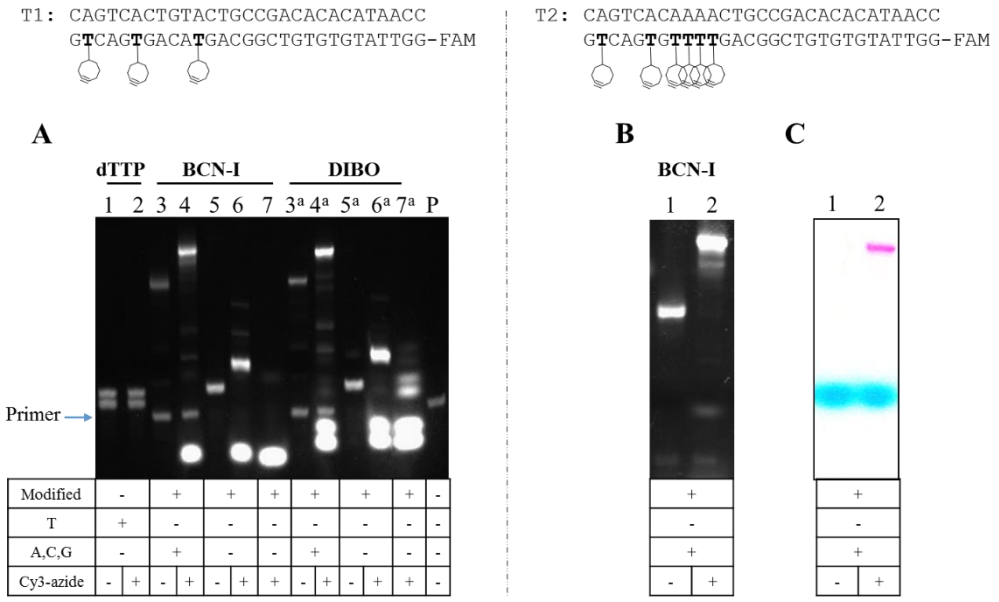


Figure 2.18. Fluorescent labelling of BCN-I- and DIBO-modified full-length products. **A.** Cy3-azide labelling of template T1 products from Gotaq polymerase at RT for 1 h. Lane P, P3 + T1; lane 1, dTTP product; lane 2, lane 1 with Cy3-azide; lane 3/3ª, cyclooctyne dU full-length products; lane 4/4ª, labelled full-length products; lane 5/5ª, cyclooctyne dU partial extension products; lane 6/6ª, labelled patial extension products; lane 7/7ª, labelled modified dUTP monomers. **B** and **C.** Cy3-azide labelling of BCN-I-modified template T2 products from KOD polymerase at RT for 1 h. **C.** Gels were visualised on a transilluminator (302 nm). **C.** Gel picture was taken under normal light. Lane 1, BCN-I dU full-length product; lane 2, labelled full-length product. (20% denaturing PAGE)

2.4 PCR using cyclooctyne dUTPs

Encouraged by the primer extension results, PCR (Chapter 1, Section 1.3.2) was investigated with BCN-I and DIBO dUTPs using KOD polymerase. This enzymatic amplification procedure is an efficient method to introduce densely packed modifications

into both strands of DNA. In PCR, the modified amplification products act as templates for the next cycle of amplification. Therefore, for efficient PCR the polymerase has to accept the cyclooctyne modified dU residues in the template as thymidine analogues in terms of base pairing/stacking, and has to accommodate them in its active site in order to continue the amplification process. Importantly, the incorporation of cyclooctyne dUTPs into DNA during PCR amplification would enable the preparation of long fluorescent probes by the subsequent SPAAC labelling when the amount of available template is limited (Figure 2.19).

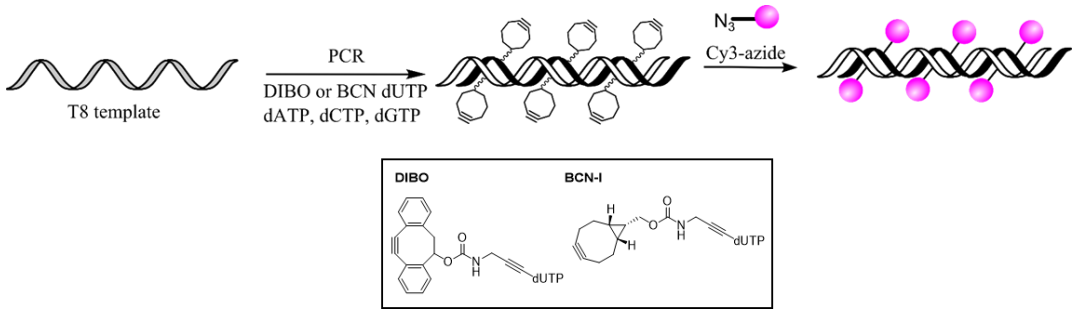


Figure 2.19. PCR and fluorescent labelling of cyclooctyne-modified products.

PCR was carried out with different ratios of BCN-I or DIBO dUTP to dTTP using KOD polymerase and the 81-mer DNA template T8. Canonical dNTPs were used as a positive control for PCR and a negative control for the fluorescent labelling reaction. BCN-I dUTP was successfully incorporated when it was mixed with dTTP in a 1:3 ratio in the presence of dATP, dGTP and dCTP (Figure 2.20.A, lane 3). Increasing the percentage of BCN-I dUTP to 50% inhibited the reaction – only a limited amount of the product was obtained (Figure 2.20.A, lane 2). PCR products were not observed using DIBO dUTP and dTTP mixtures (Figure 2.20.A, DIBO lanes). Labelling of the 25% BCN-I-modified PCR product with Cy3-azide afforded a fluorescent band without ethidium bromide staining (Figure 2.20.B, lane 3*). Subsequent staining of the agarose gel revealed the slower

electrophoretic mobility of this labelled product compared to the unlabelled amplicon, as expected.

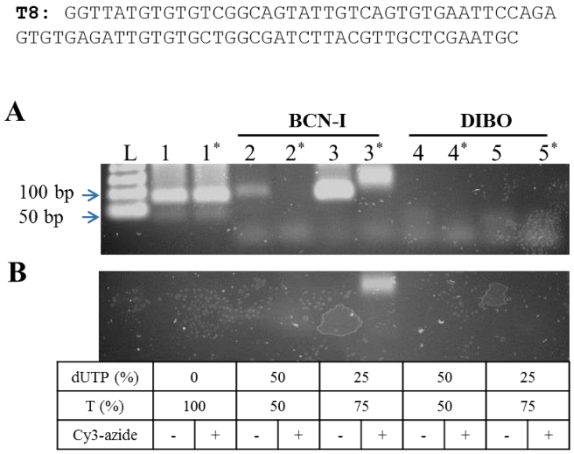


Figure 2.20. Cy3-azide labelling of PCR products from template T8, primer P6, P7 and KOD polymerase (0.5 unit, 1 mM $MgCl_2$) using different ratios of cyclooctyne dUTP to dTTP (4 nmol in total of modified dUTP/dTTP + 4 nmol each of dATP, dCTP and dGTP). Agarose gel analysis of the labelling reactions after staining with ethidium bromide (**A**) and before staining (**B**)**.** Lanes without asterisk are for unlabelled products; lanes with asterisk are for products labelled with Cy3-azide at RT for 30 min. Lane L, 50 bp ladder; lane 1/1*, unmodified amplicon; lanes 2/2* to 3/3*, 50% to 25% BCN-I-modified amplicons; lanes 4/4* to 5/5*, 50% to 25% DIBO-modified amplicons. (2% agarose gel)

2.5 Reverse transcription using cyclooctyne dUTPs

After the success with DNA polymerase reactions (primer extension and PCR), reverse transcriptases were evaluated to determine if the cyclooctyne modified dUTPs are also substrates for them (Figure 2.21). Reverse transcriptase can be used to prepare chemically modified complementary DNA from RNA templates. The cyclooctyne modified dUTPs were studied with Moloney Murine Leukemia Virus reverse transcriptase (M-MuLV, RNase H^-). This enzyme does not have the RNase activity necessary to cleave RNA strands and tolerates some modified triphosphates.[190, 191] The M-MuLV sub-family of

reverse transcriptases has higher accuracy (fidelity) and efficiency of polymerisation than other types of reverse transcriptase.[191-193]

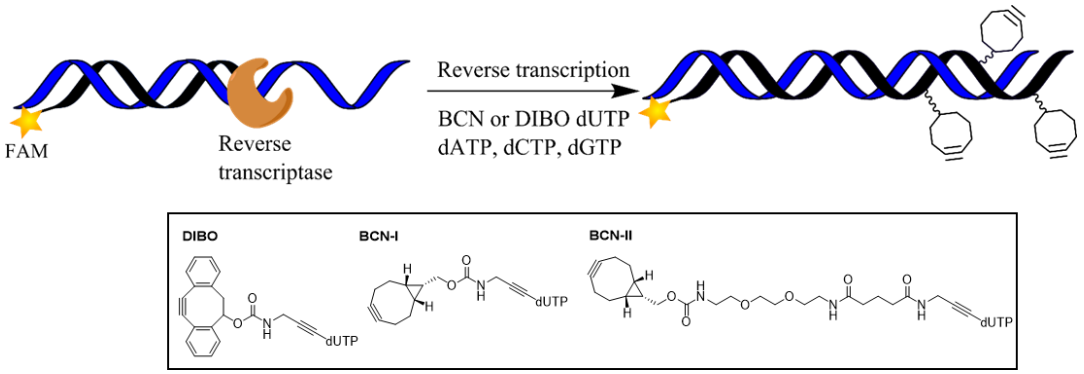


Figure 2.21. Reverse transcription using cyclooctyne dUTPs and a RNA template.

RNA templates T4 and T5 (Table 2.1) have the equivalent sequences to the DNA templates previously used in the primer extension. For the template T4, BCN-II dUTP was incorporated better than BCN-I and DIBO dUTP as judged by the production of full-length products (Figure 2.22.A, lane $2^{b}/2^{c}/2^{d}$). The BCN-II dU fully extended product was observed after 1 h incubation (Figure 2.22.A, lane 2^{d}), while for template T5, only the four-base extended product was obtained using BCN-II dUTP only (Figure 2.22.B, lane 1^{d}). Mixing BCN-II dUTP with 50% dTTP enabled the reverse transcriptase to proceed further and afforded multiple product bands (Figure 2.22, lane 3^{d}). For BCN-I and DIBO dUTPs, intense unextended primer bands were detected for all the reactions (Figure 2.22, BCN-I and DIBO lanes).

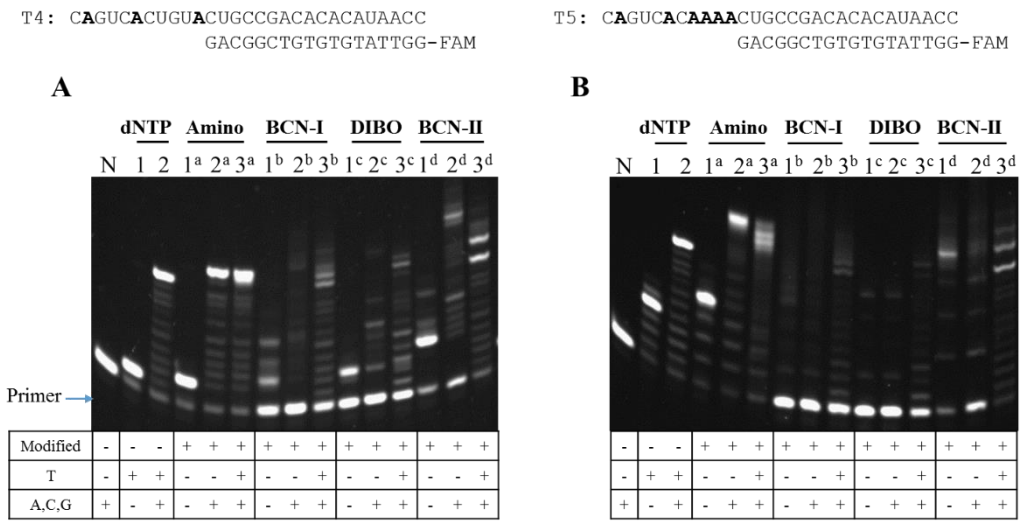


Figure 2.22. Reverse transcription using modified dUTPs, primer P3, RNA template T4 (**A**) and T5 (**B**) for M-MuLV (RNase H^-, 100 units, 1 h). 3.2 nmol for each triphosphate was used unless otherwise stated. Lane N, reaction with dATP + dCTP + dGTP; lane $1/1^a/1^b/1^c/1^d$, dTTP or modified dUTP; lane $2/2^a/2^b/2^c/2^d$, dTTP or modified dUTP + dATP + dCTP + dGTP; lane $3^a/3^b/3^c/3^d$, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP. (20% denaturing PAGE)

Increasing the reaction time to 17 h improved the efficiency of reverse transcription, as multiple bands that migrate comparably to the unmodified product were observed for the reaction with cyclooctyne dUTP and dTTP mixture (Figure 2.23, lanes 3^a and 3^b). Without the introduction of dTTP, intense unextended primer bands remained for the reactions involving modified dUTP, dATP, dCTP and dGTP (Figure 2.23, lanes 2^a and 2^b).

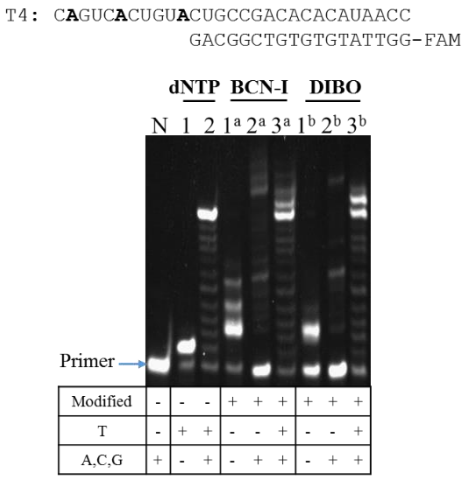


Figure 2.23. Reverse transcription using modified dUTPs, primer P3, and RNA template T4 for M-MuLV (RNase H^-, 100 units, 17 h). 3.2 nmol for each triphosphate was used unless otherwise stated. Lane N, dATP + dCTP + dGTP; lane $1/1^a/1^b$, dTTP or modified dUTP; lane $2/2^a/2^b$, dTTP or modified dUTP + dATP + dCTP + dGTP; lane $3/3^a/3^b$, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP. (20% denaturing PAGE)

2.6 Conclusions

Three cyclooctyne-modified dUTP analogues (DIBO, BCN-I and BCN-II) were synthesised from a common aminopropargyl dUTP intermediate. They were incorporated efficiently into DNA by primer extension using a template with dispersed adenine bases. In the presence of the more demanding template carrying clustered adenines sites, BCN dUTPs were better substrates than DIBO dUTP. KOD and Klenow polymerases incorporated BCN-I dUTP more efficiently than BCN-II dUTP; whereas BCN-II dUTP, which has a longer linker between the base and the cyclooctyne, showed superior incorporation with Gotaq polymerase and M-MuLV (RNase H^-) reverse transcriptase. BCN-I dUTP (25%) with dTTP (75%) supported efficient amplification in PCR reactions. Labelling of the cyclooctyne functionalised primer extension and PCR products with Cy3-azide *via* the SPAAC reaction was successful, demonstrating the efficient incorporation and stability of these modified triphosphates to the conditions encountered

in the enzymatic reactions. This strategy could be used to prepare fluorescent DNA probes and be used in other applications requiring modified DNA, for example systematic evolution of ligands by exponential enrichment (SELEX) with modified aptamers.[158]

Chapter 3 Synthesis and applications of *trans*-cyclooctene-modified deoxyuridine triphosphates

3 Synthesis and applications of *trans*-cyclooctene-modified deoxyuridine triphosphates

3.1 Background and aims

Our success in using the strain-promoted azide-alkyne cycloaddition reaction (SPAAC, Chapter 2) to label cyclooctyne modified DNA encouraged us to study other bio-orthogonal chemistry methodologies to expand the toolbox of modified nucleic acids that can be enzymatically synthesised. Another well-established click chemistry reaction is the inverse electron demand Diels-Alder reaction (IEDDA, Chapter 1, Section 1.4.4), which has reaction kinetics several orders of magnitude faster than the SPAAC reaction.[115, 131] IEDDA has been utilised for fluorescent *in cellulo* imaging of proteins and glycans,[120, 194-196] whilst DNA strands have also been functionalised *in vitro*.[68] Dienophile modified phosphoramidites (norbornene and *trans*-cyclooctene) rather than tetrazine are usually used to functionalise nucleic acids due to the bulky structure and instability of the latter.[133, 138] The cycloaddition reactions between various tetrazines and *trans*-cyclooctenes are extremely fast with rate constants of up to 30 000 $M^{-1}s^{-1}$.[136] However, the highly reactive tetrazine is not stable under aqueous or physiological conditions, therefore 4-methyl-substituted tetrazine, with a lower second-order rate constant of 820 $M^{-1}s^{-1}$, has been more widely used in biological applications due to its higher stability (Figure 3.1).[136] Interestingly fluorophores, such as fluorescein and coumarin, are quenched when they are linked to, or located nearby tetrazines,[132, 197] and the effect is eliminated after tetrazine reacts with *trans*-cyclooctene. This observation can be exploited to give good signal to background ratio in applications requiring fluorescence detection.

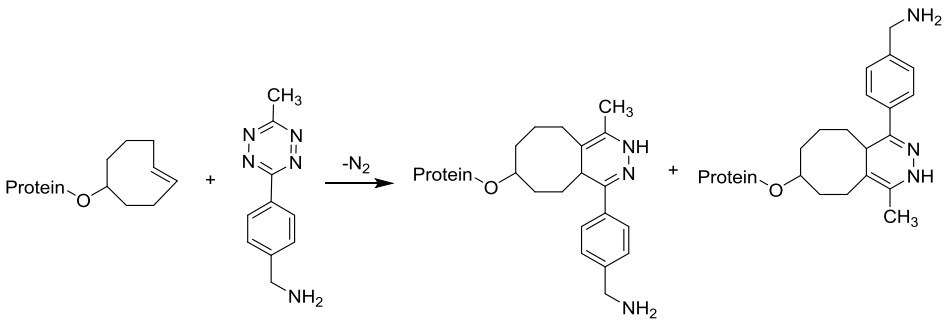


Figure 3.1. An example of the IEDDA reaction in the protein field using *trans*-cyclooctene and 4-methyl substituted tetrazine.[136]

Studies on the enzymatic incorporation of *trans*-cyclooctene (TCO) modified deoxynucleotides are limited. Indeed, at the start of this work, only a norbornene modified DNA triphosphate and a *trans*-cyclooctene cytosine RNA triphosphate had been utilised for enzymatic incorporation into nucleic acids.[134, 140] In this chapter, the compatibility of TCO modified deoxyuridine triphosphate (TCO dUTP) with primer extension, reverse transcription and PCR is investigated. After being incorporated into DNA, TCO modified strands are fluorescently labelled with tetrazine-modified dyes through the IEDDA reaction, which could be used to prepare fluorescent probes or other functionalised DNA.

3.2 Preparation of *trans*-cyclooctene-modified dUTPs

The studies in Chapter 2 confirmed that 5-aminopropargyl-2′-deoxyuridine-5′-triphosphate **2.11** (Amino dUTP) is a good substrate for different DNA polymerases (Klenow, Gotaq, Therminator™ II and KOD) and an ideal intermediate to introduce modifications.[59, 66, 168, 198] The bulky bicyclo[6.1.0]nonyne (BCN) modification, similar to *trans*-cyclooctene, demonstrated good enzymatic incorporation (Chapter 2). Therefore, it was thought likely that *trans*-cyclooctene would be accepted by DNA polymerases. Consequently, two *trans*-cyclooctene (TCO) deoxyuridine triphosphates with different lengths of linker were prepared to study their enzymatic compatibility. Amino dUTP **2.11** was labelled with either

an *N*-hydroxysuccinimide (NHS) carbonate **3.3** or ester **3.4** of *trans*-cyclooctene to afford two TCO-modified triphosphates **3.1** and **3.2** (TCO-S dUTP and TCO-L dUTP, Figure 3.2). They were treated as thymidine surrogates in the subsequent enzymatic incorporation assays.

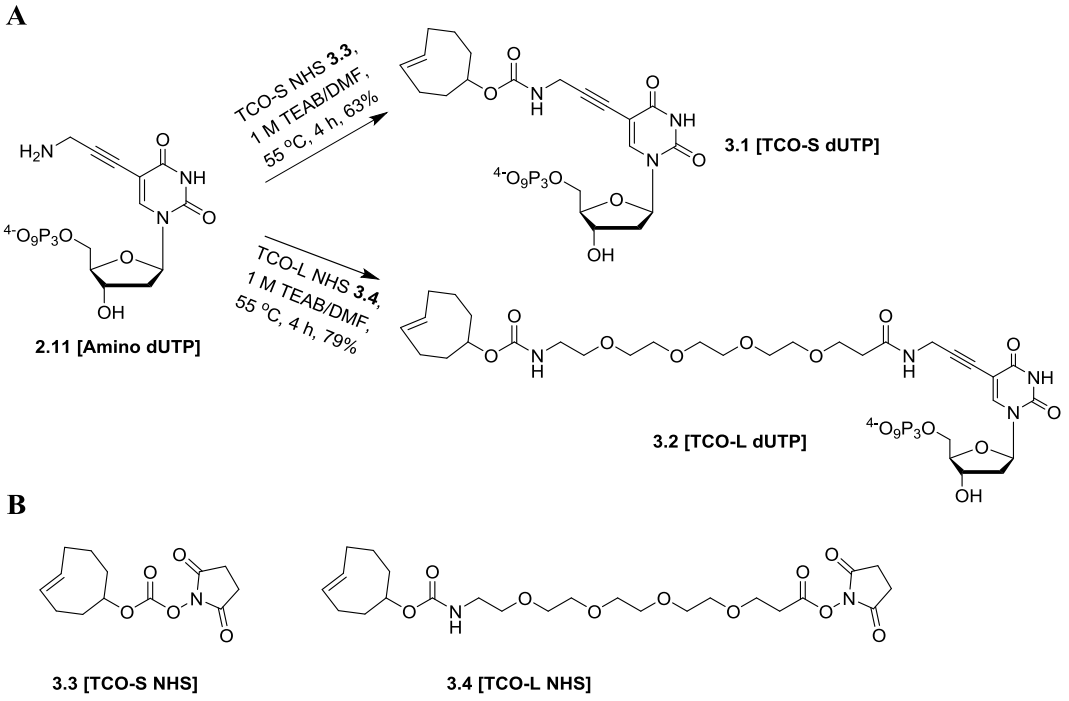


Figure 3.2. **A.** Synthesis of *trans*-cyclooctene-modified triphosphates. **B.** TCO-S NHS carbonate and TCO-L NHS ester used in the syntheses.

3.3 Primer extension using *trans*-cyclooctene dUTPs

As described in Chapter 1 (Section 1.3.2), primer extension is an efficient and simple method to introduce modified triphosphates into DNA. This is done by using equal amounts of a template and a primer. The efficiency of TCO dUTP incorporation in the primer extension assay was evaluated using the same DNA templates as previously described in Chapter 2 for ease of comparison (Figure 3.3). Template T1 has three separated modified triphosphate incorporation sites (Figure 3.4). The more challenging template T2 carries six adenines, four of which are grouped together at the start of the extension region (Figure 3.5). They were

designed to give one- or four-nucleotide extended product in the presence of dTTP or TCO dUTPs only, and yield full-length products with eleven nucleotides incorporated when dTTP or TCO dUTPs are mixed with dATP, dCTP and dGTP. The 5′-end of the primer was functionalised with 6-carboxyfluorescein (FAM) to enable quantitative analysis by analytical denaturing polyacrylamide gel electrophoresis (PAGE). Natural triphosphates dNTPs (dATP, dCTP, dGTP and dTTP) were used as positive controls.

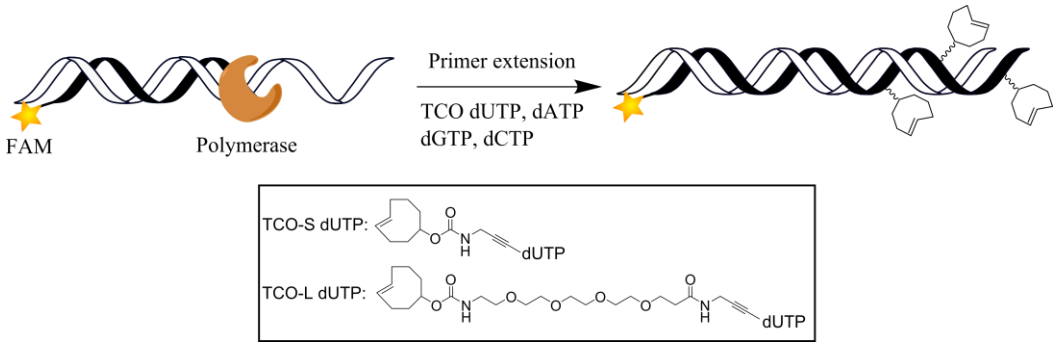


Figure 3.3. Primer extension using TCO dUTPs and a DNA template.

Primer extension with TCO dUTPs was carried out using three different polymerases (KOD at 72 ºC, Gotaq at 60 ºC, Klenow at 37 ºC); reactions were analysed by denaturing PAGE and mass spectrometry. TCO-S and TCO-L dUTPs were successfully incorporated in the presence of template T1 by KOD, Gotaq and Klenow polymerases (Figure 3.4 and Table 3.1). Bands corresponding to the full-length products were identified by correlation with the expected theoretical masses. The bulky TCO-S and TCO-L dU bases resulted in slower electrophoretic mobility of the desired products compared to their natural counterparts (Figure 3.4). For the KOD reaction (with dTTP or modified dUTP only) at 72 °C for 30 min with 2 mM $MgCl_2$, multiple digested oligonucleotide bands were observed (Figure 3.4.C, lanes 1, 1^a and 1^b). This is because of the strong 3′-5′ exonuclease activity of KOD polymerase, which is dependent on the concentration of Mg^{2+}.[199, 200] In order to decrease the rate of digestion, primer extension reactions with modified dUTPs were conducted with

1 mM $MgCl_2$. Sufficient quantities of products were generated to confirm the identity of the partially extended products by mass spectrometry (Table 3.1).

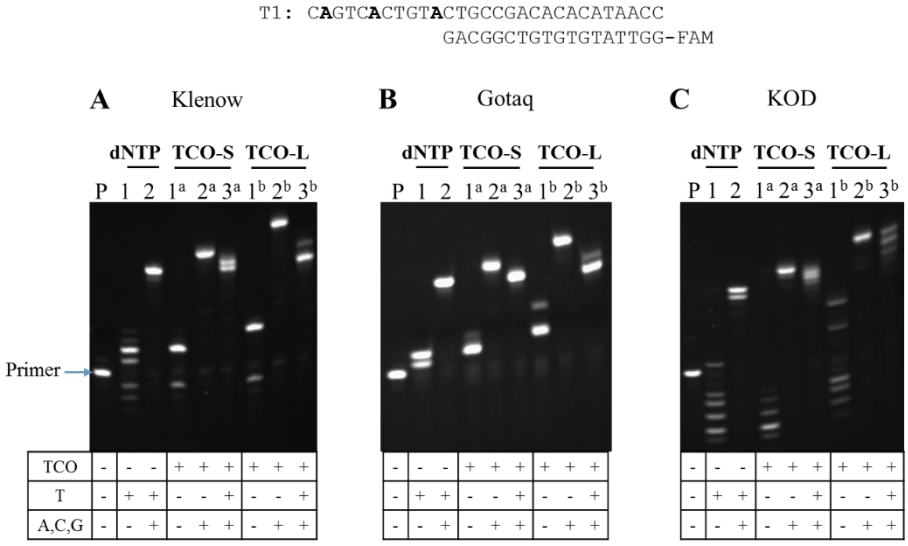


Figure 3.4. Primer extension using TCO dUTPs, primer P3 and template T1 for Klenow (**A**, 37 °C, 1.5 h), Gotaq (**B**, 60 °C, 1.5 h) or KOD (**C**, 72 °C, 1.5 h for lanes $2/2^a/2^b$ and $3^a/3^b$; 0.5 h for lane $1/1^a/1^b$) polymerases. 3.2 nmol of each triphosphate was used unless otherwise stated. Lane P, primer P3; lane $1/1^a/1^b$, dTTP or modified dUTP; lane $2/2^a/2^b$, dTTP or modified dUTP + dATP + dCTP + dGTP; lane $3^a/3^b$, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP. (20% denaturing PAGE)

Table 3.1. The mass spectrometry analysis of primer extension (Klenow, Gotaq and KOD) and reverse transcription (M-MuLV (RNase H^-) and AMV) using dTTP, TCO-S and TCO-L dUTPs. Each product mass peak was observed for all enzymes unless otherwise stated. A, C, G = dATP, dCTP, dGTP. For representative mass spectra see Appendix 8.3.

	Template T1 or T4		**Template T2 or T5**	
	dXTP	dXTP+A+C+G	dXTP	dXTP+A+C+G
	Found (Calc.)	Found (Calc.)	Found (Calc.)	Found (Calc.)
dTTP	6442(6442)	9555(9556)	7355(7355)	9553(9553)
TCO-S dUTP	6633(6633)	10129(10129)	8120(8120)	10701(10700)
TCO-L dUTP	6880(6880)	10870(10871)	9108(9109)	12184(12184)[G]

[G] Only Gotaq polymerase gave the expected product.

The number of reactive groups introduced *via* the enzymatic incorporation can be manipulated by mixing modified dUTP and dTTP in the same extension reaction.[71, 73] This method can be used to control the number of fluorophores in the final labelled DNA product to avoid collisional fluorescence quenching. This strategy requires the modified dUTP to compete with dTTP for incorporation. To evaluate this, primer extension reactions were carried out in the presence of a 1:1 ratio of TCO dUTP to dTTP in addition to dATP, dCTP and dGTP. Multiple or broad bands were observed by denaturing PAGE (Figure 3.4, lanes 3^a and 3^b), suggesting that full-length products with a mixture of dT and TCO dU were produced as opposed to a single sharp product band with dT or TCO dU only. For Klenow and Gotaq reactions, the major products migrated comparably to the natural dNTPs extension product (Figure 3.4.A and B), suggesting that these two polymerases are more selective for the natural dTTP over the TCO dUTPs. On the other hand, KOD polymerase gave multiple bands, which migrated closer to fully TCO dU modified products (Figure 3.4.C). This is probably due to KOD DNA polymerase being an engineered family B polymerase, which accepts modified triphosphates more efficiently than natural polymerases.[160, 201]

For primer extension with the more demanding template T2, once again, all three polymerases extended the primer to the end of the template in the presence of TCO-S dUTP, dATP, dCTP and dGTP (Figure 3.5, lane 2^a and Table 3.1). In contrast to T1, the full-length product of TCO-L dU was only obtained using Gotaq polymerase (Figure 3.5.B, lane 2^b and Table 3.1); nine to ten bases extended products were observed for the KOD polymerase reaction (Figure 3.5.C, lane 2^b). The long tetraethylene glycol linker of TCO-L dUTP seemed to be advantageous only for the Gotaq polymerase reactions. Furthermore, TCO-S dUTP was incorporated more efficiently than TCO-L dUTP using Klenow and KOD polymerases as judged by the presence of shorter products for the latter (Figure 3.5.A and

C). For Klenow polymerase, mixing TCO-L dUTP with dTTP in a 1:1 ratio produced multiple product bands (Figure 3.5.A, lane 3^b) which migrated more slowly than, or similar to the unmodified product (lane 2), suggesting significant incorporation of TCO-L dUTP. This result also proves that blending a modified triphosphate with its natural counterpart prevents early reaction termination, as previously demonstrated for the enzymatic incorporation of bulky cyclooctyne-modified triphosphate substrates (Chapter 2, Section 2.3).[67]

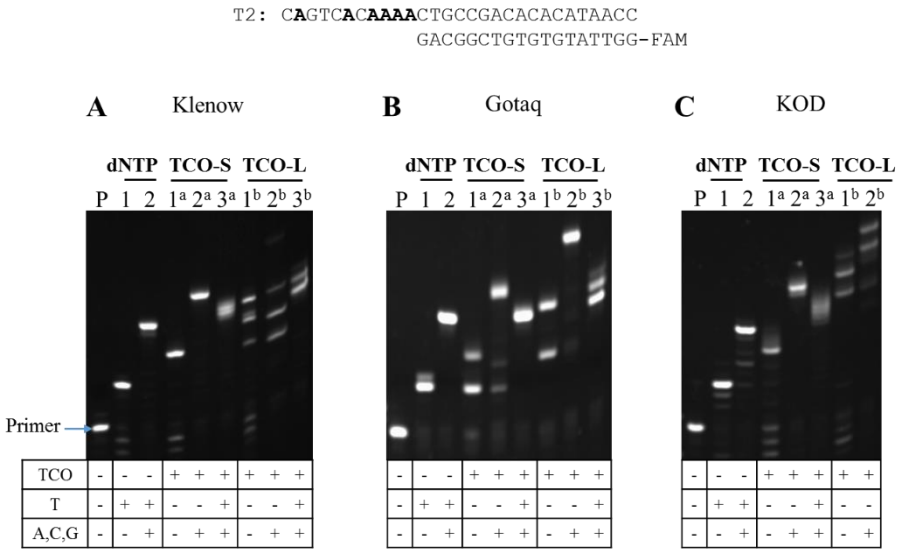


Figure 3.5. Primer extension using TCO dUTPs, primer P3 and template T1 for Klenow (**A**, 37 °C, 1.5 h), Gotaq (**B**, 60 °C, 1.5 h) or KOD (**C**, 72 °C, 2 mM $MgCl_2$, 1.5 h for lanes $2/2^a/2^b$ and $3^a/3^b$; 0.5 h for lane $1/1^a/1^b$) polymerases. 3.2 nmol of each triphosphate was used unless otherwise stated. Lane P, primer P3; lane $1/1^a/1^b$, dTTP or modified dUTP; lane $2/2^a/2^b$, dTTP or modified dUTP + dATP + dCTP + dGTP; lane $3^a/3^b$, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP. (20% denaturing PAGE)

3.3.1 Thermal stability of *trans*-cyclooctene-modified primer extension products

After the successful incorporation of TCO dUTPs, the effect of these modifications on DNA duplex stability was studied. For hybridisation applications, an ideal modified DNA probe should have similar or better annealing properties for the target nucleic acid to the

unmodified probe.[202] Fluorescent melting temperature measurements (Chapter 1, Section 1.1.5) were carried out by adding SYBR Green to the Gotaq primer extension reactions, which had been also analysed by denaturing PAGE (Figure 3.4.B and Figure 3.5.B). The melting temperatures of partially extended products using only TCO dUTPs were the same as that of the dTTP control for both template T1 and T2 (Figure 3.6, T1-Part and T2-Part). However, fully extended TCO-L dUTP products showed slightly lower melting temperatures than natural dNTP products. For template T1, three TCO-L dU modifications decreased the T_m by 1 ºC than the natural dNTP product, whilst for template T2 six modifications reduced the T_m by 2 ºC (Figure 3.6, T1-Full and T2-Full). The T_m of the TCO-S dU full-length product from template T2 with six modifications was 1 ºC lower relative to the unmodified product (Figure 3.6, T2-Full). Overall the differences in T_m are too minor to determine a trend in terms of stability for this modification.

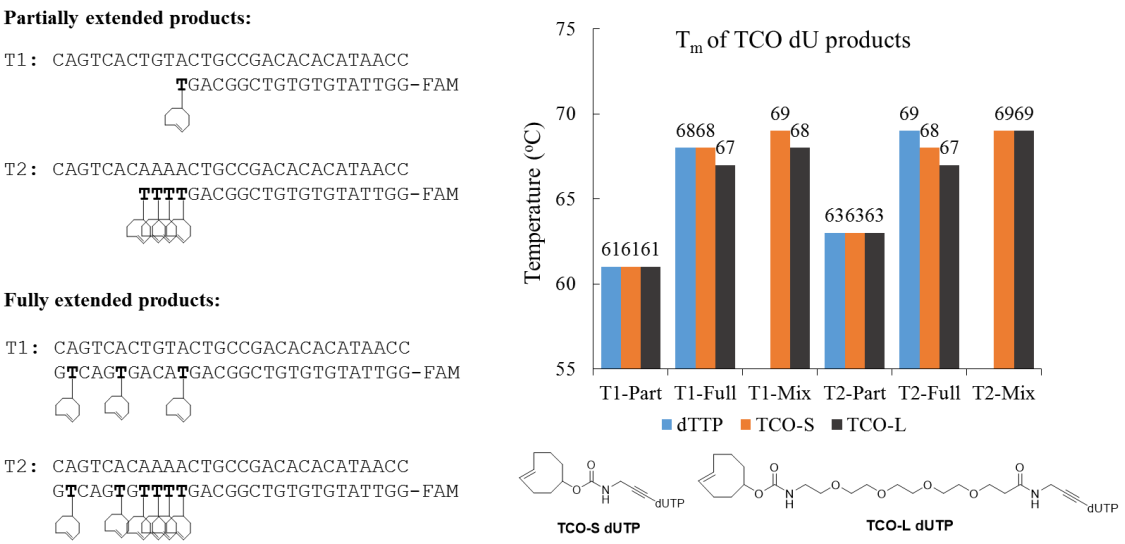


Figure 3.6. Melting temperatures of TCO-modified primer extension products from template T1 and T2 using Gotaq polymerase. Part, partially extended products; full, fully extended products; mix, reactions with dTTP and TCO dUTP (1:1) + dATP + dCTP + dGTP. Melting curves see Appendix 8.5.1.

3.3.2 Fluorescent labelling of *trans*-cyclooctyne-modified primer extension products

In Chapter 2, cyclooctyne modified DNA products were labelled with Cy3-azide *via* the SPAAC reaction. In addition to the SPAAC labelling, the BCN group has also been reported to react with tetrazine through the IEDDA reaction, which can be used to functionalise nucleic acids and proteins.[120, 141] This reaction is faster than the SPAAC reaction and would be a good alternative labelling strategy to functionalise BCN-modified oligonucleotides. However, TCO dUTP is a better substrate than BCN dUTP as shown in the primer extension studies and confirmed by PCR studies detailed later in the chapter (Section 3.4). The fluorescent labelling of *trans*-cyclooctene-modified products by the IEDDA reaction is examined in this chapter (Figure 3.7). Commercially available 4-methyl-tetrazine-sulfo-Cy3 **3.5** (Cy3-tetrazine, Figure 3.7) was used; it contains three anionic sulfonate modifications to improve aqueous solubility.[203]

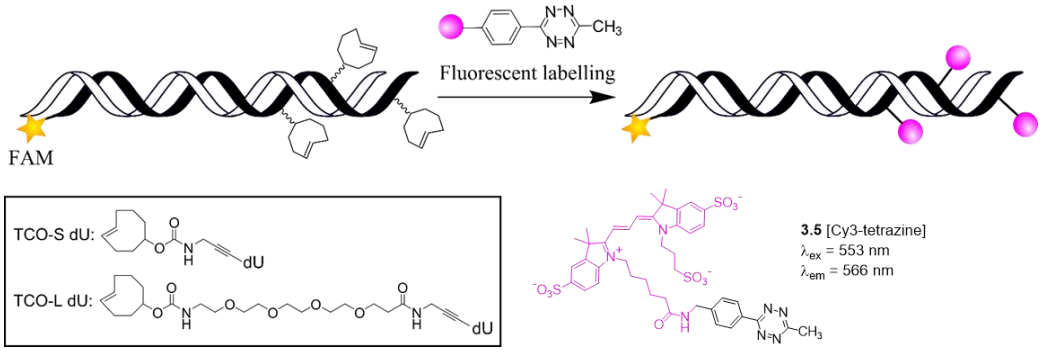


Figure 3.7. Fluorescent labelling of TCO-modified primer extension products.

Single nucleotide primer extension products from template T1 were used to evaluate the efficiency of single terminal base labelling (Figure 3.8). The IEDDA labelling reaction went to completion within 30 min (Figure 3.8.A). The labelled products exhibited much slower electrophoretic mobility compared to the unlabelled controls due to the bulky anionic

sulfo-Cy3 fluorophore. Correct mass peaks for the Cy3-tetrazine labelled products were observed (Figure 3.8.B).

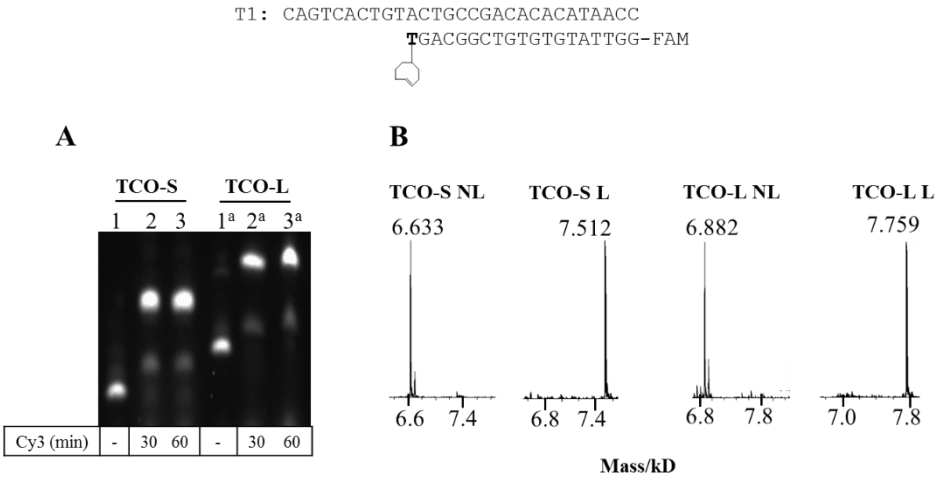


Figure 3.8. Terminal fluorescent labelling of TCO-modified partially extended products from Gotaq polymerase and template T1. **A.** Time course of fluorescent labelling of single nucleotide modified products with Cy3-tetrazine. Lane 1/1^{a}, unlabelled products; lanes 2/2^{a} and 3/3^{a}, labelling reactions for 30 and 60 min, respectively. (20% denaturing PAGE) **B.** Mass spectrometry characterisation of unlabelled (NL) and Cy3-tetrazine labelled (L) products. TCO-S NL: calc. 6633, found 6633; TCO-S L: calc. 7513, found 7512; TCO-L NL: calc. 6880, found 6882; TCO-L L: calc. 7760, found 7759.

Labelling reactions on fully extended T1 products carrying three TCO modifications were also performed (Figure 3.9). After 1 h, intense product bands labelled with two and three Cy3 dyes were observed for both TCO-S and TCO-L strands (Figure 3.9, lane 2). The TCO-L product labelled with one Cy3-tetrazine gave a weaker band than the TCO-S product (Figure 3.9, lane 2). As expected, this suggests that the longer linker of TCO-L dU is subject to less steric hindrance (from the oligonucleotide chain) and reacts with the tetrazine more efficiently than TCO-S dU. The full-length products with TCO in the middle of a DNA duplex exhibited much lower labelling efficiency compared to the above terminal labelling reactions (Figure 3.8 and Figure 3.9). Extension of the reaction time, or the use of more

equivalents (higher concentration) of Cy3-tetrazine might increase the efficiency of labelling.

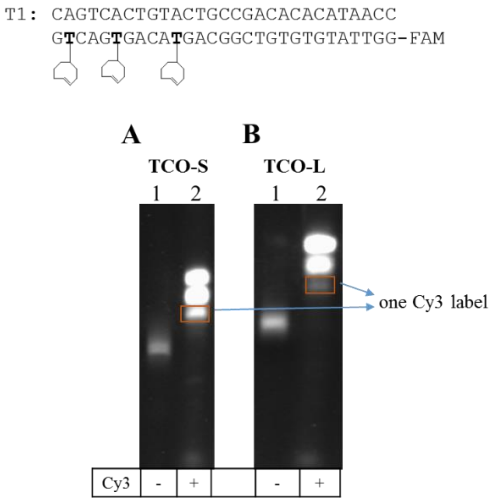


Figure 3.9. Fluorescent labelling of TCO-modified full-length products from Gotaq polymerase and template T1. TCO-S (**A**) and TCO-L (**B**) products labelled with Cy3-tetrazine. Lane 1, Unlabelled products; lane 2, labelling reactions at RT for 1 h. (20% denaturing PAGE)

3.4 PCR using *trans*-cyclooctene dUTPs

Primer extension studies have demonstrated that TCO dUTPs can be accepted as substrates by different DNA polymerases (Klenow, Gotaq and KOD). However, this is linear extension as opposed to PCR amplification, and functional groups are only introduced into one strand of the duplex. Next, PCR amplification was investigated since it requires far less template than primer extension, can afford densely functionalised DNA products, and is compatible with double-stranded templates (Figure 3.10).

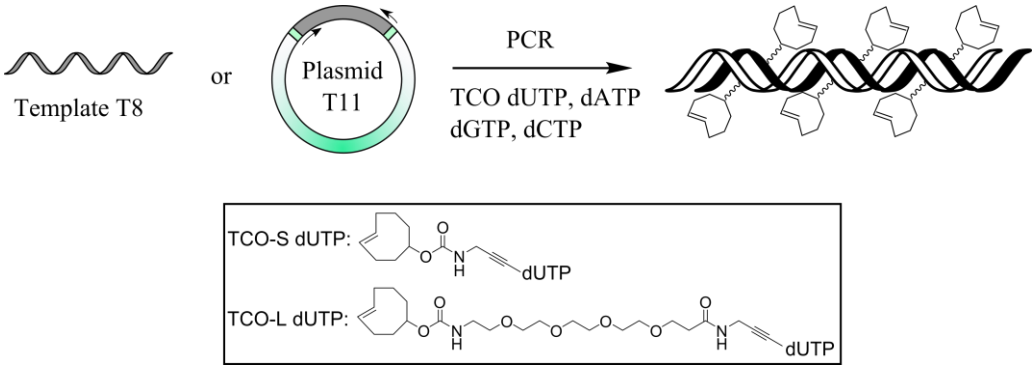


Figure 3.10. PCR using TCO dUTPs and different templates. For the plasmid template T11, the two arrows indicate the amplification direction and the region being amplified is in grey.

Table 3.2. Oligonucleotide sequences used in PCR.

Code	Sequences(5′-3′)
P6	GCATTCGAGCAACGTAAG (PCR primer for T8)
P7	GGTTATGTGTGTCGGCAG (PCR primer for T8)
T8	GGTTATGTGTGTCGGCAGTATTGTCAGTGTGAATTCCAGAGTGTGAGATTGTGTG CTGGCGATCTTACGTTGCTCGAATGC (PCR template)
P9	GTTTGGCTTTAGAGGCTGGAG (PCR primer for T11)
P10	ACTGCAATACGAATAATGGCTAC (PCR primer for T11)
T11	501 AAGATATGGG ACATAAAC**GT TTGGCTTTAG AGGCTGGAG**A AGACCCTGTA 551 AACAATCCTA TTGAATATAT TCTTGACTGT ATCAAAACCA TATACAGCAT 601 AAAACATAAA AATGGAGCAA TTAGACGTGT AAATGTAAAT ATTGCAGCTA 651 CTACTGTAGA AAACTACAAG AAATTAAAGG ATGCTGGTAT TGGAACATAT 701 ATACTTTTCC AAGAAACCTA TAACAAAAAA AGTTACGAGG AACTTCATCC 751 TACAGGTCCA AAACATGATT ATGCCTATCA TACAGAAGCA ATGGATCGTG 801 CTATGGAAGG TGGTATTGAT GATGTAGGTA TTGGGGTTTT GTTTGGACTA 851 AATATGTACA AATATGACTT TGTTGGACTT CTAATGCATG CTGAACACTT 901 GGAAGCTGCT ATGGGTGTAG GCCCTCATAC TATAAGCGTT CCTCGTATAC 951 GTCCTGCAGA TGACATTGAT CCTGAAAACT TCTCAAATGC AATATCGGAC 1001 GAGATTTTTG AAAAAATT**GT AGCCATTATT CGTATTGCAG T**TCCATACAC (amplified region of plasmid template HydGdCTD5)

Initially, PCR was carried out using a short 81-mer DNA template (T8, Table 3.2) and Gotaq polymerase. To control the number of modifications introduced, reactions were conducted in the presence of different ratios of TCO dUTP to dTTP with dATP, dCTP and dGTP. In

all cases PCR amplifications were successful, even when using only TCO-modified triphosphates with no additional dTTP (Figure 3.11.A). A more intense band was observed for TCO-L dU fully modified amplicon compared to TCO-S dU (lanes TS4 and TL4), indicating that the longer linker of TCO-L dUTP makes it a better substrate for Gotaq polymerase. The major difference in migration between the 100% TCO dU product band and the 75% product suggest that in the latter reaction mainly dTTP was incorporated in the product (Figure 3.11.A, lanes TS3, TS4, TL3 and TL4). These observations were consistent with previous primer extension results, in that Gotaq is more selective towards natural dTTP than modified nucleotides (Section 3.3).

The duplex stability of TCO-modified PCR products (T_m = 77 ºC) did not change when the TCO-S dUTP to dTTP ratio was varied from 25% to 75% (Figure 3.11.B, TS1 to TS3) and slightly increased (+1 ºC) for the 100% TCO-S dU product (78 ºC, TS4). However, there was a 6 °C difference between the full TCO-L-modified product (71 °C) and the unmodified product (77 °C). These results suggest that the TCO-S functional group is better than the TCO-L in stabilising double-stranded DNA.

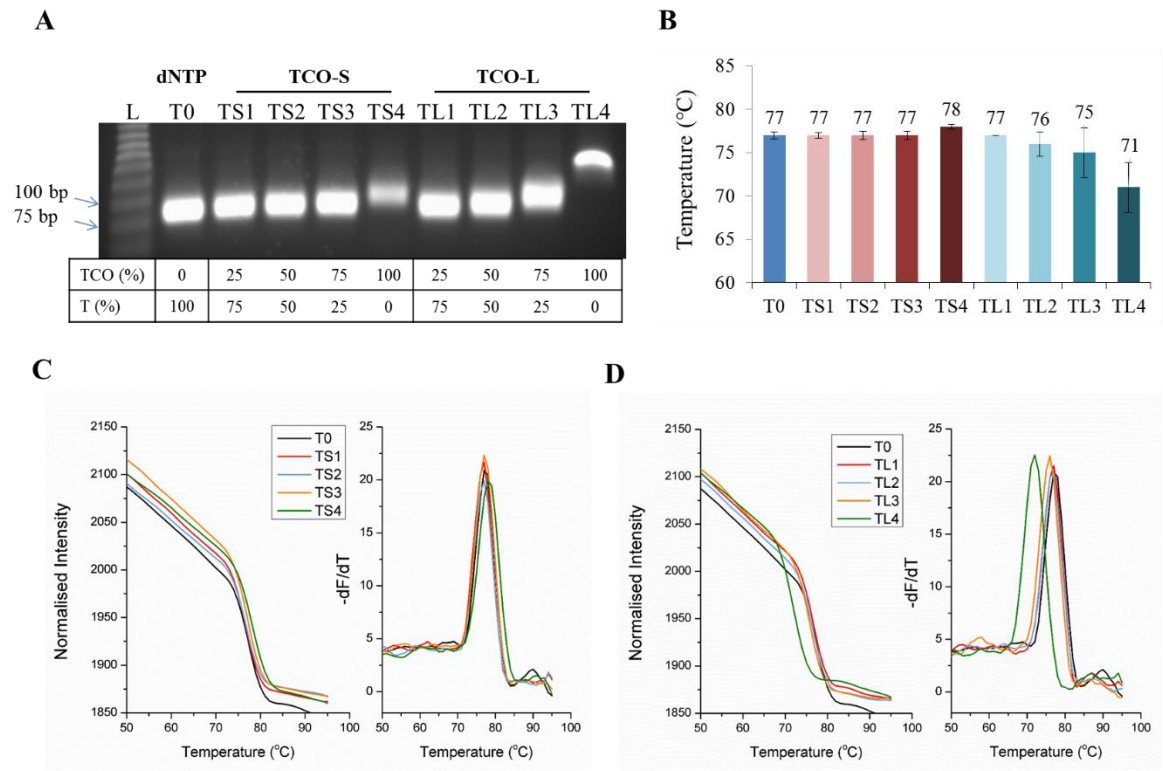


Figure 3.11. **A.** PCR amplifications using Gotaq polymerase and template T8 with different ratios of TCO dUTP to dTTP (5 nmol in total of TCO dUTP/dTTP + 5 nmol each of dATP, dCTP and dGTP). Lane L, 25 bp ladder; lane T0, unmodified amplicon; lanes TS1/TL1 to TS4/TL4, 25% to 100% TCO-S/TCO-L amplicons. (2% agarose gel containing ethidium bromide) **B.** Melting temperatures of TCO-modified PCR amplicons. Fluorescence melting curves and derivatives of TCO-S (**C**) and TCO-L (**D**) modified PCR amplicons.

To explore the feasibility of preparing longer functionalised DNA products, a 523 bp region of a plasmid containing a total of 335 A-T base pairs (T11, Table 3.2) was evaluated. Based on previous results, KOD polymerase appears to be more efficient than Gotaq polymerase for the incorporation of modified triphosphates. Taking into account the destabilising effects of modified nucleotides, PCR with KOD polymerase and template T11 was carried out using 2 mM $MgCl_2$ instead of 1 mM (condition used in Chapter 2, Section 2.4) since Mg^{2+} helps stabilise the DNA duplex.[187] The PCR amplification efficiency, as judged by the intensity of the product bands, was sensitive to the ratio of TCO dUTPs in the reaction. PCR amplicons with 50% of TCO-S dU and 25% of TCO-L dU gave intense product bands

(Figure 3.12.A, lanes PS2 and PL1). As observed with the template T8 amplicons, the DNA duplex stability was affected less by the TCO-S dUTP than by the TCO-L dUTP (Figure 3.12.B). The poor PCR amplification is probably caused by the occurrence of multiple consecutive sites of incorporation (>6) at several loci along the template (oligo dA), and/or the duplex-destabilising effect of the TCO-L modification.

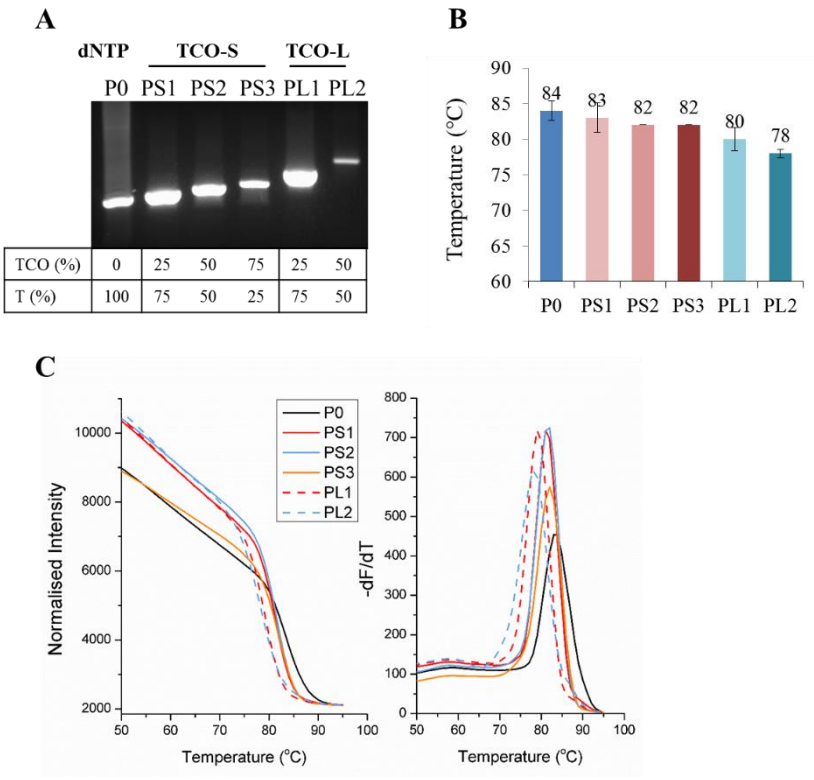


Figure 3.12. **A.** PCR amplifications using KOD polymerase and template T11 with different ratios of TCO dUTP to dTTP (10 nmol in total of TCO dUTP/dTTP + 10 nmol each of dATP, dCTP and dGTP). Lane P0, unmodified amplicon; lanes PS1 to PS3, 25% to 75% TCO-S-modified amplicons; lanes PL1 to PL2, 25% to 50% TCO-L-modified amplicons. (1.5% agarose gel containing ethidium bromide) **B.** Melting temperatures of TCO-modified PCR amplicons. **C.** Fluorescence melting curves and derivatives of TCO-modified PCR amplicons.

3.4.1 Fluorescent labelling of *trans*-cyclooctene-modified PCR products

Although it may be possible to optimise the PCR condition for the incorporation of higher percentages of TCO dUTPs, the use of lower TCO dUTP:dTTP ratios in the reaction mixture

might be a better option. With this strategy, the TCO dU modifications will be less densely distributed in the PCR products, thereby providing less steric hindrance for the subsequent labelling reaction. Also, if labelling efficiency were to remain high, lowering the density of TCO will decrease the quenching effect due to closely grouped Cy3 fluorophores (Figure 3.13). It has been reported that RNA transcription products from aminoallyl uridine triphosphate mixed with canonical uridine triphosphate (1:1 ratio) afforded better fluorescent post-labelling than more densely modified products.[71]

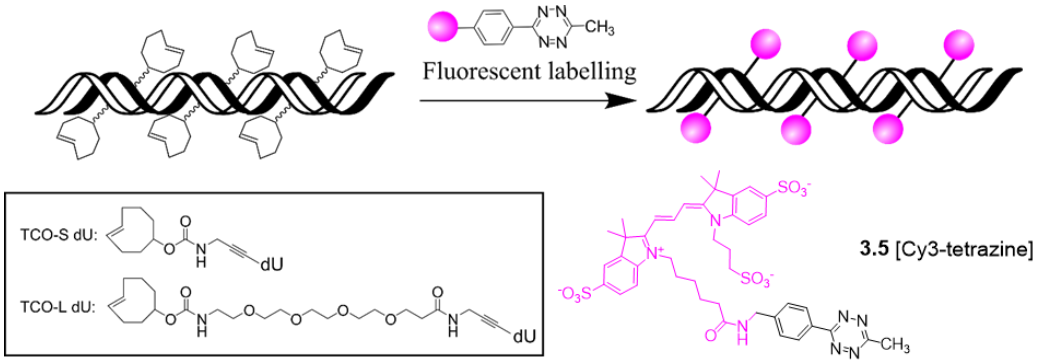
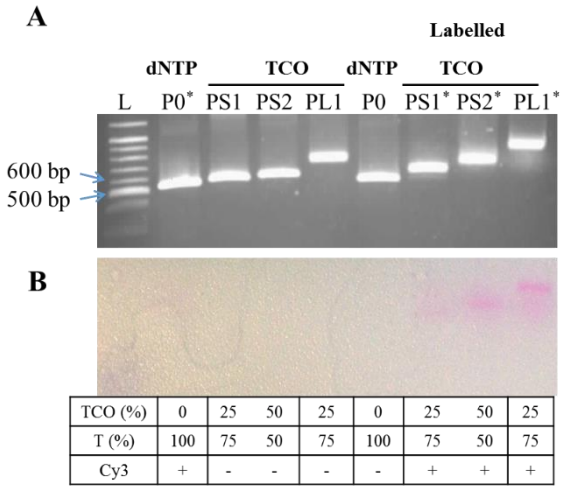


Figure 3.13. Fluorescent labelling of TCO-modified PCR products.

Fluorescent labelling was carried out on the T11-templated PCR products (523 bp). The labelled products migrated slower in proportion to increasing numbers of TCO dU modifications in the PCR amplicons (Figure 3.14.A, lanes PS1* and PS2*). After the labelling reaction, Cy3-labelled bands were visible by eye (Figure 3.14.B). The Cy3-labelled 25% TCO-L dU product (lane PL1*) afforded a more intense and sharper pink band than the 50% TCO-S dU product (lane PS2*), suggesting that the TCO-L with a longer linker was labelled more efficiently than TCO-S. Overall these results confirm that TCO-modified oligonucleotides can be used in DNA fluorescent detection and functionalisation.



TCO (%)	0	25	50	25	0	25	50	25
T (%)	100	75	50	75	100	75	50	75
Cy3	+	-	-	-	-	+	+	+

Figure 3.14. Fluorescent labelling of TCO-modified PCR amplicons from template T11. **A.** Gel was visualised by staining with ethidium bromide; **B.** Gel picture was taken under normal light. Lanes without asterisk are for unlabelled products, while lanes with asterisk are for products labelled with Cy3-tetrazine at RT for 30 min. Lane L, 100 bp ladder; lane P0/P0*, unmodified amplicon; lanes PS1/PS1* to PS2/PS2*, 25% to 50% TCO-S-modified amplicons; lane PL1/PL1*, 25% TCO-L-modified amplicons. (1.5% agarose gel)

3.5 Reverse transcription using *trans*-cyclooctene dUTPs

It is very important for applications involving RNA templates to develop methods to produce fluorescently labelled DNA probes from RNA. This can be done by reverse transcription. Moreover, the RNA strand of the modified DNA/RNA duplex can be digested by RNase enzymes to prepare functionalised single-stranded DNA.[204, 205] In this context, *trans*-cyclooctene modified dUTPs were investigated for incorporation against RNA templates (Figure 3.15) using Moloney Murine Leukemia Virus reverse transcriptase (M-MuLV, RNase H^-) and Avian Myeloblastosis Virus reverse transcriptase (AMV). An RNase inhibitor was added to AMV reactions, as the enzyme has RNase activity. Two RNA templates T4 and T5 with sequences equivalent to these used in primer extension were tested (Figure 3.16 and Figure 3.17).

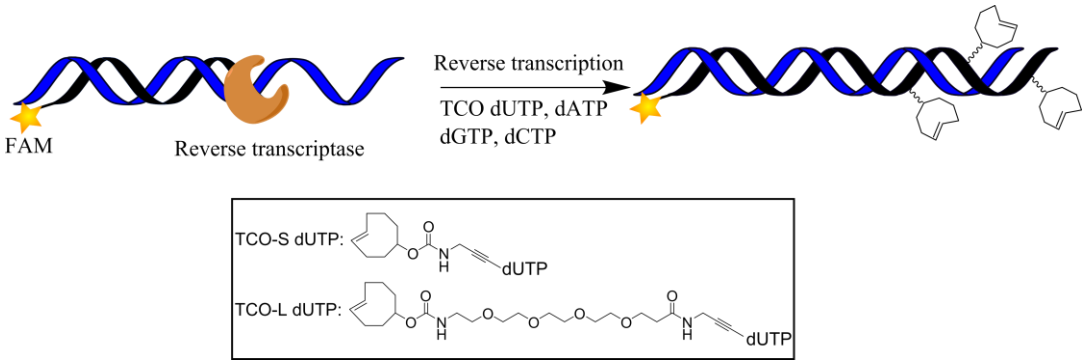


Figure 3.15. Reverse transcription using TCO dUTPs and a RNA template.

Fully extended complementary DNA products were obtained from reactions with TCO dUTP, dATP, dCTP and dGTP in the presence of the RNA template T4 as confirmed by denaturing PAGE and mass spectrometry analyses (Figure 3.16 and Table 3.1). When using AMV reverse transcriptase more intense unextended primer bands were observed for the TCO-L than the TCO-S dUTP (Figure 3.16.A), suggesting that TCO-S dUTP is incorporated more efficiently. M-MuLV (RNase H^-) produced higher conversion yields of elongated products compared to AMV reverse transcriptase (Figure 3.16). This is consistent with previous reports by Mahler and co-workers that M-MuLV (RNase H^-) reverse transcriptase gave better amplification than AMV when using natural triphosphates.[206]

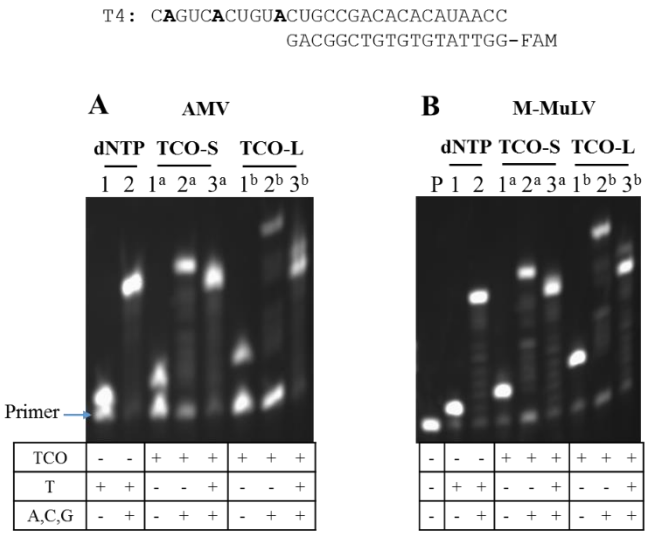


Figure 3.16. Reverse transcription using TCO dUTPs, primer P3 and template T4 with AMV (**A**, 10 units) and M-MuLV (RNase H^-, **B**, 100 units) reverse transcriptases for 15 h. 3.2 nmol of each triphosphate was used unless otherwise stated. Lane P, primer P3; lane $1/1^a/1^b$, dTTP or modified dUTP; lane $2/2^a/2^b$, dTTP or modified dUTP + dATP + dCTP + dGTP; lane $3^a/3^b$, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP. (20% denaturing PAGE)

In the case of the more demanding template T5, which has four consecutive incorporation sites (A_4) in the extension region, the TCO-S dUTP was a better substrate for both reverse transcriptases (Figure 3.17.A and B). Fully extended TCO-S dUTP products (lane 2^a) were consistently obtained in higher conversion yields than for TCO-L dUTP (lane 2^b). For TCO-L dUTP reactions, low levels of fully extended primers and multiple short strands were observed for both reverse transcriptases; more intense unextended primer bands remained for AMV reactions than for M-MuLV (RNase H^-) reactions (Figure 3.17). By using a 1:1 ratio of TCO-L dUTP:dTTP, longer extended products were observed as multiple bands (Figure 3.17, lane 3^b). Generally, all modified dUTPs were better substrates for M-MuLV (RNase H^-) than AMV, and TCO-S dUTP is a better substrate than TCO-L dUTP (Figure 3.16 and Figure 3.17).

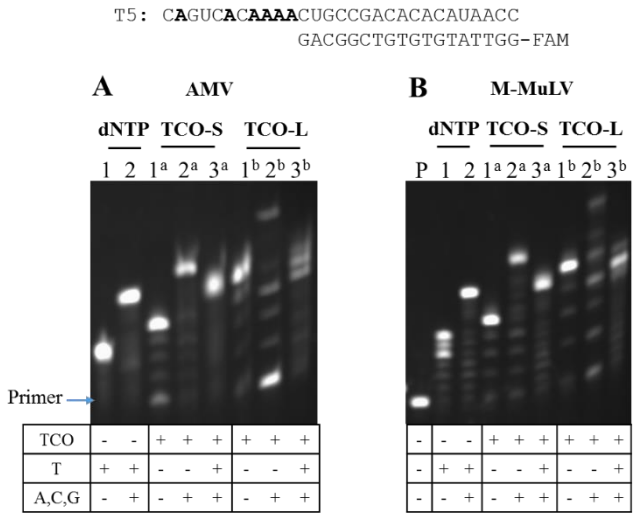


Figure 3.17. Reverse transcription using TCO dUTPs, primer P3 and template T5 with AMV (**A**, 10 units) and M-MuLV (RNase H^-, **B**, 100 units) reverse transcriptases for 15 h. 3.2 nmol of each triphosphate was used unless otherwise stated. Lane P, primer P3; lane $1/1^a/1^b$, dTTP or modified dUTP; lane $2/2^a/2^b$, dTTP or modified dUTP + dATP + dCTP + dGTP; lane $3^a/3^b$, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP. (20% denaturing PAGE)

3.6 Conclusions

Two *trans*-cyclooctene dUTPs have been synthesised efficiently by coupling their NHS carbonates/esters to amino dUTP. They were successfully incorporated into DNA by primer extension, reverse-transcription and PCR amplification. The TCO-S dUTP with a short linker between the TCO and uridine base was incorporated more efficiently than the longer TCO-L dUTP version using Klenow polymerase, KOD polymerase, M-MuLV (RNase H-) and AMV reverse transcriptases. For TCO-L dUTP reactions, primer extension through consecutive incorporation sites was only achieved using Gotaq polymerase. Moreover, the TCO-S dU stabilised the DNA duplex better than the TCO-L dU.

Cy3-tetrazine was utilised to label *trans*-cyclooctene-modified primer extension and PCR products *via* the IEDDA reaction, which can be used to prepare fluorescent probes for DNA sequence detection and to make DNA conjugates or nanostructures.[139, 158] The IEDDA click

labelling on the TCO-L-modified DNA was more efficient than labelling of the TCO-S dU products due to less steric hindrance from the oligonucleotide chain during the labelling reaction.

Independently, the Wang group also reported that TCO-S dUTP can be incorporated by different polymerases using templates with evenly distributed adenine sites in the extension region; PCR amplicon from a 90-mer template was fluorescently labelled with fluorescein isothiocynanate-tetrazine.[139] They applied this strategy to functionalise an aptamer for cell labelling studies. Their results are consistent with the findings for TCO-S dUTP in this Chapter. Compared to Wang's report, our assays include a longer TCO-linker, a more challenging template with consecutive incorporations for primer extension, a long 523 bp plasmid template and positive results from reverse transcription studies. The results in this chapter should enable a more general route to the synthesis of TCO-modified DNA. Furthermore, the efficient incorporation of TCO-S dUTP by Klenow polymerase suggests that this strategy for labelling TCO-S-modified DNA might be used to functionalise and image cellular (genomic) DNA (discussed in Chapter 6).

Chapter 4 Synthesis and applications of azide-modified deoxyuridine triphosphates

4 Synthesis and applications of azide-modified deoxyuridine triphosphates

4.1 Background and aims

There are many applications that require fluorescently labelled DNA such as fluorescence *in situ* hybridisation (FISH).[32, 39, 146] Standard FISH probes can produce weak signals when the genomic DNA or RNA target concentration is insufficient, but can be significantly improved by introducing multiple fluorophores into the probe.[207] However, to prepare short multiple labelled fluorescent probes, solid phase oligonucleotide synthesis is inefficient as the coupling of bulky fluorophore-modified phosphoramidites requires longer reaction times and the yield is lower than for unmodified monomers. This effect is amplified if multiple modifications are introduced. For FISH, enzymatic DNA probe synthesis is a viable alternative to solid phase synthesis and has certain advantages; its efficiency is higher for the synthesis of long probes, and it is suitable for the introduction of large numbers of labels. In enzymatic DNA synthesis two-step labelling processes are usually necessary because direct enzyme-mediated incorporation of fluorophore-labelled nucleoside triphosphates is problematic due to the bulky nature of fluorophores (Chapter 1, Section 1.3.2).[67] For this post-labelling strategy, the strain-promoted alkyne-azide cycloaddition reaction (SPAAC, Chapter 1, Section 1.4.3) and the inverse electron demand Diels-Alder reaction (IEDDA, Chapter1, Section 1.4.4) are efficient methodologies and are suitable for *in vivo* and *in vitro* biological applications.[68, 116, 208]

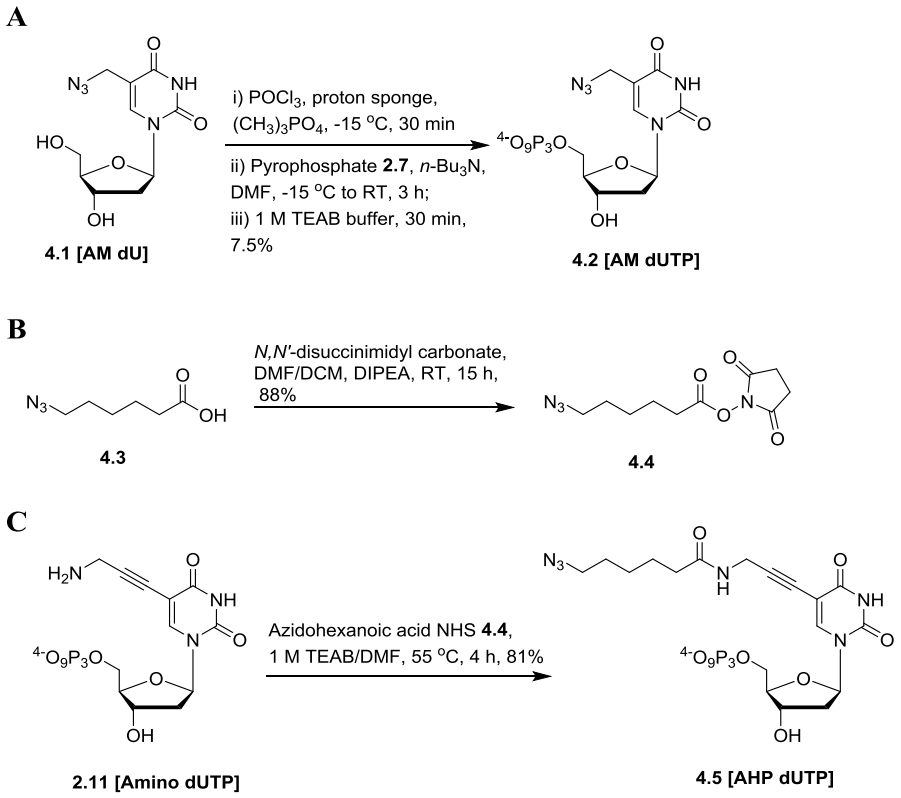
As described in Chapter 2 and Chapter 3, the initial approach was to incorporate cyclooctyne or *trans*-cyclooctene modified triphosphates into DNA and subsequently

label them with an azide or tetrazine derivative of the fluorophore. However, bulky cyclooctyne and *trans*-cyclooctene-modified nucleoside triphosphates are not ideal substrates for DNA polymerases and reverse transcriptases. Alternatively, a variety of sugar- and base-modified azide nucleotides have been enzymatically incorporated into DNA or RNA to introduce fluorophores *via* the CuAAC, SPAAC or Staudinger reactions (Chapter 1, Section 1.4.2).[68, 110-112, 114, 209] Yet an optimised DNA polymerase-compatible azide triphosphate has not been reported. Therefore, the aim of this chapter is to study the enzymatic incorporation efficiency of two azide-labelled deoxyuridine triphosphates (azidomethyl dUTP and azidohexanamidopropargyl dUTP, Scheme 4.1), and to investigate the fluorescent labelling of these azide-modified nucleic acids *via* the SPAAC reaction. Since a single-colour fluorescent probe can only detect one gene of interest at a time, cocktails of probes labelled with different fluorescent dyes are usually used to identify multiple targets.[34, 210] However the number of dyes suitable for nucleic acid detection that do not suffer from spectral overlap is limited. Therefore, the second goal of this chapter is to develop a strategy whereby different targets could be detected using probes that contain varying ratios of two fluorophores to give different coloured probes.

4.2 Preparation of azide-modified dUTPs

5-Azidomethyl-2′-deoxyuridine **4.1** (AM dU, Chapter 1, Section 1.5) can be transformed from its nucleoside to an active triphosphate in proliferating cells and incorporated into replicating DNA.[157] However, the *in vitro* enzymatic incorporation of its triphosphate and the effect of multiple additions of this modification on DNA duplex stability have not been reported. 5-Azidomethyl-2′-deoxyuridine-5′-triphosphate **4.2** (AM dUTP, Scheme 4.1.A) was therefore synthesised from AM dU **4.1** (which was synthesised by Dr Marta Gerowska in the Brown group) using the Yoshikawa reaction.[163] Phosphorus oxychloride

($POCl_3$) in trimethylphosphate was employed to generate the 5′-dichlorophosphate derivative of the nucleoside (Chapter 2, Section 2.2). Howorka and co-workers examined the regioselectivity and yield of this phosphorylation step at 0 °C and found that the yield of the 5′-monophosphate and the by-product (the 3′-monophosphate) remained constant in the first 30 min. However, the other main by-product (the 5′,3′-diphosphate) increased from 15% (5 min) to 30% (30 min).[168] This suggests that a long reaction time (e.g. -15 ºC for 2 h used in Chapter 2, Section 2.2) would not favour the production of the target 5′-monophosphate. Therefore, the first step (to form the active monophosphate) was carried out at -15 ºC for 30 min to minimise the production of the 5′,3′-diphosphate by-product. The reaction mixture was then treated with tris(tetrabutylammonium) hydrogen pyrophosphate **2.7** to afford the target triphosphate **4.2** in a yield of 7.5%.



Scheme 4.1. Synthesis of AM dUTP (**A**), 6-azidohexanoic acid NHS ester (**B**) and AHP dUTP (**C**).

In Chapter 2 and Chapter 3, 5-aminopropargyl-2′-deoxyuridine-5′-triphosphate **2.11** (amino dUTP) was found to be a good intermediate in enzymatic nucleic acid synthesis to introduce bulky functional groups (cyclooctyne and *trans*-cyclooctene) into DNA.[59, 65, 66, 168, 198] Therefore, it was also used to prepare an azide triphosphate with an alkynyl linkage to the nucleobase. 6-Azidohexanoic acid *N*-hydroxysuccinimide (NHS) ester **4.4** was synthesised from its carboxylic acid **4.3** using a reported procedure (Scheme 4.1.B).[211] 5-Azidohexanamidopropargyl-2′-deoxyuridine-5′-triphosphate **4.5** (AHP dUTP), which has a longer linker than AM dUTP, was synthesised by labelling amino dUTP **2.11** with 6-azidohexanoic acid NHS ester **4.4** (Scheme 4.1.C). It should be noted that a modified azido-7-deaza-2′-deoxyadenosine triphosphate with a similar alkynyl linker was reported by Weisbrod and Marx.[111] This dATP derivative was successfully incorporated into a single position of a DNA strand by *Pyrococcus woesei* DNA polymerase and subsequently labelled with phosphine-tagged reporter groups *via* the Staudinger reaction. Therefore, it was anticipated that AHP dUTP **4.5** would also be a good substrate for a variety of polymerases.

4.3 Primer extension using azide dUTPs

Enzymatic DNA synthesis from these two azide-modified triphosphates was examined using the same primer extension assays as described in Chapter 2 and Chapter 3 (Figure 4.1). For ease of comparison, the same DNA templates (T1 and T2) and DNA polymerases (Klenow, Gotaq and KOD) were used. The template and primer sequences are shown in Figure 4.2 along with denaturing polyacrylamide gel electrophoresis (PAGE) analyses. Template T1 carries three evenly distributed adenines in the extension region, whilst template T2 contains six adenine bases with one site containing four consecutive adenines. The primer can be extended by one (T1) or four (T2) bases in the presence of

only azide dUTP or natural dTTP to give a partially extended product. In the presence of all four triphosphates, eleven nucleotides would be incorporated to produce a fully extended product. A mix of all four natural dNTPs was used as a positive control. To visualise the extended products by analytical PAGE analysis, 6-carboxyfluorescein (FAM) was attached to the 5′-end of the primer.

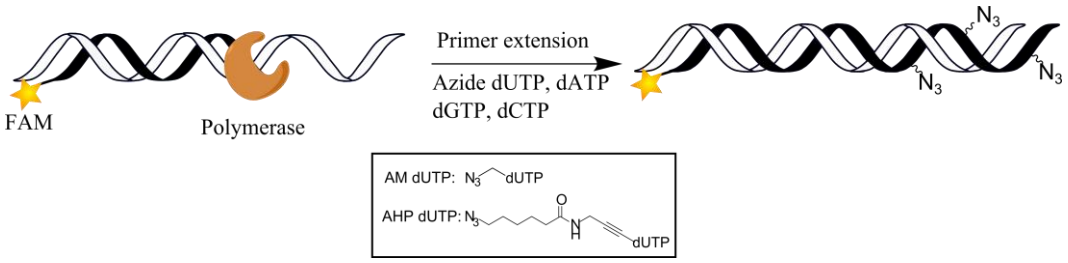


Figure 4.1. Primer extension using a DNA template and azide dUTPs.

4.3.1 Primer extension using Klenow polymerase

Primer extension with AM and AHP dUTPs was carried out using Klenow polymerase, which is a natural family A polymerase and functions most efficiently at 37 °C. For the template T1, reactions with dTTP or modified dUTP produced one- to two-base extended products due to nucleotide misincorporation *via* untemplated linear extension reaction (Figure 4.2.A, lanes 1, 1[a] and 1[b]).[134, 159, 177] The same misincorporation was observed in the control reaction with dATP, dCTP and dGTP (Figure 4.2, lane N). The resultant 3′-mismatched complex could not be extended further, and two closely running bands were observed on denaturing PAGE. Digested products were also observed for these partial extension reactions due to the 3′-5′ exonuclease activity of Klenow polymerase.

For templates T1 and T2, full-length products were obtained in the presence of modified dUTP combined with dATP, dCTP and dGTP (*i.e.* no dTTP, Figure 4.2 and Table 4.1). The electrophoretic mobility of AM dU products was similar to that of the unmodified

congener, whereas the AHP dU modified DNA strands showed slower mobility (Figure 4.2, lanes 2, 2^a and 2^b). To investigate the competitive incorporation of the modified and natural triphosphates, the primer extension reactions were repeated with mixtures of azide dUTP and natural dTTP (1:1) in addition to dATP, dCTP and dGTP. Broader product bands from the AHP dUTP and dTTP mixture were observed (Figure 4.2, lane 3^b), due to the incorporation of a mixture of AHP dUTP and dTTP. This phenomenon was not clearly observed when AM dUTP was used due to the minor difference of electrophoretic mobility between the structurally similar AM dU and dT.

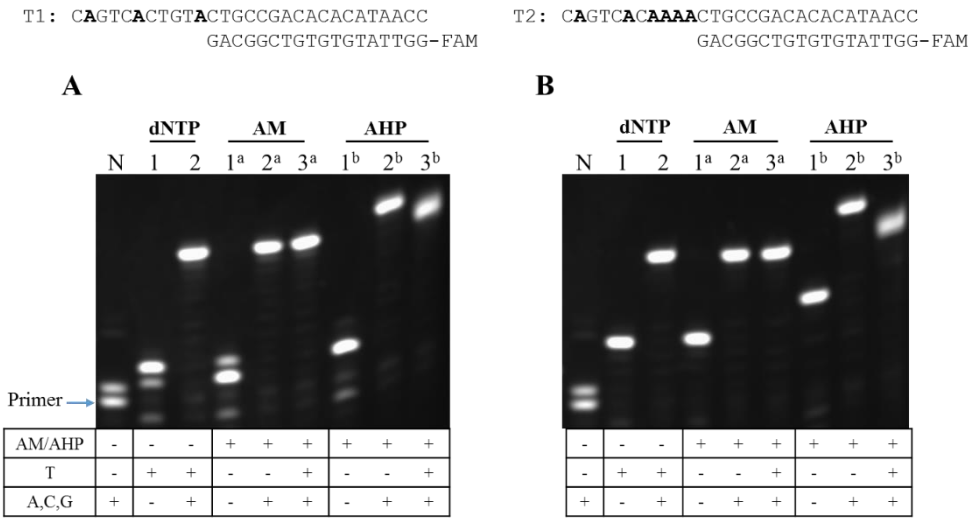


Figure 4.2. Primer extension using azide dUTPs, primer P3, template T1 (**A**) or T2 (**B**) for Klenow polymerase (37 ºC, 1 h). 3.2 nmol of each triphosphate was used unless otherwise stated. Lane N, reactions with dATP + dCTP + dGTP; lane $1/1^a/1^b$, dTTP or modified dUTP; lane $2/2^a/2^b$, dTTP or modified dUTP + dATP + dCTP + dGTP; lane $3^a/3^b$, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP. (20% denaturing PAGE)

Table 4.1. Mass spectrometry analysis of primer extension and reverse transcription reactions using azide dUTPs. A, C, G = dATP, dCTP, dGTP. For representative mass spectra (HPLC-MS) see Appendix 8.4.

	Template T1 or T4		**Template T2 or T5**	
	dUTP Found (Calc.)	dUTP+A+C+G Found (Calc.)	dUTP Found (Calc.)	dUTP+A+C+G Found (Calc.)
AM	6483(6483)^{G,K,M}	9679(9679)^{all}	7519(7519)^{K,G,D}	9799(9799)^{K,G,D}
AHP	6621(6620)^{all}	10090(10090)^{all}	8067(8067)^{all}	10621(10621)^{all}

Note: The superscripts indicate that the mass was confirmed for the specific polymerase used. All = all four polymerases. G = Gotaq, K = Klenow, D = KOD, M = M-MuLV (RNase H^-).

4.3.2 Primer extension using Gotaq polymerase

A thermostable family A polymerase (Gotaq) was also examined in the primer extension assay with AM and AHP dUTPs. Gotaq polymerase functions efficiently at higher temperatures (75 °C) and is commonly used in the polymerase chain reaction (PCR). The initial primer extension with Gotaq polymerase was carried out at 72 °C, close to its optimum temperature. Denaturing PAGE analysis demonstrated that azide-modified full-length products were obtained for template T1 as for the previous Klenow reactions (Figure 4.3, confirmed by MS). For template T2, in the presence of AM dUTP in combination with dATP, dCTP and dGTP, a shorter product than the unmodified full-length product was obtained, whereas for AHP dUTP reactions, only two-base extended products were observed (Figure 4.3.B, lanes 2, 2^a and 2^b). These results indicate that AM dUTP is incorporated more efficiently by Gotaq polymerase than AHP dUTP. Promisingly, in the presence of a 1:1 mixture of modified dUTP and dTTP, fully extended product bands were observed, showing that mixing modified dUTP with natural dTTP prevents extension termination, as previously reported.[67] To identify the optimum time required for full extension with AHP dUTP, a time course experiment was conducted

with both templates. For template T1, the reaction was complete within 4 h, while for template T2 overnight incubation gave the best conversion yield (Figure 4.4).

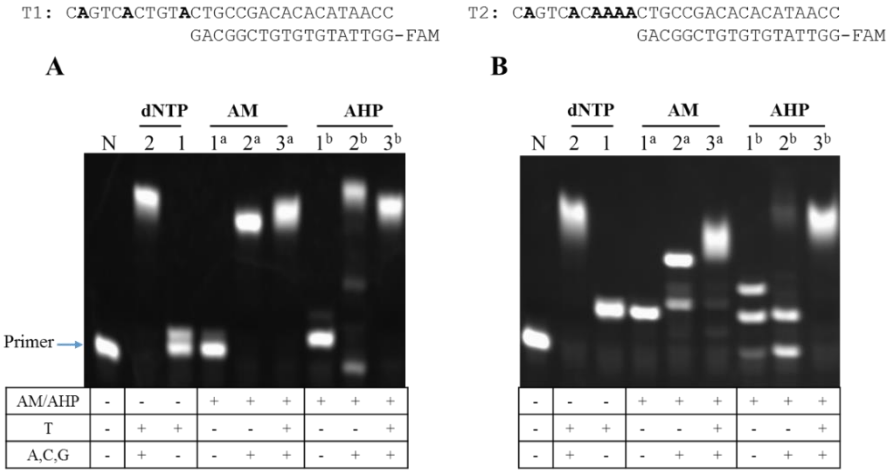


Figure 4.3. Primer extension using azide dUTPs, primer P3, template T1 (**A**) or T2 (**B**) for Gotaq polymerase (72 °C, 1 h). 3.2 nmol of each triphosphate was used unless otherwise stated. Lane N, reaction without triphosphates; lane $1/1^a/1^b$, dTTP or modified dUTP; lane $2/2^a/2^b$, dTTP or modified dUTP + dATP + dCTP + dGTP; lane $3^a/3^b$, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP. (20% denaturing PAGE)

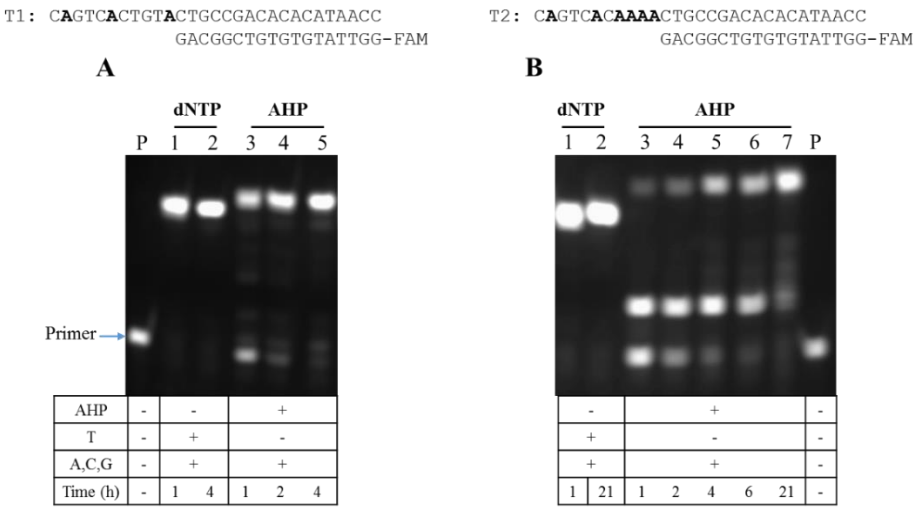


Figure 4.4. Time course of primer extension using AHP dUTP and Gotaq polymerase (72 °C). **A.** Reactions using primer P3 and template T1. Lane P, primer P3; lanes 1 to 2, dNTPs (1 and 4 h); lanes 3 to 5, AHP dUTP + dATP + dCTP + dGTP (1, 2 and 4 h, respectively). **B.** Reactions using primer P3 and template T2. Lane P, primer P3; lanes 1 to 2, dNTPs (1 and 21 h); lanes 3 to 7, AHP dUTP + dATP + dCTP + dGTP (1, 2, 4, 6 and 21 h, respectively).

The thermal stability of AM- and AHP-modified DNA duplexes was then studied by monitoring the melting temperatures of Gotaq primer extension products in the presence of SYBR Green (Figure 4.5). The T_m of partially and fully extended products were in the range of 62 °C to 70 °C in 1 × Gotaq green buffer. This suggests that the duplex products were partially denatured under the primer extension condition (72 °C), which might affect the incorporation efficiency of modified triphosphates. The melting temperature analyses also showed that DNA containing the AHP modification has a similar or slightly better stability to the unmodified DNA due to the extra π-stacking of the alkyne linkage on the surrounding bases, whereas the AM modification destabilises the DNA duplex on average by 0.5 °C per incorporation (Figure 4.5). By mixing AM dUTP with dTTP to reduce the number of modified nucleotides in the DNA strand, the T_m increased by 1 °C for the template T1 product and by 3 °C for the template T2 product compared to the fully modified products (Figure 4.5, mix). Consistent with the above observations, Sagi and co-workers reported the destabilising effect of a methyl group and the stabilising effect of the 1-propynyl group when attached to the 5-position of pyrimidine bases.[169] These observations were later confirmed by the thermal stability of PCR products (Section 4.4).

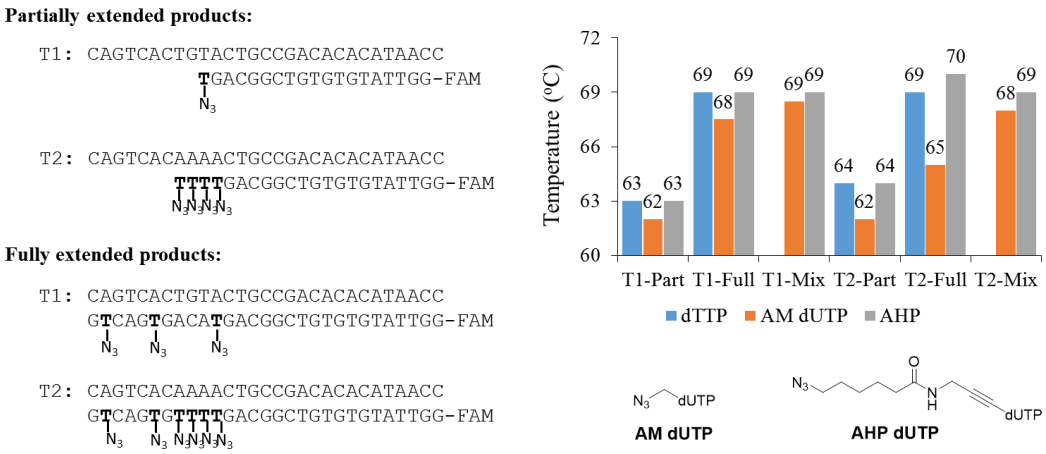


Figure 4.5. Melting temperatures of azide-modified primer extension products from template T1 and T2 using Gotaq polymerase. Part, partially extended products; full, fully extended products; mix, reactions with dTTP and azide dUTP (1:1) + dATP + dCTP + dGTP. Melting curves see Appendix 8.5.2.

As a consequence, Gotaq primer extension reactions were repeated at 60 °C for both template T1 and T2 to ensure the formation of the primer-template DNA duplex. Denaturing PAGE analyses showed that azide dU full-length products from both templates were obtained in 1.5 h (Figure 4.6, lanes 2^a and 2^b). Even with reduced polymerase activity at 60 °C, AM and AHP dUTPs were incorporated more efficiently than at 72 °C (Figure 4.3, lanes 2^a and 2^b). This proves that forming a stable primer-template duplex is important for Gotaq polymerase. This simple observation is easy to overlook in primer extension assays. The expected mass peaks of the partially and fully extended products of AM and AHP dUTPs were observed by analysing the reaction mixtures (Table 4.1); an extra addition of dA at the end of the full-length product was detected, which is a known feature of Gotaq polymerase.

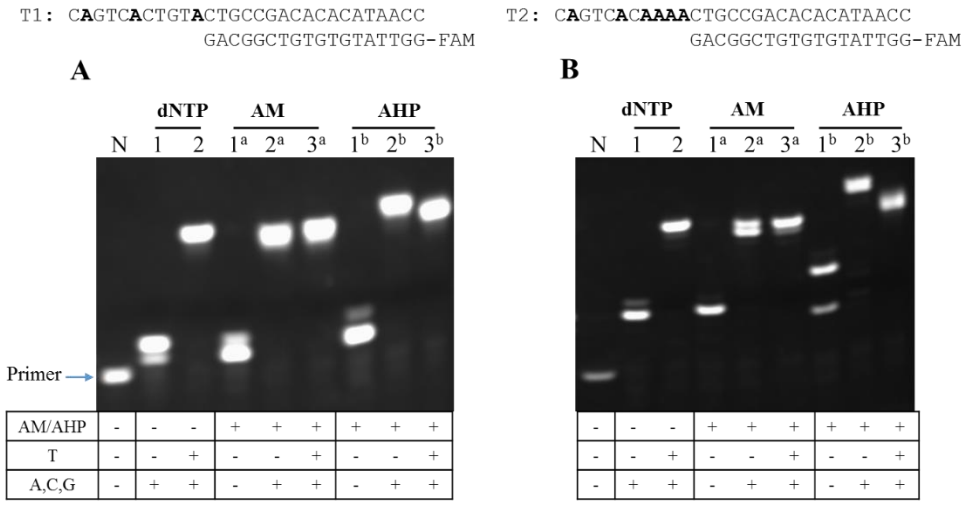


Figure 4.6. Primer extension using azide dUTPs, primer P3, template T1 (**A**) or T2 (**B**) for Gotaq polymerase (60 °C, 1 h). 3.2 nmol of each triphosphate was used unless otherwise stated. Lane N, reaction without triphosphates; lane 1/1^{a}/1^{b}, dTTP or modified dUTP; lane 2/2^{a}/2^{b}, dTTP or modified dUTP + dATP + dCTP + dGTP; lane 3^{a}/3^{b}, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP. (20% denaturing PAGE)

4.3.3 Primer extension using KOD polymerase

KOD polymerase is an engineered family B polymerase, designed to incorporate modified triphosphates more efficiently than natural polymerases are able to. The initial primer extension with azide dUTPs was carried out at 72 °C for 1.5 h. Full-length products with AM/AHP dU modifications were successfully obtained (Figure 4.7). However, multiple digested primer bands were present for the reaction with dTTP or azide dUTP only, due to the strong 3′-5′ exonuclease activity of KOD polymerase (Figure 4.7, lanes 1, 1^{a} and 1^{b}).

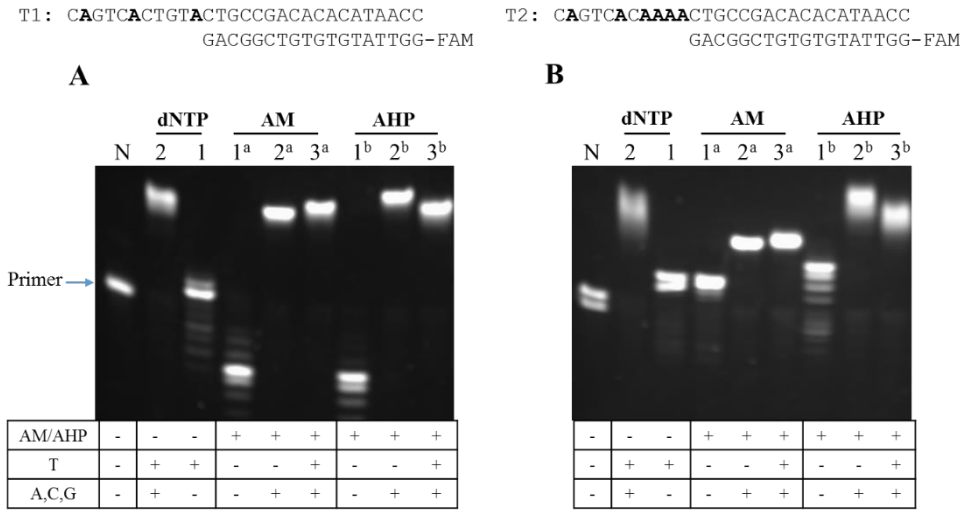


Figure 4.7. Primer extension using azide dUTPs, primer P3, template T1 (**A**) or T2 (**B**) for KOD polymerase (72 °C, 1.5 h). 3.2 nmol of each triphosphate was used unless otherwise stated. Lane N, reaction without triphosphates; lane $1/1^a/1^b$, dTTP or modified dUTP; lane $2/2^a/2^b$, dTTP or modified dUTP + dATP + dCTP + dGTP; lane $3^a/3^b$, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP. (20% denaturing PAGE)

To minimise the exonuclease activity, optimisation of the primer extension reaction conditions for KOD polymerase was carried out. (Figure 4.8). It was envisaged that decreasing the reaction temperature from 72 °C to 60 °C would decrease exonuclease activity; however, this might also adversely affect incorporation activity. Therefore, the concentration of $MgCl_2$ was increased to 2 mM to compensate. In addition to these changes, the reaction time for modified dUTP or dTTP reactions was decreased to 30 min. Partially extended products with fewer digested by-products were observed (Figure 4.8, lanes 1, 1^a and 1^b). Fully extended products were observed as major products for modified dUTP reactions containing dATP, dCTP and dGTP after 1.5 h (Figure 4.8, lanes 2, 2^a and 2^b).

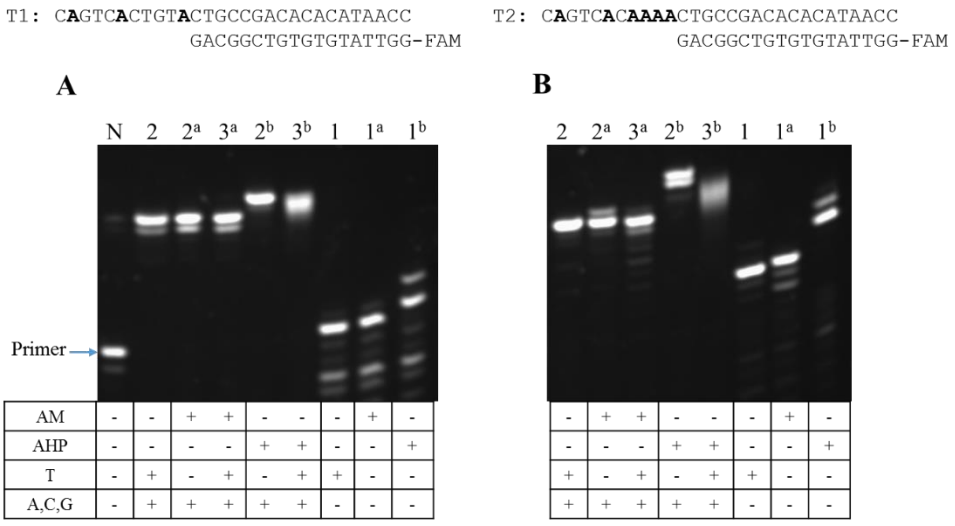


Figure 4.8. Primer extension using azide dUTPs, primer P3, template T1 (**A**) or T2 (**B**) for KOD polymerase (60 °C). 3.2 nmol of each triphosphate was used unless otherwise stated. Lane N, reaction without triphosphates; lane 1/1^{a}/1^{b}, dTTP or modified dUTP (30 min); lane 2/2^{a}/2^{b}, dTTP or modified dUTP + dATP + dCTP + dGTP (1.5 h); lane 3^{a}/3^{b}, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP (1.5 h). (20% denaturing PAGE)

4.3.4 Kinetics analysis of AHP dUTP single nucleotide incorporation

Steady state kinetics for single nucleotide incorporation were determined for Gotaq and KOD polymerases using template T1 and AHP dUTP (Figure 4.9.A), which is a better monomer than the AM modification for duplex stabilisation. The kinetic measurements were carried out using the procedure described in Chapter 2 (Section 2.3.5). The data were then fitted to the Michaelis-Menten equation to calculate the Michaelis constant (K_M) and maximal reaction rate (V_{max}) (Figure 4.9.B). The supplier-defined unit of enzyme concentration was used to calculate the catalytic constant k_{cat} (Table 4.2). The relative catalytic efficiency k_{cat}/K_M indicates the enzymatic incorporation efficiency of a triphosphate by a polymerase. For KOD polymerase, the k_{cat}/K_M of AHP dUTP (0.82) was similar to dTTP (1.00), suggesting that AHP dUTP was almost as good a substrate as dTTP. By comparison, for Gotaq polymerase the k_{cat}/K_M value of AHP dUTP

incorporation was 0.07 relative to dTTP (1.00). The binding affinity between AHP dUTP (K_M = 197 ± 23 μM) and Gotaq was 3.2 times higher than that of the dTTP (K_M = 61 ± 10 μM, k_{cat} = 1.60 ± 0.16 % s^{-1}); the turnover reaction rate of AHP dUTP (k_{cat} = 0.36 ± 0.02 % s^{-1}) was 4.4 times slower. These results demonstrate that Gotaq is more selective for the unmodified nucleotide.

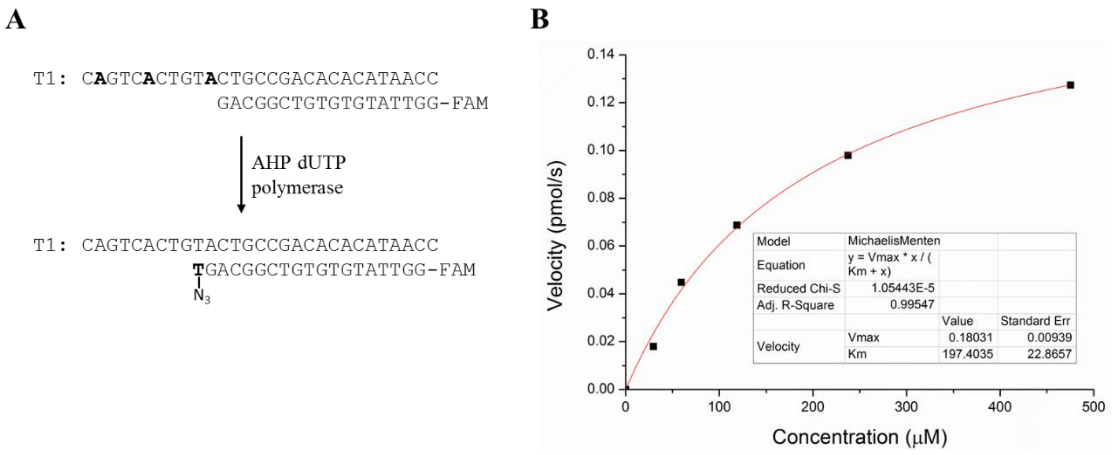


Figure 4.9. **A.** Single nucleotide incorporation during primer extension. **B.** Michaelis-Menten kinetics for AHP dUTP incorporation using Gotaq polymerase. For dTTP incorporation using Gotaq polymerase see Figure 2.15; For KOD polymerase results see Appendix 8.6.2.

Table 4.2. The kinetics of dTTP and AHP dUTP incorporation using Gotaq and KOD polymerases at 60 °C. The supplier defined polymerase concentration (unit) was used to calculate the k_{cat}. The k_{cat}/K_M values are expressed relative to dTTP.

Polymerase	dXTP	V_{max} pmol s^{-1}	K_M μM	k_{cat} pmol s^{-1} $unit^{-1}$	k_{cat}/K_M Relative values
Gotaq	dTTP	0.80 ± 0.05	61 ± 10	1.60 ± 0.10	1.00
	AHP dUTP	0.18 ± 0.01	197 ± 23	0.36 ± 0.02	0.07
KOD	dTTP	0.24 ± 0.01	1.4 ± 0.3	0.48 ± 0.02	1.00
	AHP dUTP	0.24 ± 0.01	1.7 ± 0.2	0.48 ± 0.02	0.82

4.3.5 Fluorescent labelling of azide-modified primer extension products

After AM and AHP dUTPs have been incorporated into DNA strands, the azide can be functionalised efficiently *via* the CuAAC, SPAAC or Staudinger reactions.[68, 110-112, 114] Here, the SPAAC reaction was used to label the azide-modified DNA products with bicyclo[6.1.0]nonyne (BCN) functionalised fluorophores (Figure 4.10). BCN has C2-symmetric alkyne and only produces two isomeric products, which simplifies the analysis of SPAAC products (Chapter 1, Section 1.4.3). Commercially available amino-triethylene glycol BCN **4.6** was treated separately with the NHS active esters of three fluorophores with mutually distinct fluorescent spectra, to prepare the corresponding novel BCN fluorophores (Cy3-BCN **4.7**, FAM-BCN **4.8**, Cy5-BCN **4.9**, Figure 4.11). The van Delft group and the Luedtke group have used this labelling strategy to prepare BCN-functionalised AlexaFluor fluorophores to image cellular glycans and DNA.[122, 157]

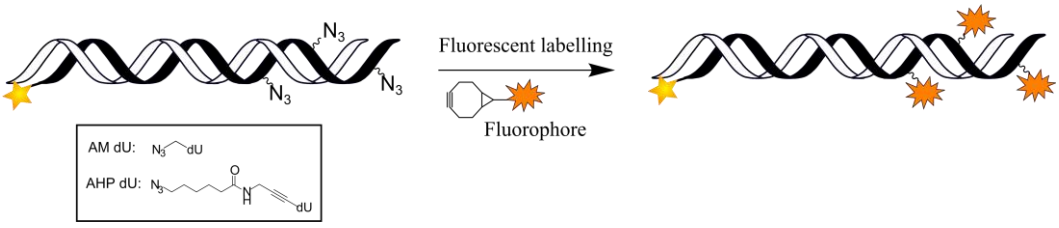


Figure 4.10. Fluorescent labelling of azide-functionalised primer extension products with BCN-modified fluorophores.

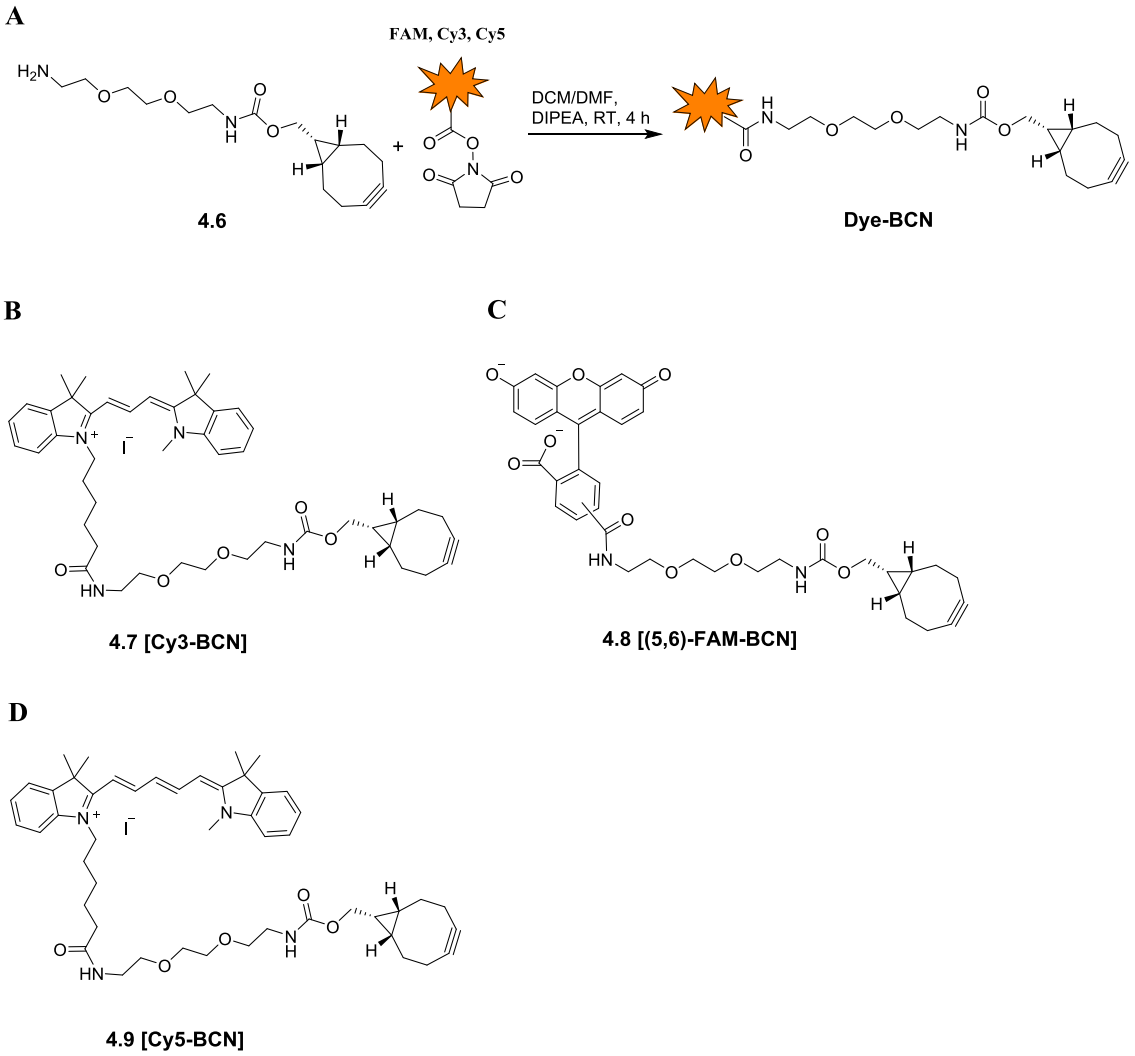


Figure 4.11. **A.** Synthesis protocol for BCN-functionalised fluorophores. **B.** Chemical structures of Cy3-BCN, 5(6)-carboxyfluorescein-BCN (for ease of description 5(6)-FAM-BCN is written as FAM-BCN subsequently) and Cy5-BCN.

Cy3-BCN was first used to label the azide-functionalised primer extension products. Labelling reactions were carried out on the partially and fully extended products from both templates (Figure 4.12). Extension products obtained from unmodified dNTPs were used as negative controls. Partially extended products of both AM and AHP dUTPs were labelled efficiently and no unlabelled starting materials were observed (Figure 4.12.B and D, lanes 2 and 4). Compared to the unlabelled primer extension products (Figure 4.12.A and C, lanes 2 and 4), the Cy3-BCN labelled strands exhibited much slower

electrophoretic mobility due to the bulky fluorophores. For the full-length products, the AHP-modified DNAs were labelled with more than one Cy3-BCN, while AM dU products remained mainly unlabelled (Figure 4.12.B and D, lanes 3 and 5). The products containing the AHP dU nucleotides exhibited much higher labelling efficiency than the AM dU products, especially when the azide groups were in the middle of DNA strands. This may be due to the longer linkage between the azide and the uracil base for the AHP moiety that protrudes further out from the major groove of the DNA duplex. Consequently, there is less steric hindrance for Cy3-BCN labelling. With this in mind, subsequent fluorescent labelling studies were only carried out on AHP-modified products.

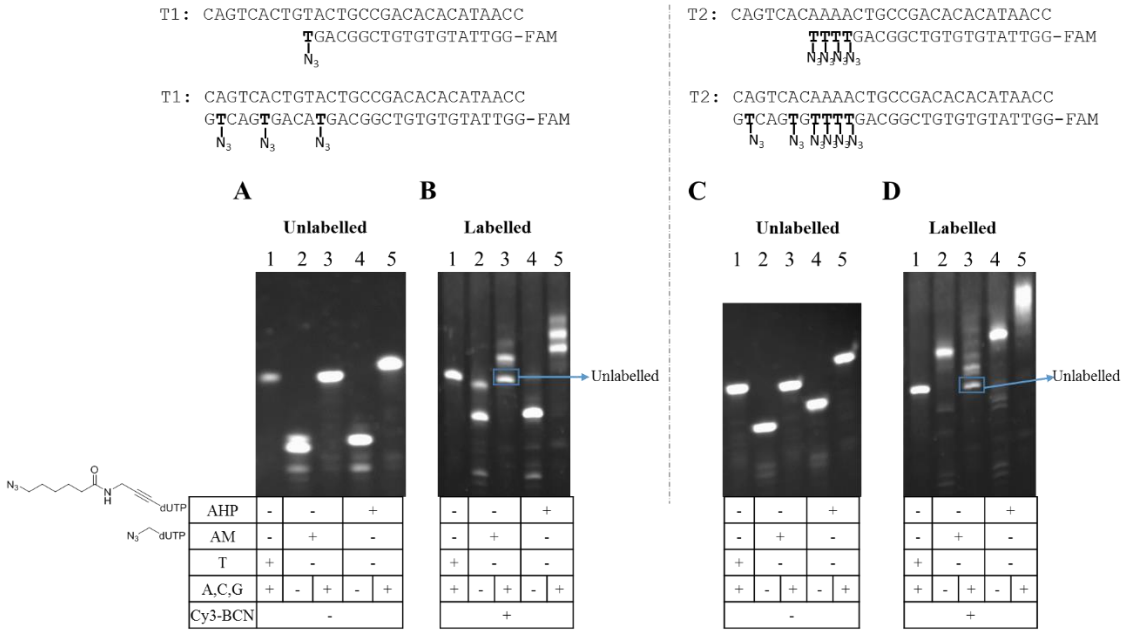


Figure 4.12. Fluorescent labelling of azide-modified primer extension products with Cy3-BCN at RT for 2 h. Unlabelled products from template T1 (**A**) and T2 (**C**). Cy3-BCN labelled products from template T1 (**B**) and T2 (**D**). Lane 1, unmodified product, negative control; lane 2, AM dU partially extended product; lane 3, AM dU full-length product; lane 4, AHP dU partially extended product; lane 5, AHP dU full-length product. (20% denaturing PAGE)

The labelling efficiency of different BCN-functionalised fluorophores (Cy3-BCN, FAM-BCN, Cy5-BCN) was then examined using AHP dUTP primer extension products from both template T1 and T2 (Figure 4.13 and Figure 4.14). Denaturing PAGE analyses showed that the AHP dU product was labelled with at least one Cy3-BCN within 1 h (Figure 4.13.A and Figure 4.14.A, lane 4). By comparison, intense unlabelled bands were observed for FAM-BCN and Cy5-BCN labelling reactions (Figure 4.13.A and Figure 4.14.A, lanes 5 and 6). The negative charge of the 5(6)-FAM moiety might decrease the reactivity of BCN toward the azide modifications in the anionic DNA major groove. However, an advantage of the FAM labelled products is that they exhibited better aqueous solubility than the cyanine labelled products. The hydrophobicity and bulky structure of Cy5 might lead to its lower reactivity in the labelling reaction. For all labelling reactions, increasing the reaction time to 4 h gave a higher ratio of densely (or fully) labelled products (Figure 4.13.B and Figure 4.14.B, lanes 4 to 6). Expected mass peaks of double FAM-BCN labelled products (calc. 11455, found 11457) and triple FAM-BCN labelled products (calc. 12137, found 12139) were observed.

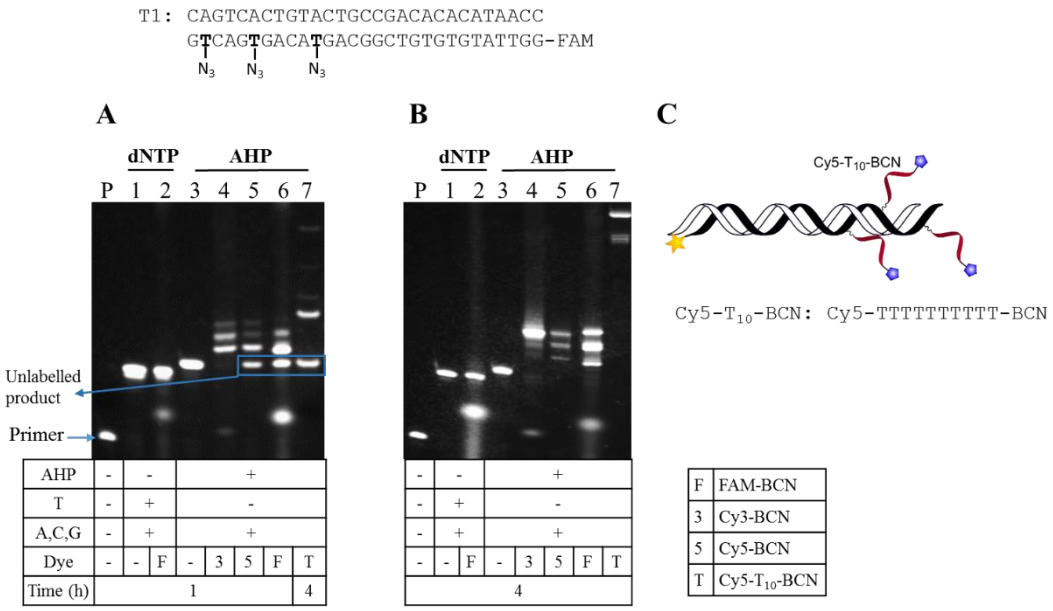


Figure 4.13. Fluorescent labelling of AHP dU full-length products from template T1. **A.** Reactions with Cy3-BCN, Cy5-BCN, FAM-BCN (5 eq) at RT for 1 h and Cy5-T_{10}-BCN (1 eq) at 55 °C for 4 h. **B.** Reactions with Cy3-BCN, Cy5-BCN, FAM-BCN (5 eq) at RT for 4 h and Cy5-T_{10}-BCN (2 eq) at 55 °C for 4 h. Lane P, primer P3; lane 1, unmodified product; lane 2, unmodified product mixed with FAM-BCN, negative control; lane 3, AHP dU product; lanes 4 to 7, AHP dU products labelled individually with Cy3-BCN, Cy5-BCN, FAM-BCN and Cy5-T_{10}-BCN. **C.** Cy5-T_{10}-BCN labelled product. (20% denaturing PAGE)

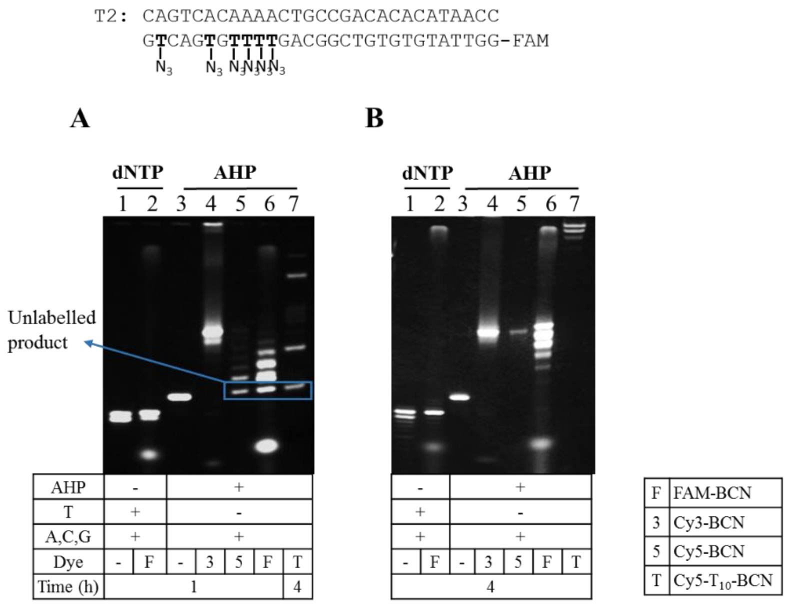


Figure 4.14. Fluorescent labelling of AHP dU full-length products from template T2. **A.** Reactions with Cy3-BCN, Cy5-BCN, FAM-BCN (5 eq) at RT for 1 h and Cy5-T_{10}-BCN (1 eq) at 55 °C for 4 h. **B.** Reactions with Cy3-BCN, Cy5-BCN, FAM-BCN (5 eq) at RT for 4 h and Cy5-T_{10}-BCN (2 eq) at 55 °C for 4 h. Lane 1, unmodified product; lane 2, unmodified product mixed with FAM-BCN, negative control; lane 3, AHP dU product; lanes 4 to 7, AHP dU products labelled individually with Cy3-BCN, Cy5-BCN, FAM-BCN and Cy5-T_{10}-BCN (20% denaturing PAGE)

For Cy5-BCN labelling reactions, weaker product bands were observed from the 4 h reaction mixture than from the 1 h reaction (Figure 4.13 and Figure 4.14, lane 5). This is presumably caused by the poor aqueous solubility of DNA containing multiple Cy5 moieties. In an attempt to overcome this solubility problem, a T_{10} oligonucleotide carrying Cy5 at the 5′-end and BCN at the 3′-end was used in the labelling reaction (Figure 4.13.C). In principle, this approach should also separate the Cy5 dyes and reduce fluorescence quenching. Unlike the conditions for the Cy5-BCN labelling reactions (5 eq, RT, 1 h), only one equivalent of Cy5-T_{10}-BCN was used for this labelling reaction which was carried out at 55 °C for 4 h. Fluorescent bands with much slower mobility than Cy5-BCN labelled DNA products were observed (Figure 4.13.A and Figure 4.14.A, lanes 5 and 7). The labelled Cy5-T_{10}-BCN strands increased the size of the AHP dU products,

forming branched DNA structures (Figure 4.13.C). Increasing the ratio of Cy5-T_{10}-BCN to DNA products to 2:1 improved the labelling reaction further. The main products bands migrated close to the loading well and no unlabelled products were detected (Figure 4.13.B and Figure 4.14.B, lane 7). The Cy5-T_{10}-BCN labelled products gave more intense bands than the Cy5-BCN labelled products (Figure 4.13.B and Figure 4.14.B, lanes 5 and 7). These results confirm that the oligothymidine linker improves the aqueous solubility of the labelled products. This strategy also proved to be an efficient and easy method to make branched DNA probes.[212, 213]

In conclusion, primer extension studies have demonstrated that the AM and AHP dUTPs are accepted as good substrates by a variety of DNA polymerases (Klenow, Gotaq and KOD). This is an efficient strategy to introduce multiple modifications in the linear extension region of DNA. The AHP modification with an alkynyl linkage stabilises the DNA duplex, whereas the AM dU has a destabilising effect. Post-labelling of AM- and AHP-modified primer extension products using the SPAAC reaction demonstrated that AHP dU with a longer linker between the azide and the uracil base was labelled more efficiently. AHP dU full-length DNA products were fluorescently labelled with different BCN fluorophores. Cy3-BCN exhibited higher reactivity than FAM-BCN and Cy5-BCN, but FAM labelled products had better aqueous solubility. The labelling reaction between the AHP product and Cy5-T_{10}-BCN showed that introducing the T_{10} oligomer improves the fluorescent property and aqueous solubility of Cy5-labelled DNA, and could be used to prepare branched DNA nanostructures. This strategy is well worth pursuing further in the future.

4.4 PCR using azide dUTPs

The success in the enzymatic incorporation of AM and AHP dUTPs using thermostable polymerases (Gotaq and KOD) in primer extension prompted the study of these two azide triphosphates in PCR (Chapter 1, Section 1.3.2). In applications for which there is insufficient template or which requires densely modified products, PCR is the method of choice for introducing modifications into DNA duplexes. It results in the labelling of both strands simultaneously. In this section, PCR amplification in the presence of AM or AHP dUTP is examined with two different templates (Figure 4.15 and Table 4.3).

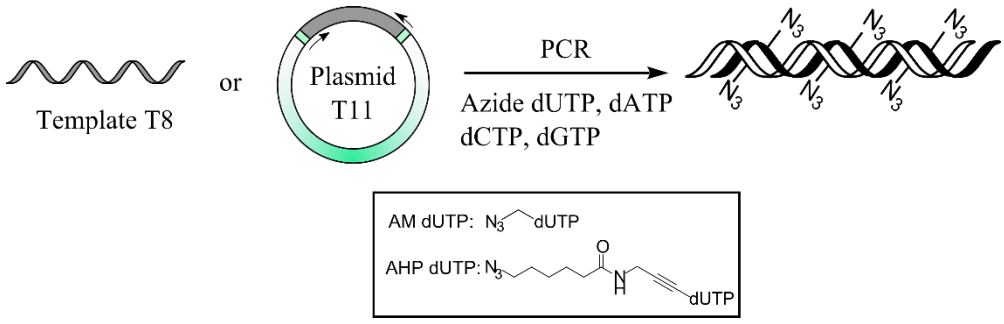


Figure 4.15. PCR using azide dUTPs and different templates. The two arrows indicate the amplification direction and the amplification region is in grey.

Table 4.3. Oligonucleotide sequences used in PCR.

Code	Sequences (5′-3′)
P6	GCATTCGAGCAACGTAAG (PCR primer for T8)
P7	GGTTATGTGTGTCGGCAG (PCR primer for T8)
T8	GGTTATGTGTGTCGGCAGTATTGTCAGTGTGAATTCCAGAGTGTGAGATTGTGTGCTGG CGATCTTACGTTGCTCGAATGC (PCR template)
P9	GTTTGGCTTTAGAGGCTGGAG (PCR primer for T11)
P10	ACTGCAATACGAATAATGGCTAC (PCR primer for T11)
T11	501 AAGATATGGG ACATAAAC**GT TTGGCTTTAG AGGCTGGAG**A AGACCCTGTA 551 AACAATCCTA TTGAATATAT TCTTGACTGT ATCAAAACCA TATACAGCAT 601 AAAACATAAA AATGGAGCAA TTAGACGTGT AAATGTAAAT ATTGCAGCTA 651 CTACTGTAGA AAACTACAAG AAATTAAAGG ATGCTGGTAT TGGAACATAT 701 ATACTTTTCC AAGAAACCTA TAACAAAAAA AGTTACGAGG AACTTCATCC 751 TACAGGTCCA AAACATGATT ATGCCTATCA TACAGAAGCA ATGGATCGTG 801 CTATGGAAGG TGGTATTGAT GATGTAGGTA TTGGGGTTTT GTTTGGACTA 851 AATATGTACA AATATGACTT TGTTGGACTT CTAATGCATG CTGAACACTT 901 GGAAGCTGCT ATGGGTGTAG GCCCTCATAC TATAAGCGTT CCTCGTATAC 951 GTCCTGCAGA TGACATTGAT CCTGAAAACT TCTCAAATGC AATATCGGAC 1001 GAGATTTTTG AAAAAATT**GT AGCCATTATT CGTATTGCAG T** (amplified region of plasmid HydGdCTD5 template)

The amplification compatibility of AM and AHP dUTPs in PCR was explored initially using an 81-mer DNA template T8 (Table 4.3) and Gotaq polymerase. Different percentages of modified dUTP mixed with natural dTTP were used to control the number of reactive groups introduced into PCR products for subsequent fluorescent labelling. In all cases, PCR amplifications were successful, even when using 100% modified azide dUTP in combination with dATP, dCTP and dGTP (Figure 4.16.A, lanes TM4 and TH4). Melting experiments on the amplicons showed that the AHP dU modification preserved the stability of the DNA duplex, whereas the AM dU destabilised the duplex relative to unmodified DNA proportionally to the number of modifications in the PCR products (Figure 4.16.B). These results are consistent with the previous melting temperature studies on primer extension products. The effects of modified nucleotides on duplex stability should be taken into account when designing PCR protocols for the incorporation of modified dUTPs.

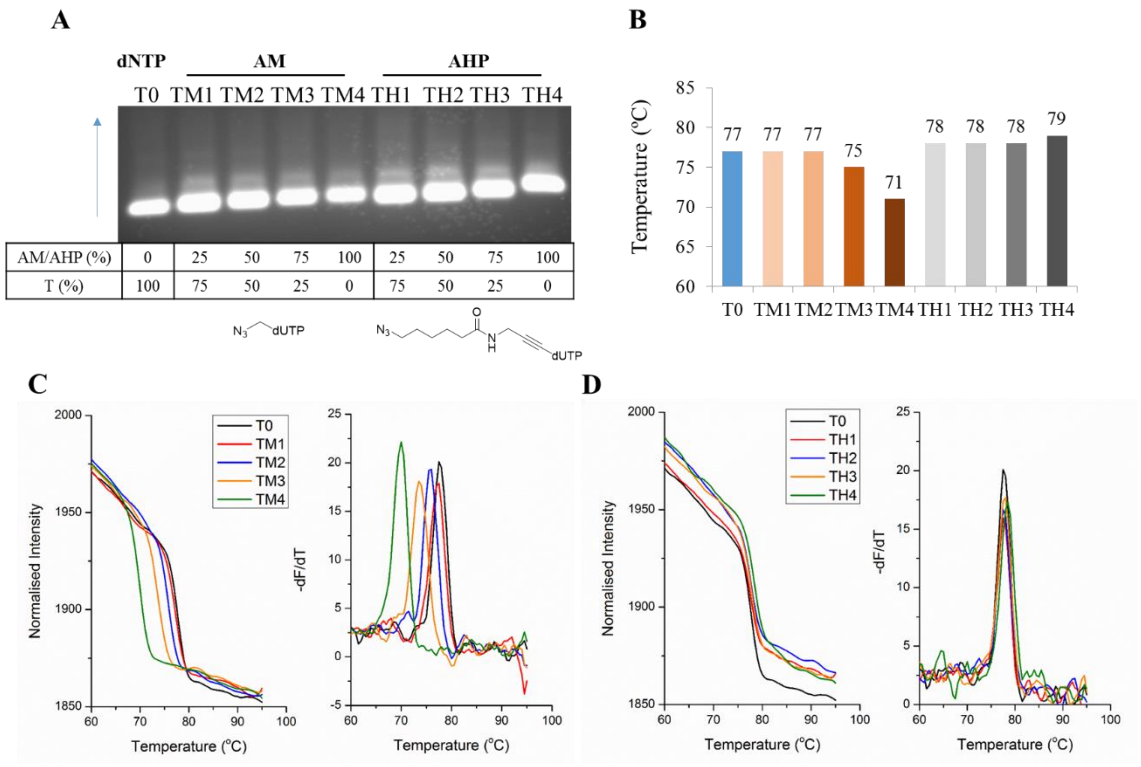


Figure 4.16. **A.** PCR using Gotaq polymerase and template T8 (primer P6 and P7) with different ratios of azide dUTP to dTTP (5 nmol in total of azide dUTP/dTTP + 5 nmol each of dATP, dCTP and dGTP). Lane T0, unmodified amplicon; lanes TM1/TH1 to TM4/TH4, 25% to 100% azide-modified amplicons. (2% agarose gel containing ethidium bromide) **B.** Melting temperatures of azide-modified PCR amplicons. Melting curves and derivatives of AM dU (**C**) and AHP dU (**D**) modified PCR amplicons.

Encouraged by these results, PCR amplification using a longer, more demanding plasmid template T11 (Table 4.3) was investigated using KOD polymerase, which amplifies DNA more efficiently than Gotaq polymerase. Initial PCR with different ratios of azide dUTP to dTTP afforded fully modified products for AHP dUTP reactions but not for AM dUTP (Figure 4.17). The inefficient amplification using AM dUTP might be due to its duplex destabilising effect, a problem that might be solved by increasing the divalent Mg^{2+} cation concentration. Magnesium ions bind to the negatively charged phosphate backbone more tightly than monovalent cations, and therefore has a positive effect on the stability of the DNA duplex.[187]

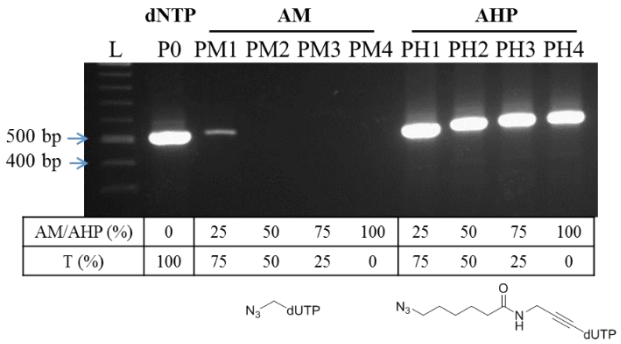


Figure 4.17. PCR using KOD polymerase (1 mM $MgCl_2$) and template T11 (primer P9 and P10) with different ratios of azide dUTP to dTTP (10 nmol in total of azide dUTP/dTTP + 10 nmol each of dATP, dCTP and dGTP). Lane L, 100 bp ladder; lane P0, unmodified amplicon; lanes PM1/PH1 to PM4/PH4, 25% to 100% azide-modified amplicons. (1.5% agarose gel containing ethidium bromide)

As a consequence, real-time PCR using 2 mM $MgCl_2$ was conducted. Promisingly, PCR with 100% of either AM or AHP dUTP gave intense product bands (Figure 4.18.A, lanes PM4 and PH4). Melting temperatures of modified PCR products increased progressively with higher ratios of AHP dUTP in the reaction; in contrast the AM modifications had a destabilising effect, giving a 10 ºC reduction in T_m between the fully modified and unmodified amplicons (Figure 4.18.B).

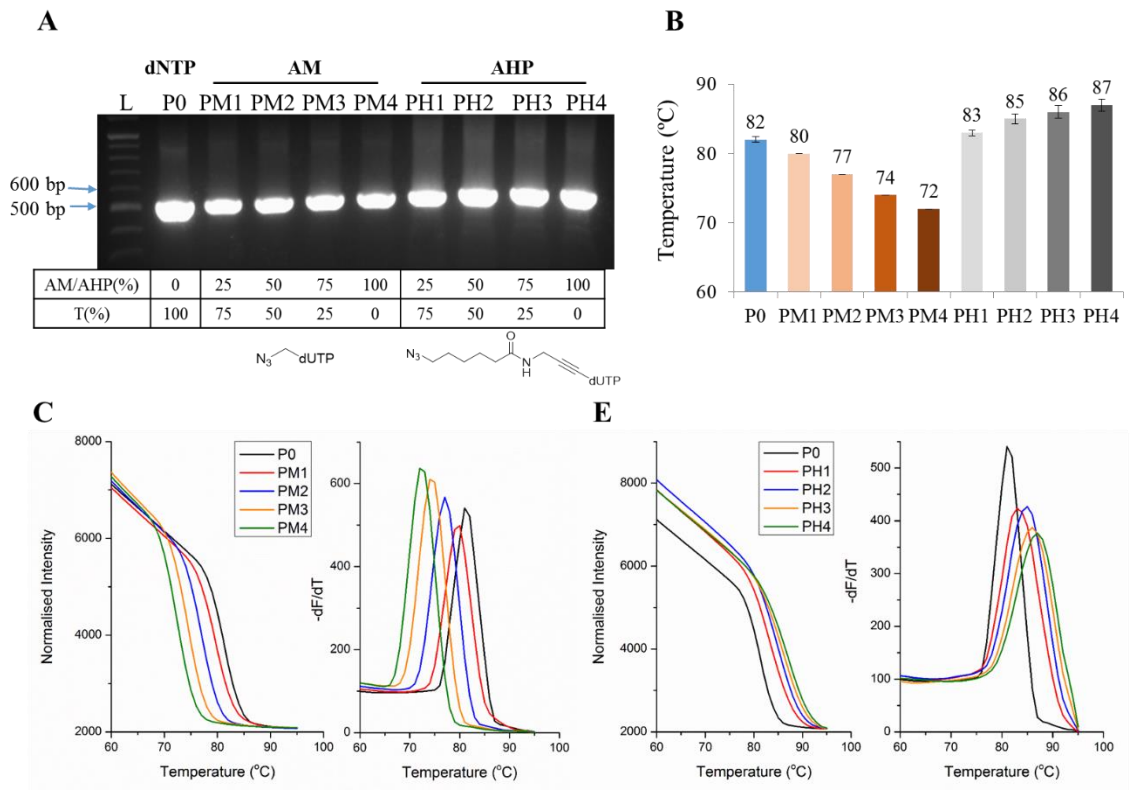


Figure 4.18. **A.** PCR using KOD polymerase (2 mM $MgCl_2$) and template T11 (primer P9 and P10) with different ratios of azide dUTP to dTTP (10 nmol in total of azide dUTP/dTTP + 10 nmol each of dATP, dCTP and dGTP). Lane L, 100 bp ladder; lane P0, unmodified amplicon; lanes PM1/PH1 to PM4/PH4, 25% to 100% azide-modified amplicons. (1.5% agarose gel containing ethidium bromide) **B.** Melting temperatures of azide-modified PCR amplicons. Melting curves and derivatives of AM dU (**C**) and AHP dU (**D**) modified PCR amplicons.

To confirm the fidelity of base pairing during PCR, AHP-modified amplicons with different AHP dU to dT ratios were directly analysed by automated Sanger sequencing.[214] Only the 25% AHP-modified product (PH1) had the full and correct sequence. With more modifications in the products (higher AHP dU percentages), the sequencing enzyme terminated earlier (stopping after 351 bp (50% PH2), 173 bp (75% PH3), 281 bp (100% PH4)). This is not surprising as the sequencing polymerase was not designed to read through templates containing modified bases. Therefore, to circumvent the problem the functionalised products were used as templates for a second-round of PCR amplification using natural triphosphates and Gotaq polymerase (Figure 4.19). This produced templates

that were suitable for Sanger sequencing and gave perfect sequencing results for all products (Appendix 8.7).

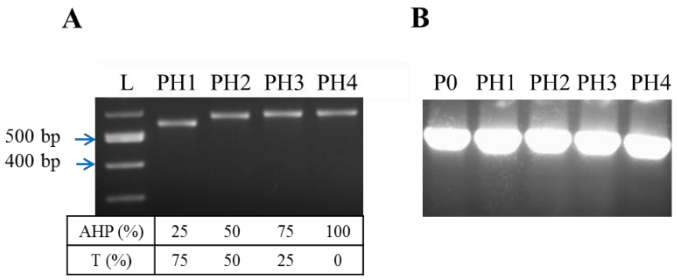


Figure 4.19. Second-round of PCR using dNTPs and AHP-modified amplicons as templates. **A.** Different percentages of AHP dU-modified templates. **B.** Amplicons from the second-round of PCR using Gotaq polymerase. Lane L, 100 bp ladder; lane P0, unmodified amplicon template; lanes PH1 to PH4, 25% to 100% AHP-modified templates. (1.5% agarose gel containing ethidium bromide)

4.4.1 Fluorescent labelling of AHP-modified PCR products (T11)

After the success in obtaining AM- and AHP-modified PCR products, the next step was to fluorescently label these PCR amplicons (523 bp) to prepare long fluorescent DNA probes that at present cannot be synthesised by solid phase oligonucleotide synthesis. As the AM modification exhibited poor labelling efficiency in the primer extension labelling reaction, only AHP dU PCR amplicons were studied in this section (Figure 4.20).

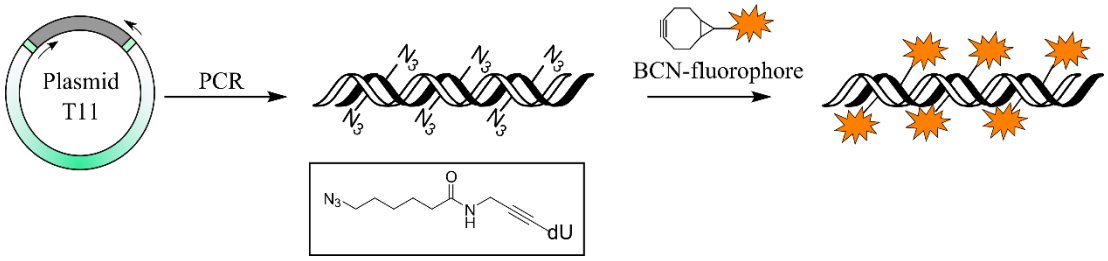


Figure 4.20. Fluorescent labelling of AHP-modified PCR products (T11).

Due to the hydrophobic nature of Cy3 and Cy5, it was anticipated that the aqueous solubility of fully labelled DNA products could be problematic. Hence, the nature of the

labelling reaction buffer was initially investigated. For the ease of availability, Cy3-BCN labelling of 50% AHP-modified products (PH2) was carried out in 1 × KOD or Gotaq green buffer from 10 min to 60 min (Figure 4.21.A). Intense labelled product bands with altered electrophoretic mobility were observed for the reactions in Gotaq green buffer, suggesting these products had been labelled with different numbers of Cy3-BCN units. However, for the labelling reaction in KOD buffer, a faint band was observed in the 10 min reaction, but no bands thereafter, presumably due to the poor aqueous solubility of the Cy3-BCN labelled products as noted previously. Whilst these results demonstrated that the Gotaq green buffer is a better reaction medium. However, when the DNA was heavily labelled by increasing the reaction time, aqueous solubility was still an issue. No product was observed after 4 h labelling for 100% AHP-modified amplicons (Figure 4.21.B, lane PH4). For the 75% AHP-modified product, a faint band was detected after 2 h (lane 4). Therefore, further labelling reactions were restricted to 0.5 to 1 h, which gives sufficient labelling and keeps the products soluble.

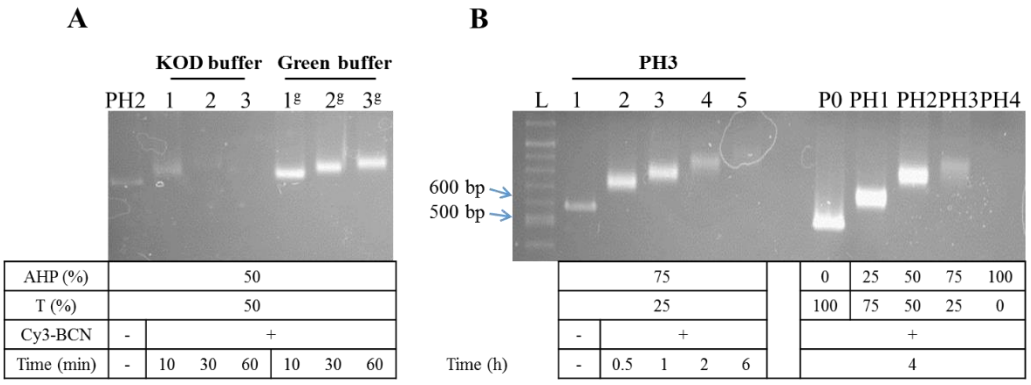


Figure 4.21. Cy3-BCN fluorescent labelling of AHP-modified PCR amplicons. **A.** Labelling of 50% AHP-modified amplicons (PH2) in KOD polymerase buffer 1 (1 ×) and Gotaq green buffer (1 ×). Lane PH2, unlabelled PH2 amplicon; lanes 1/1^{g} to 3/3^{g}, labelling of PH2 amplicons at RT for 10, 30 and 60 min respectively. **B.** Labelling of AHP-modified PCR products at RT. Lane 1, unlabelled 75% AHP dU amplicon (PH3); lanes 2 to 5, labelling of PH3 amplicons for 0.5, 1, 2 and 6 h respectively; lane P0, unmodified amplicon mixed with Cy3-BCN, negative control; lanes PH1 to PH4, labelling of 25% to 100% AHP-modified amplicons for 4 h. (1.5% agarose gel containing ethidium bromide)

4.4.1.1 Cy3-BCN labelled AHP-modified PCR products

Based on these initially optimised labelling conditions, the AHP-functionalised PCR products were labelled with Cy3-BCN in 1 × Gotaq green buffer for 30 min (Figure 4.22.A). Labelled products showed highly fluorescent bands without staining with ethidium bromide (Figure 4.22.B). The 50% AHP dU PCR product gave the strongest fluorescent signal (Figure 4.22.E, PH2-Cy3). With all dT units in the PCR product replaced by AHP dU (PH4-Cy3), the fluorescent signal decreased, presumably due to collisional quenching between fluorophores or poor aqueous solubility of the densely labelled products. This was consistent with the post-labelling of amino functionalised RNA strands reported by ’t Hoen and co-workers.[71] They also demonstrated that fluorescent labelling of the product, which was generated from enzymatic incorporation of an aminoallyl uridine triphosphate (aa-UTP) and unmodified UTP in a 1:1 ratio, gave the optimal fluorescent signal.

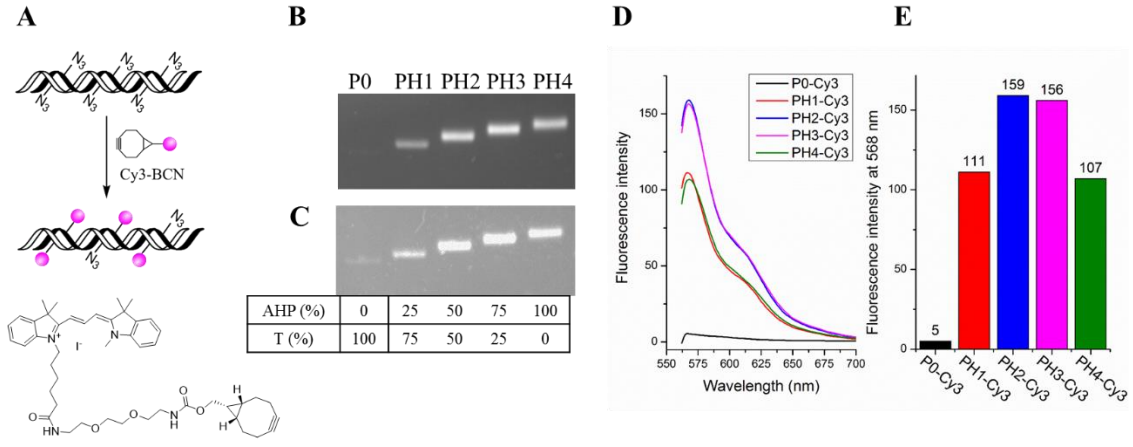
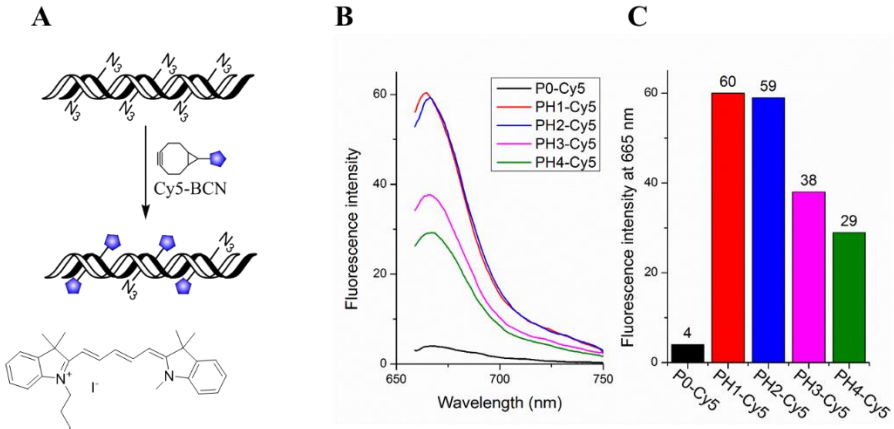


Figure 4.22. Cy3-BCN fluorescent labelling of AHP-modified PCR amplicons. **A.** Scheme for Cy3-BCN labelling. Agarose gel analysis of labelling reactions before staining (**B**) and after staining with ethidium bromide (**C**). (1.5% agarose gel) **D.** Fluorescence emission spectra of labelled PCR amplicons, excited at 545 nm. **E**. The fluorescence intensity at 568 nm for each sample. P0/P0-Cy3, unmodified amplicon mixed with Cy3-BCN, negative control; PH1/PH1-Cy3 to PH4/PH4-Cy3, 25% to 100% AHP-modified amplicons labelled with Cy3-BCN at RT for 30 min.

4.4.1.2 Cy5-BCN and Cy5-T$_{10}$-BCN labelled AHP-modified PCR products

Cy5-BCN was also used to label AHP-modified PCR products from the template T11 (Figure 4.23.A). However, the agarose gel electrophoresis purification procedure used for Cy3-BCN labelling failed with Cy5-BCN. Labelled or unlabelled PCR amplicons were not detected after gel extraction. Given that Cy5 is more hydrophobic than Cy3, solubility might be an even bigger influence on labelling yield. Therefore, instead of gel extraction to purify the labelling reaction mixture, excess Cy5-BCN monomer was removed by ethanol precipitation and steady state fluorescence of the samples was directly monitored. The 25% AHP dU product exhibited the most intense signal (Figure 4.23.C, PH1-Cy5), with intensity decreasing as the AHP dU modification progressively increased. This might be caused by the strong fluorescence collisional-quenching effect of Cy5, which has also been observed in other reports of densely Cy5-modified nucleic acids and proteins.[71, 215]



	P0	PH1	PH2	PH3	PH4
AHP (%)	0	25	50	75	100
T (%)	100	75	50	25	0

Figure 4.23. Cy5-BCN fluorescent labelling of AHP-modified PCR amplicons. **A.** Scheme for Cy5-BCN labelling. **B.** Fluorescence emission spectra of labelled PCR amplicons, excited at 645 nm. **C.** The fluorescence intensity at 665 nm for each sample. P0-Cy5, unmodified amplicon mixed with Cy5-BCN, negative control; PH1-Cy5 to PH4-Cy5, 25% to 100% AHP-modified amplicons labelled with Cy5-BCN at RT for 30 min.

Once again, to overcome the aqueous solubility and fluorescence quenching issues associated with the Cy5 fluorophore, oligo T_{10} with Cy5 and BCN groups on the termini was used to label the AHP-functionalised PCR amplicons (Figure 4.24.A). At room temperature no desired products were obtained, whereas, when the reactions were left at 55 °C for 4 h in a heating block, densely labelled products were observed (Figure 4.24.B, lanes PH1 to PH4). In the latter case, the solvent evaporated to dryness gradually, which promoted the SPAAC labelling reaction. For the AHP dU products with 50% and 100% modifications, intense Cy5 fluorescence signals were observed in the loading well (Figure 4.24.C); the 25% AHP dU labelled product afforded the most intense fluorescent band (lane PH1).

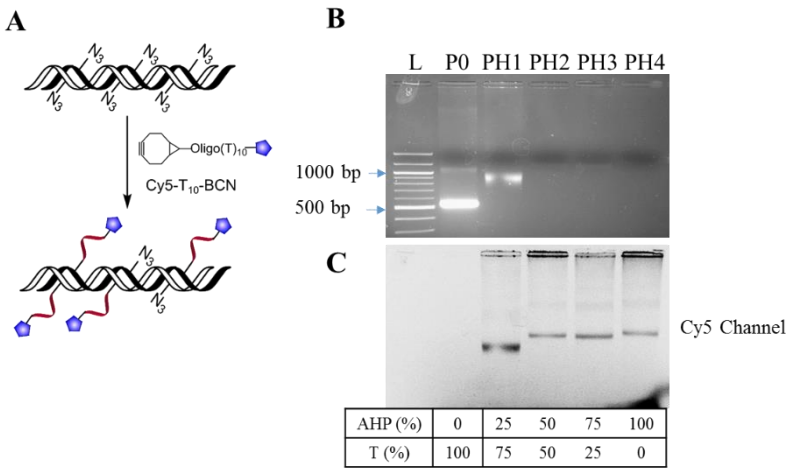
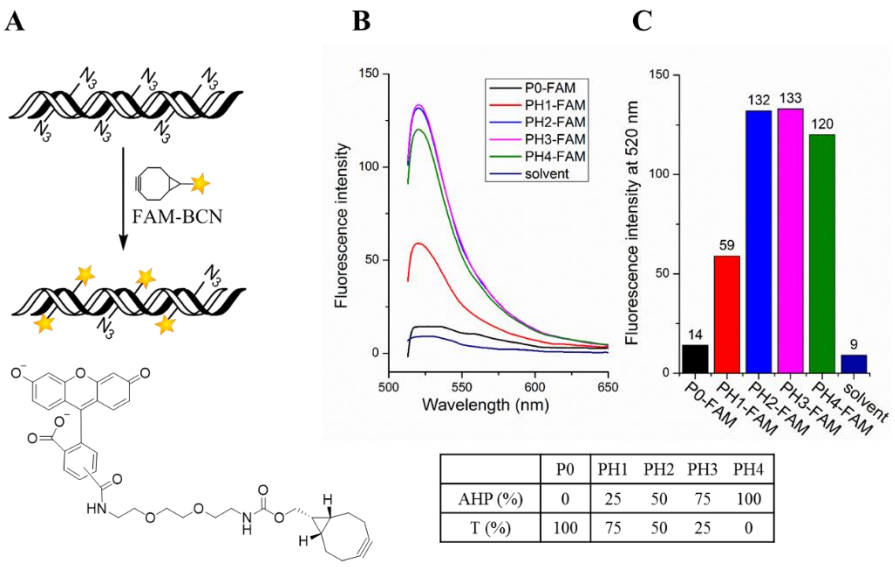


Figure 4.24. Cy5-T_{10}-BCN fluorescent labelling of AHP-modified PCR amplicons. **A.** Scheme for Cy5-T_{10}-BCN labelling. Agarose gel analysis of labelling reactions visualised on a transilluminator (**B**, 302 nm) and by Cy5 channel (**C**). Lane P0, unmodified amplicon, negative control; lanes PH1 to PH4, 25% to 100% AHP-modified amplicons labelled with Cy5-T_{10}-BCN at 55 ºC for 4 h. (1.5% agarose gel containing ethidium bromide)

4.4.1.3 FAM-BCN labelled AHP-modified PCR products

Fluorescein, which has far better aqueous solubility than Cy3 and Cy5, was also investigated in the labelling reaction (Figure 4.25.A). As with Cy3, the

AHP-functionalised PCR products were efficiently labelled within 30 min. When recording the fluorescence measurements, it should be noted that the pH strongly influences FAM fluorescence. The FAM moiety has cationic, neutral, anionic and dianionic forms (Figure 4.26),[216] of which the most fluorescent is the dianionic form at pH ~ 8 and above (ε_{490} = 79000 $M^{-1}cm^{-1}$, Φ = 0.9).[217] Therefore, fluorescence measurements were recorded in 75 mM TEAB buffer at pH 8.0. The 50% and 75% AHP dU amplicons gave the most intense fluorescent signals (Figure 4.25.C, PH2-FAM and PH3-FAM). However, for the 100% AHP dU product, the fluorescence signal slightly decreased, most likely due to fluorescence quenching (PH4-FAM).



	P0	PH1	PH2	PH3	PH4
AHP (%)	0	25	50	75	100
T (%)	100	75	50	25	0

Figure 4.25. FAM-BCN fluorescent labelling of AHP-modified PCR amplicons. **A.** Scheme for FAM-BCN labelling. **B.** Fluorescence emission spectra of labelled PCR amplicons, excitation at 480 nm. **C.** Fluorescence intensity at 520 nm for each sample. P0-FAM, unmodified amplicon mixed with FAM-BCN, negative control; PH1-FAM to PH4-FAM, 25% to 100% AHP-modified amplicons labelled with FAM-BCN at RT for 30 min; solvent, 75 mM TEAB buffer (pH 8.0).

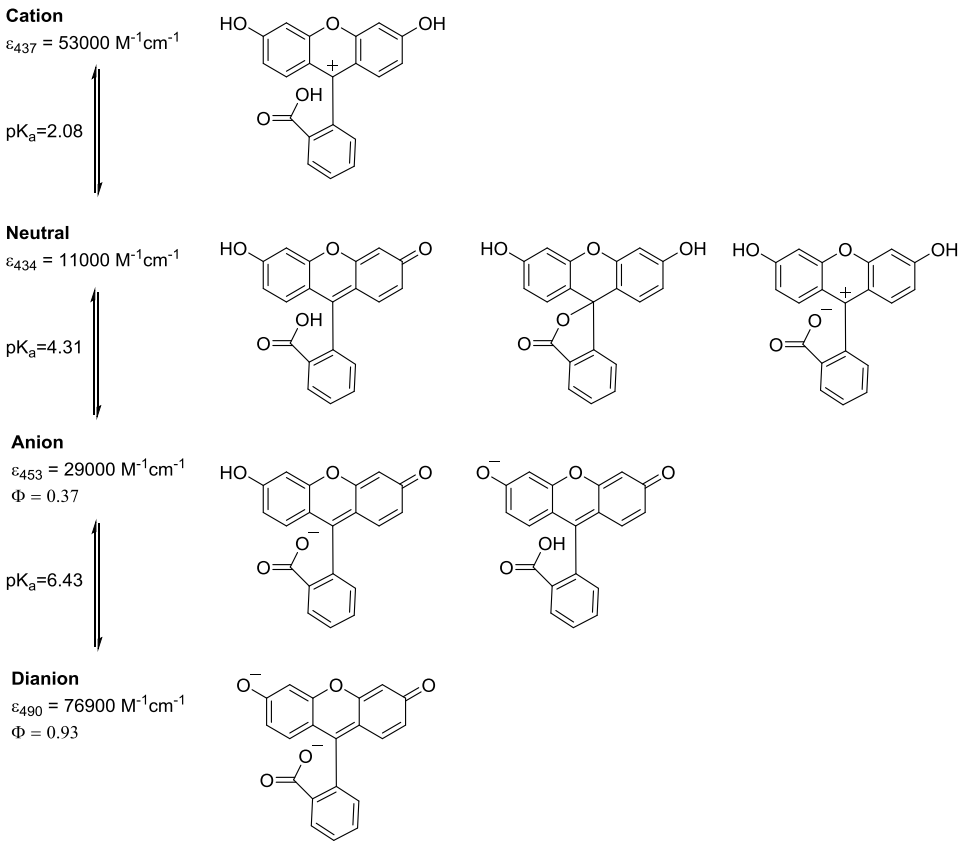


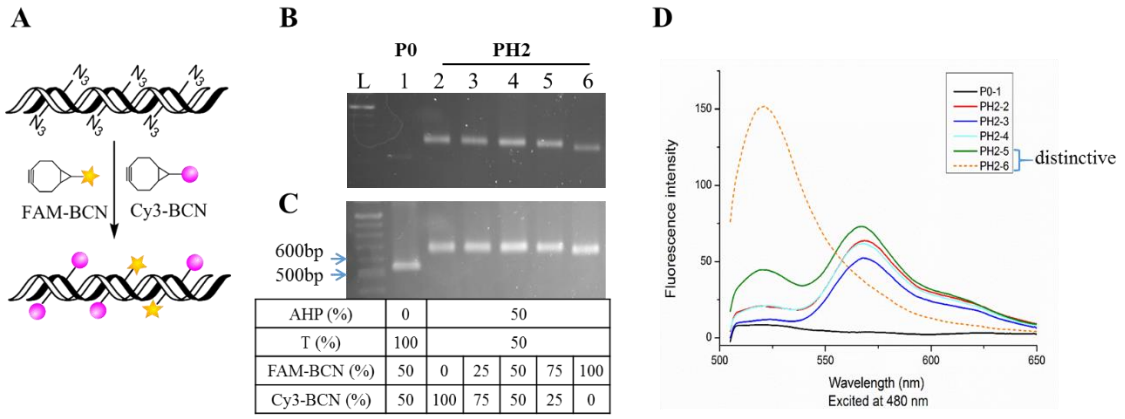
Figure 4.26. Chemical structures of fluorescein in its cationic, neutral, anionic and dianionic states.[217]

4.4.1.4 Cy3 and FAM dual labelled AHP-modified PCR products

For genomic FISH, single fluorophore labelled probes (*i.e.* multiple units of one type of fluorophore) can be used to detect specific gene sequences. However, to detect multiple gene targets simultaneously, differently coloured fluorescent probes with distinct spectra are required.[34] However, the number of fluorescent dyes that can be used for nucleic acid detection is limited due to spectral overlap. This problem can potentially be resolved by introducing multiple coloured fluorophores (e.g. FAM and Cy3) into a single probe. To achieve this, different fluorophores can be attached to DNA *via* orthogonal chemical reactions which require specific functional groups on different phosphoramidites/dNTPs during the DNA assembly reaction (Chapter 1, Section 1.4.5).[67, 102, 138] This will produce probes with subtly different fluorescence emission spectra, akin to mixing different

primary colours to produce various shades. However, these strategies require the synthesis of different modified monomers (dNTPs). Here, a labelling strategy that 'blends' two different coloured BCN labelling fluorophores in a single reaction to give fluorescent probes with specific emission spectra is developed.

To maintain the labelling efficiency and minimise fluorescence quenching, the PCR products from 50% AHP dUTP and 50% dTTP (PH2, 523 bp, average 168 azide modifications per duplex) were used in dual fluorophore labelling reactions with different ratios of Cy3-BCN and FAM-BCN (Figure 4.27.A). Agarose gel analysis showed, as expected, that the electrophoretic mobility of different combinations of Cy3 and FAM labelled products varied because of their different structures and charge states (Figure 4.27.B and C). To detect fluorescent signals of both dyes simultaneously by fluorimetry, samples were dissolved in TEAB buffer (75 mM, pH 8.0) and excited at 480 nm. These samples were then reused after lyophilisation for gel electrophoresis analysis. When the ratio of FAM-BCN changed from 100% to 75%, the fluorescence emission spectra changed (Figure 4.27.D, PH2-5 and PH2-6), which demonstrates that diverse fluorescently labelled products can be prepared simply by labelling with fluorescent dyes in different combinations and ratios. This principle could be applied to two-colour microarray-based hybridisation to detect more than two gene sequences simultaneously. In the mixing range of FAM (100% to 75%) and Cy3 (0% to 25%), further alternative combinations could be used to obtain different coloured fluorescent probes.



AHP (%)	0	50				
T (%)	100	50				
FAM-BCN (%)	50	0	25	50	75	100
Cy3-BCN (%)	50	100	75	50	25	0

Figure 4.27. Cy3-BCN and FAM-BCN dual labelling of 50% AHP-modified PCR amplicons. **A.** Scheme for dual labelling. Agarose gel analysis of the dual labelled products before staining (**B**) and after staining with ethidium bromide (**C**). Lane L, 100 bp ladder. (1.5% agarose gel) **D.** Fluorescence emission spectra of the dual labelled products, excited at 480 nm (75 mM TEAB buffer, pH 8.0). 1/P0-1, unmodified amplicon mixed with Cy3-BCN (50%), negative control; 2/PH2-2 to 6/PH2-6, 50% AHP-modified amplicons labelled with different combinations of Cy3-BCN (100% to 0%) and FAM-BCN (0% to 100%) respectively.

To summarise, PCR amplification of AM and AHP dUTPs was successfully carried out even when the modified dUTP was not supplemented with dTTP. The AHP dUTP was amplified more efficiently than the AM dUTP and the AHP-modified amplicon duplexes demonstrated good thermal stability. The PCR amplicons (523 bp) from reactions with different ratios of AHP dUTP:dTTP all gave 100% correct sequencing results (*i.e.* demonstrate faithful incorporation during PCR). These AHP-functionalised PCR products were efficiently post-labelled with different BCN-modified fluorescent reporter groups, and afforded intense fluorescence signals. A single step two-dye labelling method was developed to prepare DNA probes with distinct fluorescence spectra. This is a potential methodology to synthesise various fluorescent probes using limited numbers of fluorophores for FISH applications.

4.5 Synthesis of single and dual labelled fluorescent ssDNA probes

The probes discussed so far were double-stranded, however, single-stranded probes are more sensitive than the denatured duplex DNA for hybridisation based detection.[218] In addition to automated solid phase oligonucleotide synthesis (Chapter 1, Section 1.3.1), there are several methods to prepare single-stranded DNA (ssDNA). This is usually performed using PCR amplification followed by exonuclease digestion of one strand, magnetic beads separation (streptavidin-biotin binding), denaturing PAGE separation or denaturing HPLC.[219] Several of these methods are explored to prepare fluorescently labelled single-stranded probes from AHP-functionalised PCR products in this Section.

4.5.1 Synthesis of single-stranded probes *via* exonuclease digestion

λ-exonuclease is derived from *E. coli* exonuclease and preferentially digests the 5′-monophosphate labelled strand of a DNA duplex.[220, 221] It has a toroidal structure consisting of a homotrimer.[222] The DNA duplex substrate is positioned in the centre of the toroid and the exonuclease threads through the DNA, enabling efficient processivity.[222, 223] O'Sullivan and co-workers compared several different ssDNA production methods and demonstrated that the combination of λ-exonuclease digestion and asymmetric PCR (APCR) afforded single-stranded products with higher recovery yield and purity.[219] Compared to PCR, APCR generates a mixture of double- and single-stranded products by using unequal concentrations of the two primers (Chapter 1, Section 1.3.2). Therefore, the AHP-modified ssDNA was initially synthesised by exonuclease digestion of PCR products or by a combination of exonuclease digestion and APCR (Figure 4.28). One PCR/APCR primer contained a 5′-phosphate group making this strand a better substrate for λ-exonuclease.

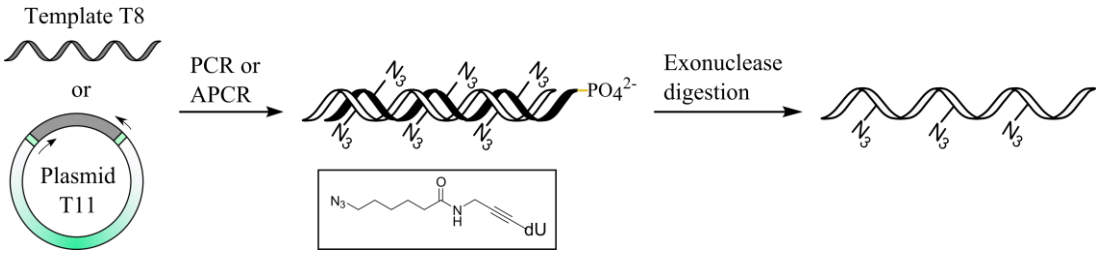


Figure 4.28. λ-exonuclease digestion of PCR or APCR products from AHP dUTP, template T8 or T11.

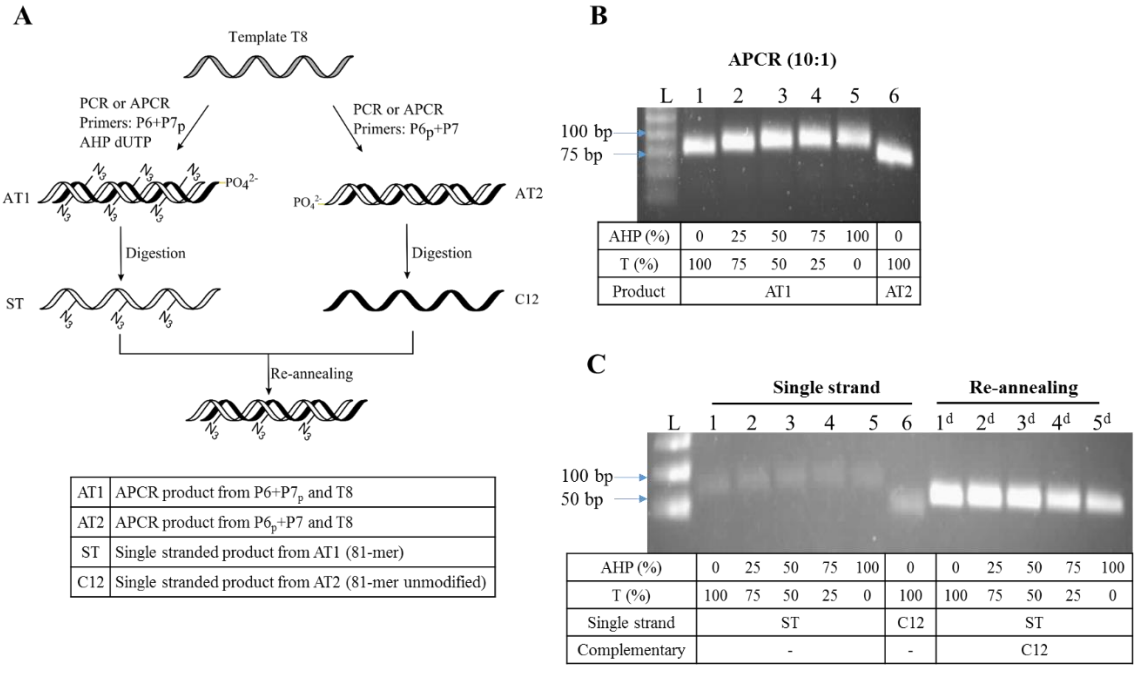
Table 4.4. Oligonucleotide sequences used in PCR/APCR and single strand synthesis.

Code	Sequences(5′-3′)
P6	GCATTCGAGCAACGTAAG (primer for T8)
P7	GGTTATGTGTGTCGGCAG (primer for T8)
$P6_p$	Phosphate-GCATTCGAGCAACGTAAG (primer for λ-exonuclease digestion)
$P7_p$	Phosphate-GGTTATGTGTGTCGGCAG (primer for λ-exonuclease digestion)
$P6_b$	Biotin-GCATTCGAGCAACGTAAG (primer for streptavidin binding)
T8	GGTTATGTGTGTCGGCAGTATTGTCAGTGTGAATTCCAGAGTGTGAGATTGTGTGCT GGCGATCTTACGTTGCTCGAATGC (template)
P9	GTTTGGCTTTAGAGGCTGGAG (primer for T11)
P10	ACTGCAATACGAATAATGGCTAC (primer for T11)
$P9_p$	Phosphate-GTTTGGCTTTAGAGGCTGGAG (primer for λ-exonuclease digestion)
$P10_p$	Phosphate-ACTGCAATACGAATAATGGCTAC (primer for λ-exonuclease digestion)
T11	Plasmid HydGdCTD5 template (Table 4.3)

4.5.1.1 Synthesis of short AHP-modified single strands

Single-stranded probes with 20 to 50 nucleotides are commonly used in hybridisation applications.[224] Therefore, the preparation of AHP-modified single strands was carried out by λ-exonuclease digestion of APCR products from a short template T8 (Figure 4.29.A). In addition to the AHP-modified single strand (ST), an unmodified and complementary C12 strand was also prepared as a control. Agarose gel electrophoresis revealed that a substantial amount of double-stranded amplicons were produced in the

APCR amplifications when a ratio of 10:1 of normal to 5′-phosphorylated primers was used (Figure 4.29.B). However, this could be due to ssDNA detection being less sensitive than double-stranded DNA (dsDNA) using ethidium bromide staining. After λ-exonuclease digestion, only weak duplex bands were observed (Figure 4.29.C, single strand lanes). To confirm that single-stranded probes had been produced, the AHP-modified (ST) and complementary unmodified (C12) single strands generated by APCR and λ-exonuclease digestion were annealed to give intense duplex bands (Figure 4.29.C, re-annealing lanes).



AT1	APCR product from P6+P7$_p$ and T8
AT2	APCR product from P6$_p$+P7 and T8
ST	Single stranded product from AT1 (81-mer)
C12	Single stranded product from AT2 (81-mer unmodified)

AHP (%)	0	25	50	75	100	0
T (%)	100	75	50	25	0	100
Product	AT1					AT2

AHP (%)	0	25	50	75	100	0	0	25	50	75	100
T (%)	100	75	50	25	0	100	100	75	50	25	0
Single strand	ST					C12	ST				
Complementary	-					-	C12				

Figure 4.29. **A.** Scheme for exonuclease digestion of APCR/PCR amplicons and re-annealing of the AHP-modified single strand (ST) with the complementary unmodified strand (C12). **B.** APCR using KOD polymerase and template T8 (10:1 of primer P6:P7$_p$/P7:P6$_p$) with different ratios of AHP dUTP to dTTP. Lane L, 25 bp ladder. **C.** λ-exonuclease digestion of APCR products and re-annealing of the AHP-modified single strands (ST) with C12. Lane L, 50 bp ladder. Lane 1/1^d/6, unmodified product; lanes 2/2^d to 5/5^d, 25% to 100% AHP-modified products. (2% agarose gel containing ethidium bromide)

To further support the above observation, agarose gel electrophoresis was repeated but with SYBR Gold staining, which can detect both ssDNA and dsDNA. These gels showed that APCR produced both double- and single-stranded products (Figure 4.30.A). As expected, the yield and purity of ssDNA was enhanced by λ-exonuclease digestion (Figure 4.30.B, single strand lanes).[219] When the AHP-modified ssDNA (ST) and complementary strand (C12) were annealed, a fainter faster migrating band was also observed in addition to the intense duplex band as a result of an excess of primers after APCR (Figure 4.30.B, re-annealing (C12) lanes). This was confirmed as the annealing product of ssDNA with excess primers P6 and P7 (Figure 4.30.B, re-annealing (primers) lanes). To simplify the above annealing strategy, the AHP-modified ssDNA was also annealed to the complementary template T8 and the expected duplex products were consistently observed (Appendix 8.8). These results confirm that both C12 (generated from digestion of APCR product) and T8 (chemically synthesised) strands can be used in the re-annealing reaction to form the desired duplex products. The preparation of complementary DNA strand by digestion is important for the preparation of long ssDNA (523-mer, Section 4.5.1.2). Such long DNA strands cannot be synthesised by solid phase chemical methods.

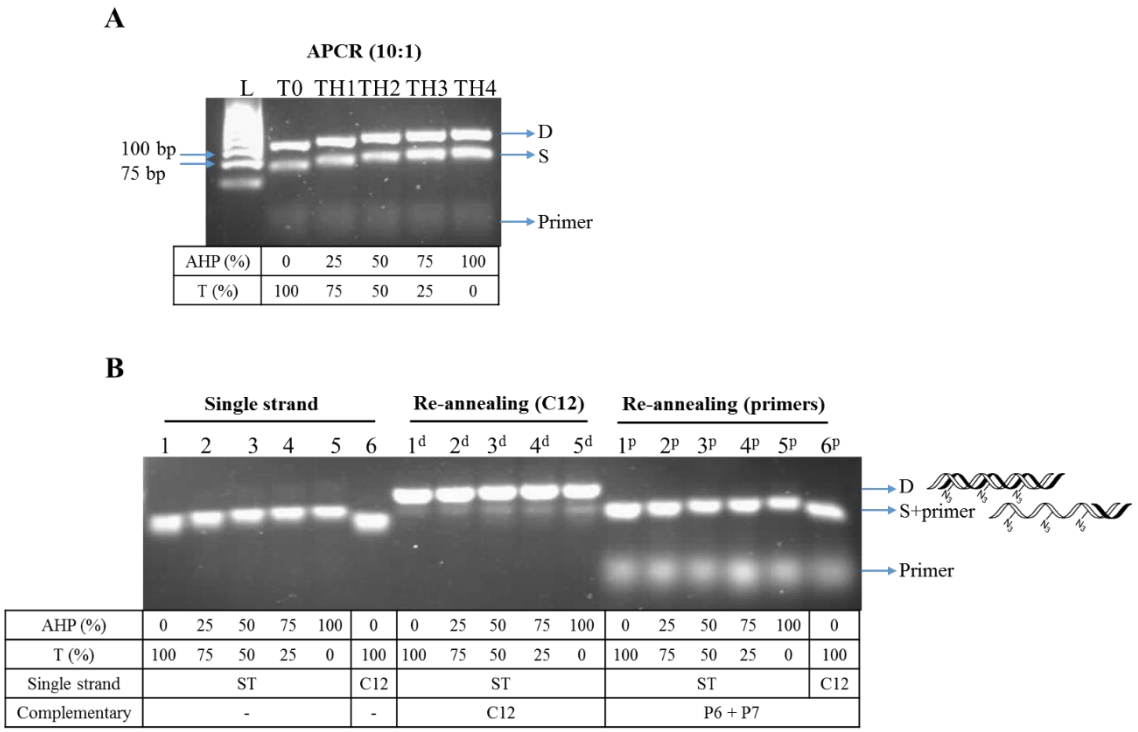
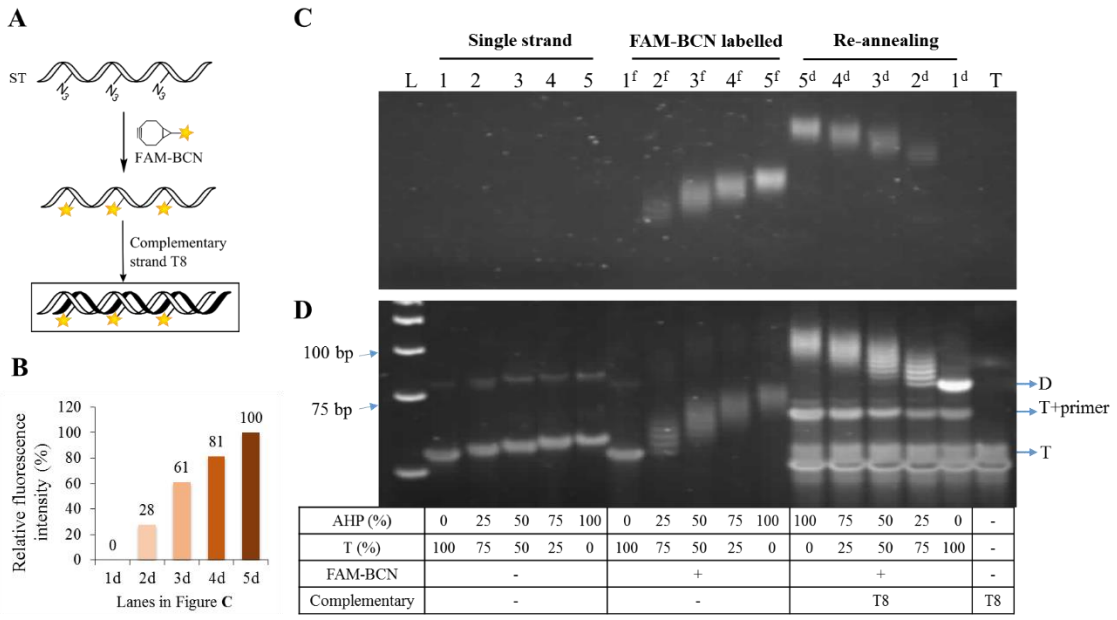


Figure 4.30. **A.** APCR using KOD polymerase and template T8 (10:1 of primer P6:P7$_p$) with different ratios of AHP dUTP to dTTP. Lane L, 25 bp ladder; lane T0, unmodified amplicon; lanes TH1 to TH4, 25% to 100% AHP-modified amplicons. D = double strands; S = single strands. **B.** λ-exonuclease digestion of APCR amplicons and re-annealing of the AHP-modified single strands (ST) with complementary strand C12 or primers (P6 and P7). Lanes 1/1^d/1^p and 6/6^p, unmodified products; lanes 2/2^d/2^p to 5/5^d/5^p, 25% to 100% AHP-modified products. S + primer = single strands annealed with primer. (2% agarose gel containing SYBR Gold)

4.5.1.1.1 Fluorescent labelling of short AHP-modified single strands

After optimisation of the preparation of AHP-modified single strands, these were labelled with FAM-BCN to prepare fluorescent probes *via* the SPAAC labelling reaction. FAM-BCN labelling with AHP-modified ssDNA was performed at 55 °C to improve the labelling efficiency. The labelled products were re-annealed with the template T8 to investigate their binding to the complementary strand (Figure 4.31.A). Native PAGE, which offers better resolution than agarose gels, was used to analyse the labelling and re-annealing reactions. Fluorescently labelled products were observed without staining

(Figure 4.31.C, FAM-BCN labelled lanes), indicating that ssDNA containing AHP dU modifications was efficiently labelled with FAM-BCN. With lower percentages of AHP dU modifications, broader fluorescent bands were observed compared to the fully modified products, indicating the varied numbers of azides inside each ssDNA (Figure 4.31.C, FAM-BCN labelled lanes). These labelled products were successfully re-annealed with template T8 and afforded fluorescent dsDNA (Figure 4.31.C, re-annealing lanes). The fluorescence intensity of the re-annealed duplexes increased with higher numbers of AHP dU modifications (Figure 4.31.B). Importantly, this shows that single-stranded oligonucleotides, even with large numbers of bulky fluorophores, can efficiently anneal to their complementary strands and could be subsequently used for DNA fluorescence detection applications.



AHP (%)	0	25	50	75	100	0	25	50	75	100	100	75	50	25	0	-
T (%)	100	75	50	25	0	100	75	50	25	0	0	25	50	75	100	-
FAM-BCN	-					+					+					-
Complementary	-					-					T8					T8

Figure 4.31. FAM-BCN fluorescent labelling of AHP-modified single strands (ST) and re-annealing of the labelled products with template T8. **A.** Scheme for labelling and re-annealing. **B.** Relative fluorescence intensity for the re-annealed duplex products (lanes 1^d to 5^d in Figure **C**). Native PAGE analysis before staining (**C**) and after staining with SYBR Gold (**D**). Lane L, 25 bp ladder; lane $1/1^f/1^d$, unmodified product; lanes $2/2^f/2^d$ to $5/5^f/5^d$, 25% to 100% AHP-modified products; lane T, template T8. D = duplex products; T + primer, excess primer annealed to template T8; T = template T8. (12% native PAGE)

4.5.1.1.2 Dual labelling of short AHP-modified single strands

After the success in obtaining one colour single-stranded probes, a dual labelling strategy using different combinations of two fluorophores was examined as described previously for double-stranded AHP-modified PCR amplicons (Section 4.4.1). 50% AHP-modified ssDNA (81-mer, average 5.5 modifications per strand) was initially labelled with FAM-BCN and Rhodamine-B-BCN **4.10** (Figure 4.32.A). Rhodamine-B-BCN was used since it has similar absorption and fluorescent emission spectra to Cy3-BCN but also has greater aqueous solubility (Figure 4.32).

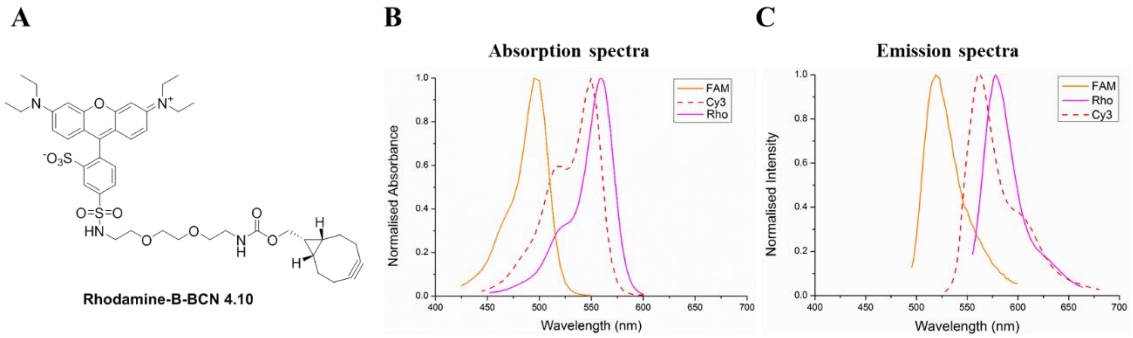
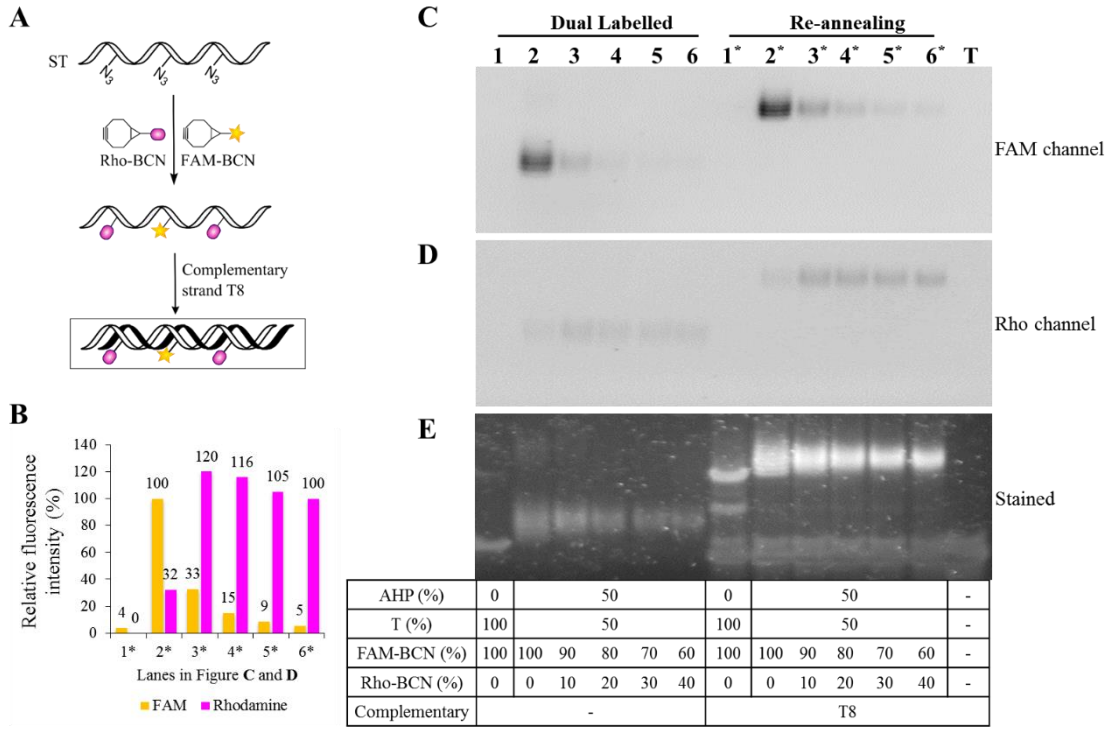


Figure 4.32. **A.** Chemical structure of Rhodamine-B-BCN (purchased from Synaffix). Absorption spectra (**B**) and emission spectra (**C**) of FAM-BCN (0.1 M TEAB buffer, pH 8.0), Cy3-BCN (EtOH) and Rhodamine-B-BCN (EtOH).

As FAM-BCN showed lower labelling efficiency in the dual-labelling of AHP-modified PCR amplicons (Section 4.4.1), a higher ratio of FAM-BCN to Rhodamine-B-BCN was used throughout. Fluorescent bands were observed by native PAGE without staining; FAM-BCN and Rhodamine-B-BCN labelled products showed fluorescent bands in their corresponding detection channels (Figure 4.33.C and D). The dual labelled probes were re-annealed with the template T8 and the fluorescence intensity of the duplex products was quantified by ImageJ software. When the percentage of FAM-BCN was changed from 100% to 90%, the relative fluorescence intensity of the FAM-signals decreased

dramatically by three-fold (Figure 4.33.B, 2* and 3*). This may be due to the lower reactivity of FAM-BCN relative to Rhodamine-B-BCN. The fluorescence intensity of Rhodamine-B decreased slightly when the percentage of Rhodamine-B-BCN was varied from 10% to 40% (Figure 4.33.B, 3* to 6*). This may be the result of spectral overlap of the two fluorophores or the collisional quenching effect of densely labelled Rhodamines (Figure 4.32.B and C). These results demonstrate that this dual-labelling strategy can afford fluorescent probes with different fluorescence profiles, which could be used in FISH applications to target different DNA or RNA sequences.



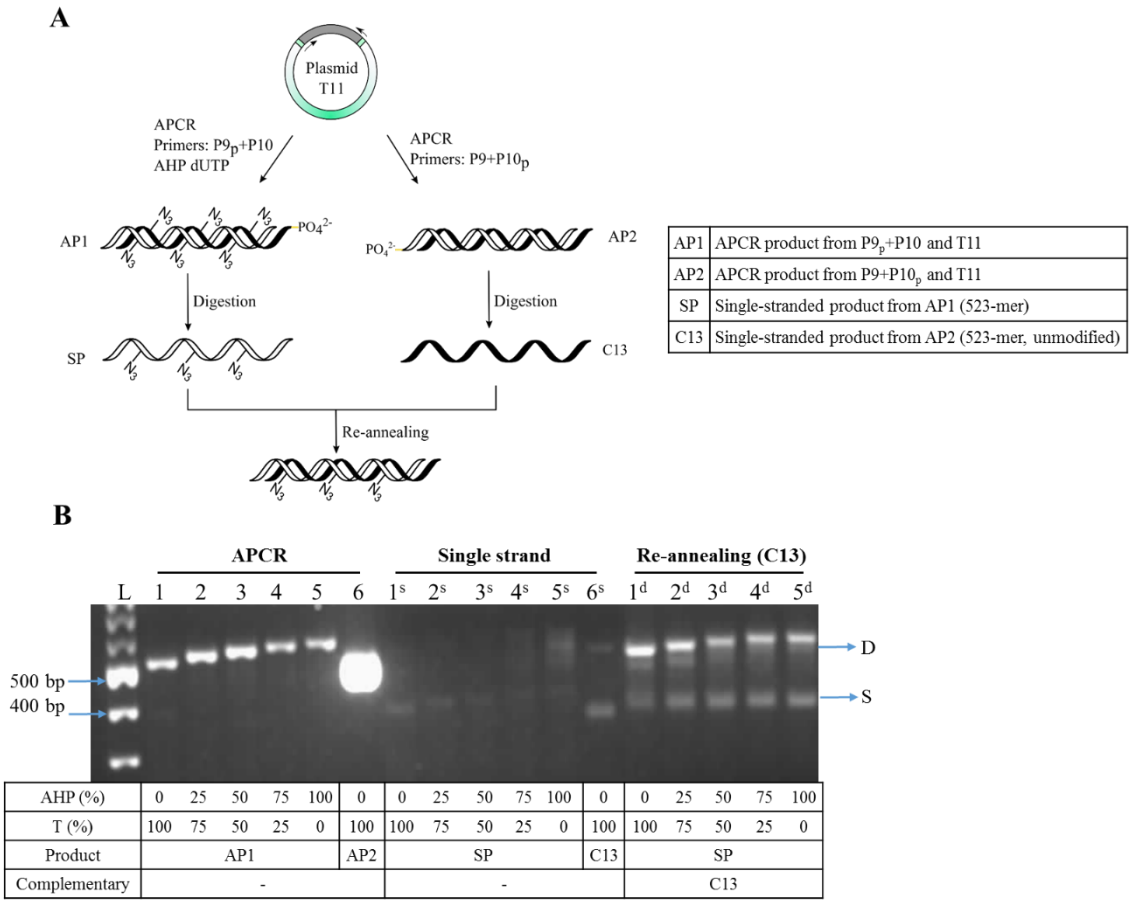
AHP (%)	0	50					0	50					-
T (%)	100	50					100	50					-
FAM-BCN (%)	100	100	90	80	70	60	100	100	90	80	70	60	-
Rho-BCN (%)	0	0	10	20	30	40	0	0	10	20	30	40	-
Complementary	-						T8						

Figure 4.33. Dual labelling of 50% AHP-modified single strands (ST) with different combinations of FAM-BCN (100% to 60%) and Rhodamine-B-BCN (0% to 40%). **A.** Scheme for dual labelling and re-annealing. **B.** Relative fluorescence intensity for the re-annealed duplex products in FAM and Rhodamine channels (lanes 1* to 6* in Figure **C** and **D**). Native PAGE analysis visualised before staining by FAM channel (**C**), Rhodamine channel (**D**) and after staining with SYBR Gold (**E**). Lane 1/1*, unmodified product, negative control; lanes 2/2* to 6/6*, 50% AHP-modified products. (12% native PAGE)

4.5.1.2 Synthesis of long AHP-modified single-stranded probes

Short single-stranded fluorescent probes of the kind described above are very useful for target DNA sequence detection. However, in the case of genome analysis, longer intensely fluorescent probes are commonly used.[225, 226] The same digestion and fluorescent labelling techniques were therefore used to synthesise long fluorescent probes from AHP-modified APCR products (T11) (Figure 4.34.A). Of the primers used, using primer P9 as the limiting primer was more efficient than using primer P10 as judged by the intensity of APCR product bands (Figure 4.34.B, APCR lanes). This is possibly due to the higher T_m (54.3 °C) of P9 relative to P10 (52.8 °C).[63, 227, 228] The T_m of an oligonucleotide indicates its binding affinity to its complementary strand and is concentration dependent. Therefore, when primer P9 is used as a limiting primer in APCR, it anneals to the template strand better and amplifies more efficiently compared to primer P10.

During the λ-exonuclease digestion, as the number of AHP dU units increased in the phosphorylated strand, digestion became progressively slower, and a faint duplex band was present in the 100% AHP dU single-stranded sample (Figure 4.34.B, lane 5^s). Nonetheless, ssDNA products were confirmed by re-annealing of the modified strand (SP) with unmodified complementary C13 (lane 6^s) to form hybrid modified/unmodified DNA duplexes (Figure 4.34.B, re-annealing (C13) lanes).

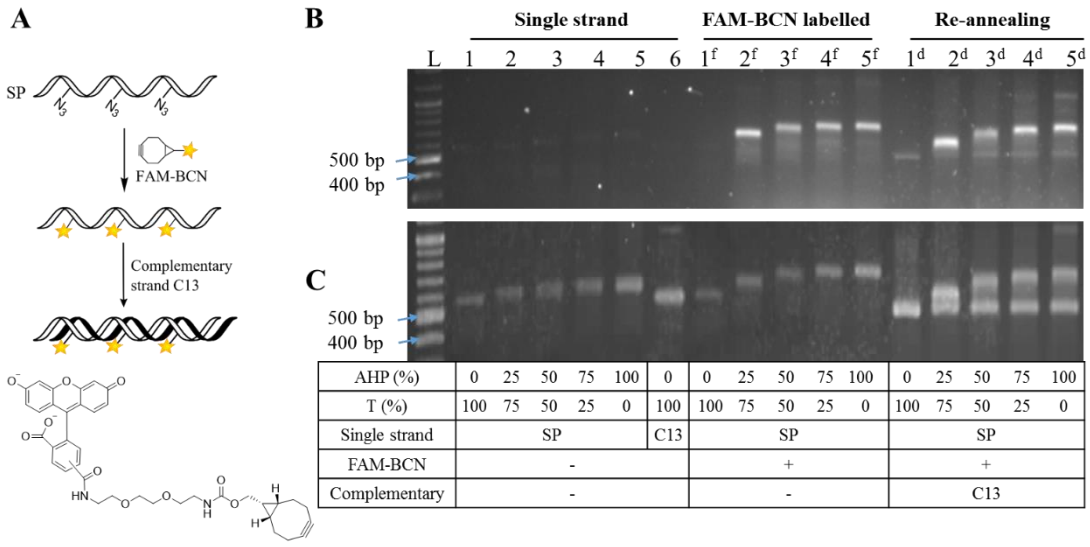


AHP (%)	0	25	50	75	100	0	0	25	50	75	100	0	0	25	50	75	100
T (%)	100	75	50	25	0	100	100	75	50	25	0	100	100	75	50	25	0
Product	AP1					AP2	SP					C13	SP				
Complementary	-						-						C13				

Figure 4.34. λ–exonuclease digestion of APCR amplicons (T11, 10:1 of P10:P9$_p$/P9:P10$_p$). **A.** Scheme for digestion and re-annealing of the AHP-modified single strands (SP) with the complementary unmodified strand C13. **B.** Agarose gel analysis of digestion and re-annealing. Lane L, 100 bp ladder; lanes 1/1^s/1^d and 6/6^s, unmodified products; lanes 2/2^s/2^d to 5/5^s/5^d, 25% to 100% AHP-modified products. D = duplex products; S = single strands. (1.5% agarose gel containing ethidium bromide)

FAM-BCN was then used to label the AHP-modified long ssDNA (523-mer), followed by re-annealing with complementary strand C13. As expected, fluorescent bands for the labelled single strands and the re-formed duplexes were visualised without the need for dye-staining (Figure 4.35.B). These results show that this strategy can be used to prepare very long single-stranded fluorescent probes. It could be applied to the Cy3-BCN or Cy5-BCN dyes if additional hydrophilic modifications can be attached to the

fluorophores to improve their aqueous solubility such as the oligothymine linker used previously or sulfonic acid groups directly attached to the dye.[203]



AHP (%)	0	25	50	75	100	0	0	25	50	75	100	0	25	50	75	100
T (%)	100	75	50	25	0	100	100	75	50	25	0	100	75	50	25	0
Single strand	SP					C13	SP					SP				
FAM-BCN	-						+					+				
Complementary	-						-					C13				

Figure 4.35. FAM-BCN fluorescent labelling of AHP-modified single strands (SP). **A.** Scheme for labelling and re-annealing. Agarose gel analysis before staining (**B**) and after staining with ethidium bromide (**C**). Lane L, 100 bp ladder; lanes $1/1^{f}/1^{d}$ and 6, unmodified products, controls; lanes $2/2^{f}/2^{d}$ to $5/5^{f}/5^{d}$, 25% to 100% AHP-modified products. (1.5% agarose gel) The labelling reactions were also analysed by agarose gel containing ethidium bromide (Appendix 8.9).

4.5.2 Synthesis of dual labelled single-stranded probes using streptavidin magnetic bead separation

The previous fluorescent labelling of ssDNA, generated by exonuclease digestion of APCR or PCR products involves laborious purification procedures between each step. To simplify the digestion and labelling strategy, another widely used method to prepare single-stranded oligonucleotides, streptavidin-coated magnetic bead separation is described in this section.[229, 230] Streptavidin is a protein from *Streptomyces avidinii* and has an extremely high affinity for biotin *via* non-covalent binding.[231] By using one 5′-biotin modified PCR primer and one unmodified primer, PCR amplicons can be captured on streptavidin magnetic beads (Figure 4.36). Alkaline or heat denaturation can

then be used to denature the DNA duplex. The un-biotinylated strand is liberated into solution and can be collected, while the biotinylated complementary single strand remains attached to the magnetic beads. Furthermore, the unwanted excess triphosphates from the PCR can be removed easily by washing the streptavidin magnetic beads.

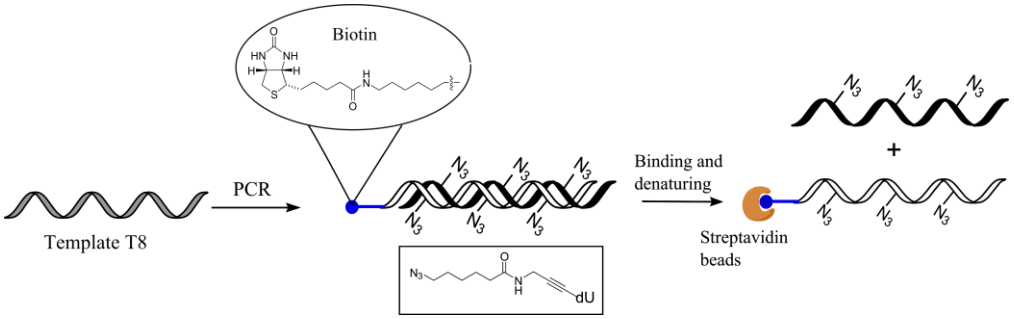
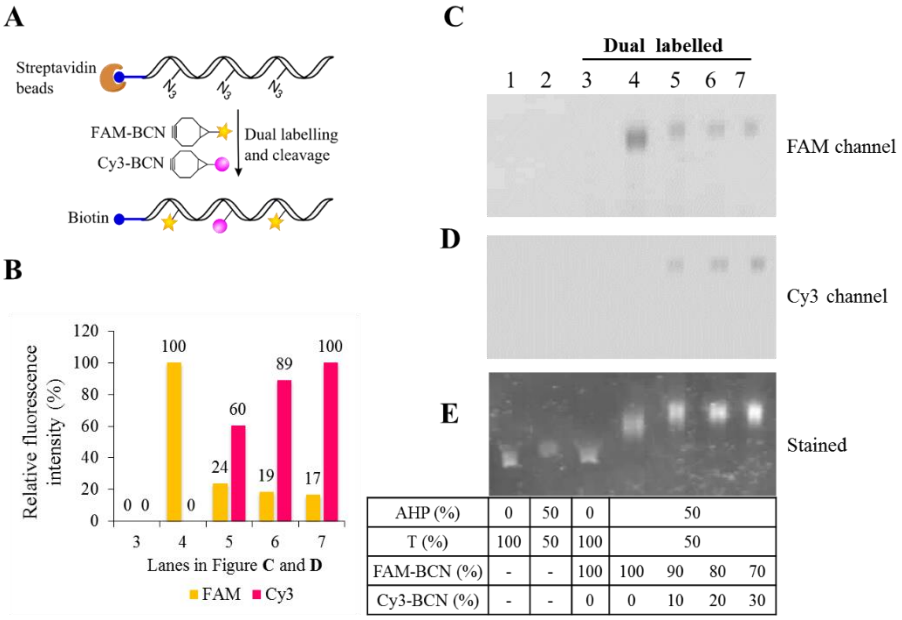


Figure 4.36. Streptavidin magnetic bead separation to make single-stranded products from AHP-modified PCR amplicons with biotin modified strands.

For the preparation of fluorescent probes, labelling the biotinylated strand with fluorophores is a better strategy than labelling the un-biotinylated strand due to the simpler purification by washing the streptavidin magnetic beads to remove excess reagents. Therefore, PCR using AHP dUTP with dTTP in a 1:1 ratio was carried out in the presence of template T8, one normal primer (P6) and one 5′-biotin modified primer ($P7_b$, Table 4.4). PCR amplicons were bound to the streptavidin beads before sodium hydroxide denaturation. The remaining single strands on the magnetic beads were labelled with two dyes simultaneously. Initially dual labelling was examined using different combinations of FAM-BCN and Cy3-BCN (Figure 4.37.A). The unmodified product was used as the negative control for fluorescent labelling. After completion of the labelling reaction, samples were washed with 70% ethanol. The labelled oligonucleotides were then cleaved from the beads by treatment with formamide. The dual labelled products were first analysed by denaturing PAGE (Figure 4.37.C, D and E)

and quantified using ImageJ software. As observed previously, a sharp 4.2-fold decrease in FAM-fluorescence intensity occurred upon reducing the ratio of FAM-BCN from 100% to 90%, whilst the Cy3 signal intensity increased progressively as Cy3-BCN was changed from 10% to 30% (Figure 4.37.B). This implies that the product from the bead-based labelling strategy is not significantly different to that obtained by APCR and λ-exonuclease digestion.

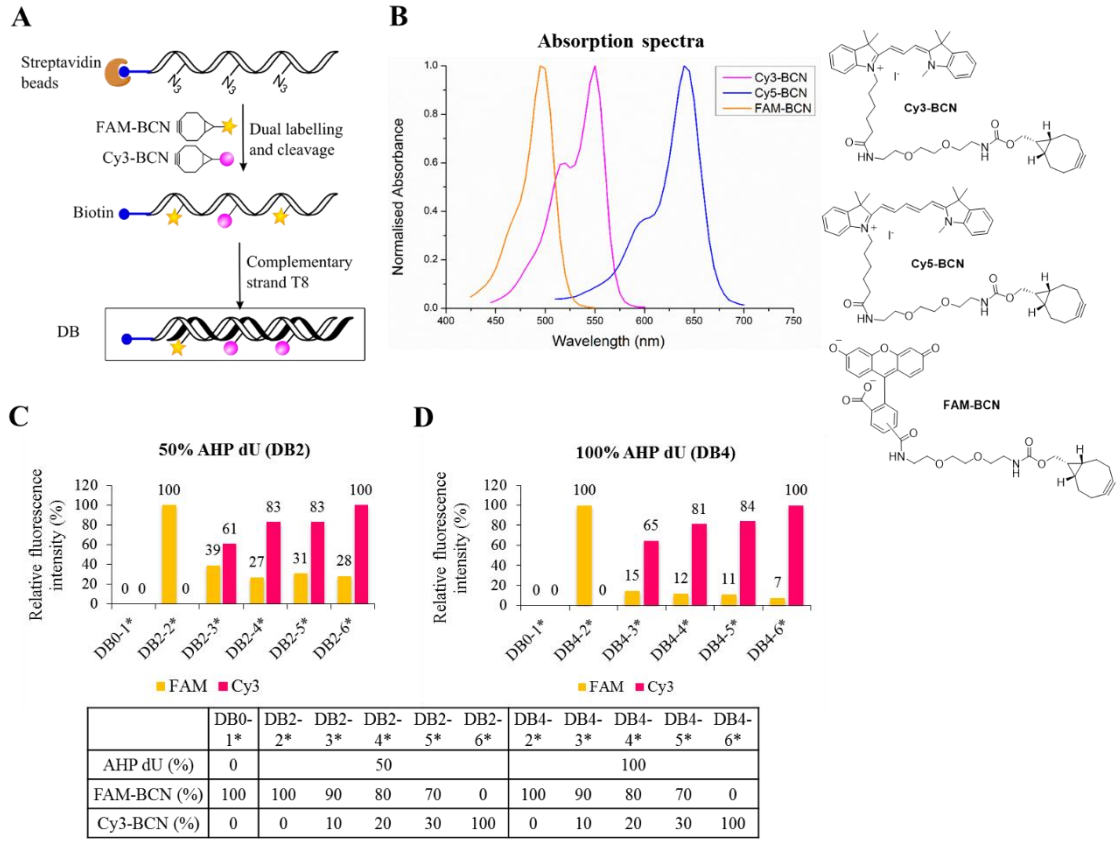


AHP (%)	0	50	0	50			
T (%)	100	50	100	50			
FAM-BCN (%)	-	-	100	100	90	80	70
Cy3-BCN (%)	-	-	0	0	10	20	30

Figure 4.37. FAM-BCN and Cy3-BCN dual labelling of 50% AHP-modified single strands from streptavidin magnetic separation. **A.** Scheme for dual labelling. **B.** Relative fluorescence intensity for single-stranded product bands in FAM and Cy3 channels (lanes 3 to 7 in Figure **C** and **D**). Denaturing PAGE analysis visualised before staining by FAM channel (**C**), Cy3 channel (**D**) and after staining with SYBR Gold (**E**). Lane 1, unmodified product; lane 2, unlabelled 50% AHP-modified product; lane 3, unmodified product mixed with 100% FAM-BCN, negative control; lanes 4 to 7, 50% AHP-modified products labelled with different combinations of FAM-BCN (100% to 70%) and Cy3-BCN (0% to 30%) respectively. (8% denaturing PAGE)

To check the purity and annealing properties of the dual labelled products, the labelled 50% AHP-modified ssDNA was annealed to the complementary strand (T8) and analysed by native PAGE (Figure 4.38.A and Appendix 8.10.1). The fluorescent quantification of

the re-annealed duplexes (DB) showed the same trend as the denaturing PAGE results (Figure 4.37.B and Figure 4.38.C). The signal observed in the FAM channel for 100% Cy3-BCN labelled product (DB2-6*) was caused by the overlapping absorption spectra of FAM and Cy3 (Figure 4.38.B and C).



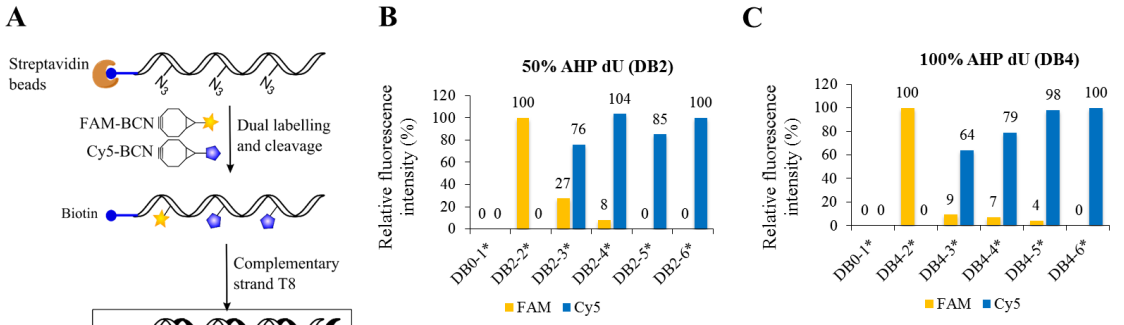
	DB0-1*	DB2-2*	DB2-3*	DB2-4*	DB2-5*	DB2-6*	DB4-2*	DB4-3*	DB4-4*	DB4-5*	DB4-6*
AHP dU (%)	0	50					100				
FAM-BCN (%)	100	100	90	80	70	0	100	90	80	70	0
Cy3-BCN (%)	0	0	10	20	30	100	0	10	20	30	100

Figure 4.38. FAM-BCN and Cy3-BCN dual labelling of 50% and 100% AHP-modified single strands from streptavidin magnetic separation. **A.** Scheme for dual labelling and re-annealing. **B.** Absorption spectra of FAM-BCN (75 mM TEAB buffer, pH 8.0), Cy3-BCN (EtOH) and Cy5-BCN (EtOH). **C** and **D.** Relative fluorescence intensity for the re-annealed duplex products in FAM and Cy3 channels (DB, original native PAGE see Appendix 8.10.1 and 8.10.2). DB0-1*, unmodified product with 100% FAM-BCN, negative control; DB2-2*/DB4-2* to DB2-6*/DB4-6*, AHP-modified products labelled with different combinations of FAM-BCN (100% to 0%) and Cy3-BCN (0% to 100%) respectively.

In comparison with the 50% AHP-modified products, the fully AHP dU products (containing eleven modifications per strand) afforded more intense (brighter) ssDNA probes as shown in the single fluorophore labelling reactions (Figure 4.31). Dual labelling

reactions were therefore carried out using the 100% AHP dU products. Surprisingly upon labelling, a 6.7-fold reduction in FAM fluorescence was observed when FAM-BCN was decreased from 100% to 90% (Figure 4.38.D, DB4-2* and DB4-3*). This reduction is larger than that of the 50% AHP dU products (2.7-fold, Figure 4.38.C, DB2-2* and DB2-3*). On the other hand, the Cy3 fluorescence of both 50% and 100% AHP dU products increased progressively as more Cy3-BCN was used. This may be due to the lower activity of FAM-BCN relative to Cy3-BCN.

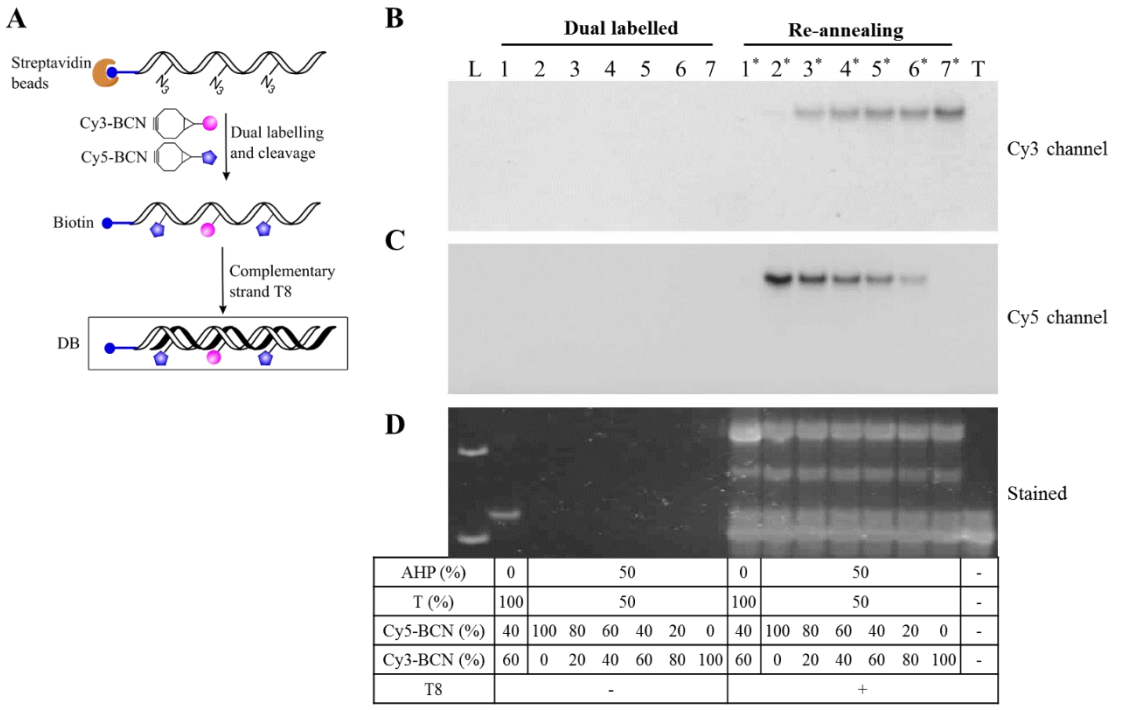
The dual labelling strategy was then expanded to a FAM-BCN and Cy5-BCN system (Figure 4.39), in which the two fluorophores do not have overlapping absorption spectra (Figure 4.38.B). Similar results to those obtained for the FAM-Cy3 system were observed, thus indicating that a combination of the negatively charged FAM-BCN and positively charged Cy3 or Cy5 is not an ideal combination. This is possibly due to the unfavourable electrostatics between negatively charged FAM and the anionic oligonucleotide backbone.



	DB0-1*	DB2-2*	DB2-3*	DB2-4*	DB2-5*	DB2-6*	DB4-2*	DB4-3*	DB4-4*	DB4-5*	DB4-6*
AHP dU (%)	0	50					100				
FAM-BCN (%)	70	100	90	80	70	0	100	90	80	70	0
Cy5-BCN (%)	30	0	10	20	30	100	0	10	20	30	100

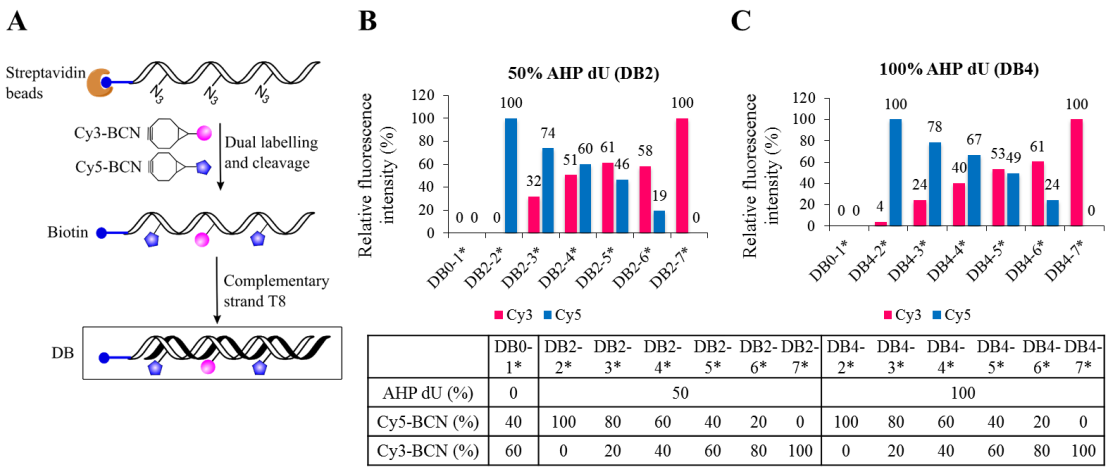
Figure 4.39. FAM-BCN and Cy5-BCN dual labelling of 50% and 100% AHP-modified single strands from streptavidin magnetic separation. **A.** Scheme for dual labelling and re-annealing. **B** and **C.** Relative fluorescence intensity for the re-annealed duplex products in FAM and Cy5 channels (DB, original native PAGE see Appendix 8.10.3 and 8.10.4). DB0-1*, unmodified product with 70% FAM-BCN, negative control; DB2-2*/DB4-2* to DB2-6*/DB4-6*, AHP-modified products labelled with different combinations of FAM-BCN (100% to 0%) and Cy5-BCN (0% to 100%) respectively.

Consequently, dual labelling using Cy3-BCN and Cy5-BCN was investigated (Figure 4.40.A). No fluorescent bands were observed for the dual labelled single strands, presumably due to the collisional fluorescence quenching between each fluorophore in the unstructured ssDNA (Figure 4.40.B and C). In contrast, when annealed to the complementary T8 strand, the labelled fluorophores are held apart by the rigid duplex, reducing collisional quenching and giving rise to intense bands on the native PAGE. This favourable 'off-on' system upon duplex formation should provide better detection for hybridisation applications by reducing the background signal. Most importantly, the fluorescence of the re-formed duplex demonstrated a progressive decrease in Cy5 signal when the percentage of Cy5-BCN was varied from 100% to 0%, while the Cy3 fluorescence intensity increased in proportion to the percentage of Cy3-BCN used (Figure 4.41.B and C). This clearly demonstrates the feasibility of making fluorescent probes with subtly different fluorescence characteristics using only two fluorophore labels.



AHP (%)	0	50						0	50						-
T (%)	100	50						100	50						-
Cy5-BCN (%)	40	100	80	60	40	20	0	40	100	80	60	40	20	0	-
Cy3-BCN (%)	60	0	20	40	60	80	100	60	0	20	40	60	80	100	-
T8	-							+							

Figure 4.40. Dual labelling of 50% AHP-modified single strands from streptavidin magnetic separation with different combinations of Cy5-BCN (100% to 0%) and Cy3-BCN (0% to 100%). **A.** Scheme for dual labelling and re-annealing. Native PAGE analysis visualised before staining by Cy3 channel (**B**), Cy5 channel (**C**) and after staining with SYBR Gold (**D**). Lane L, 25 bp ladder; lane 1/1*, unmodified product, negative control; lanes 2/2* to 7/7*, 50% AHP-modified products; lane T, complementary T8. (12% native PAGE)



	DB0-1*	DB2-2*	DB2-3*	DB2-4*	DB2-5*	DB2-6*	DB2-7*	DB4-2*	DB4-3*	DB4-4*	DB4-5*	DB4-6*	DB4-7*
AHP dU (%)	0	50						100					
Cy5-BCN (%)	40	100	80	60	40	20	0	100	80	60	40	20	0
Cy3-BCN (%)	60	0	20	40	60	80	100	0	20	40	60	80	100

Figure 4.41. Cy3-BCN and Cy5-BCN dual labelling of 50% and 100% AHP-modified single strands from streptavidin magnetic separation. **A.** Scheme for dual labelling and re-annealing. **B** and **C.** Relative fluorescence intensity for the re-annealed duplex products in Cy3 and Cy5 channels (DB, original native PAGE see Figure 4.40 and Appendix 8.10.5). DB0-1*, unmodified product with 40% Cy5-BCN, negative control; DB2-2*/DB4-2* to DB2-7*/DB4-7*, AHP-modified products labelled with different combinations of Cy5-BCN (100% to 0%) and Cy3-BCN (0% to 100%) respectively.

To summarise, a broad range of single-colour and dual-colour ssDNA probes (81-mer and 523-mer) with different fluorescent profiles were successfully obtained. These could potentially be used to detect multiple targets simultaneously in FISH applications. The advantage of using streptavidin magnetic bead separation over λ-exonuclease digestion is clear in that it simplifies purification by enabling excess reagents to be washed away; allows densely labelled products bound to the beads thereby preventing aggregation due to the presence of multiple hydrophobic cationic Cy3 and Cy5 labels; and provides an additional pull down tag in the form of biotin.

4.6 Reverse transcription using azide dUTPs

Finally, to investigate the biocompatibility of AM and AHP dUTPs with RNA dependent enzymes, reverse transcription was carried out (Figure 4.42). This enables the synthesis

of modified DNA products from RNA templates, and is very important for RNA detection as well as the RNA-aptamer field.[232-234]

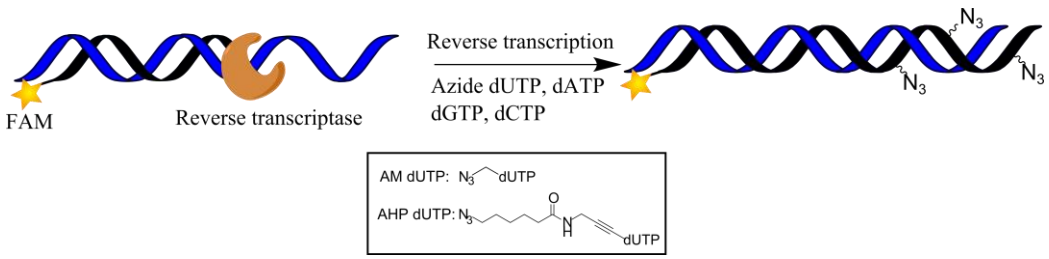


Figure 4.42. Reverse transcription using a RNA template and azide dUTPs.

The AM and AHP dUTPs were tested for incorporation against two synthetic RNA templates, carrying the same sequence as the primer extension templates (Figure 4.43 and Figure 4.44). Moloney Murine Leukemia Virus reverse transcriptase (M-MuLV, RNase H^-) without RNase activity, and Avian Myeloblastosis Virus reverse transcriptase (AMV) were employed. AM and AHP dUTPs were incorporated successfully in the presence of both templates (T4 and T5) and produced fully extended products with M-MuLV and AMV reverse transcriptases (Figure 4.43 and Figure 4.44). M-MuLV (RNase H^-) reactions gave nearly quantitative conversion to the targeted products, whereas with AMV some residual primers remained (Figure 4.43 and Figure 4.44). Therefore M-MuLV (RNase H^-) appears to be a better choice and with further work this could be exploited for DNA functionalisation. For example, these azide-modified RNA/DNA duplexes are known to be good substrates for RNase, which can selectively digest the RNA strand and leave the complementary modified DNA strand intact.[204, 205] This could be investigated as a future strategy to prepare ssDNA probes.

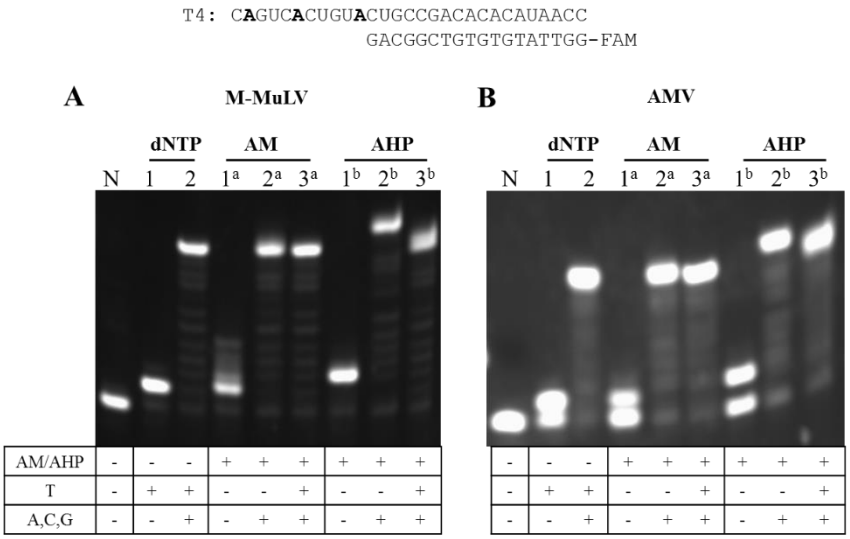


Figure 4.43. Reverse transcription with azide dUTPs, primer P3 and RNA template T4 using M-MuLV (RNase H$^-$, **A**) and AMV (**B**) reverse transcriptases. 3.2 nmol of each triphosphate was used unless otherwise stated. Lane N, reaction without triphosphates; lane 1/1^a/1^b, dTTP or modified dUTP; lane 2/2^a/2^b, dTTP or modified dUTP + dATP + dCTP + dGTP; lane 3^a/3^b, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP. (20% denaturing PAGE)

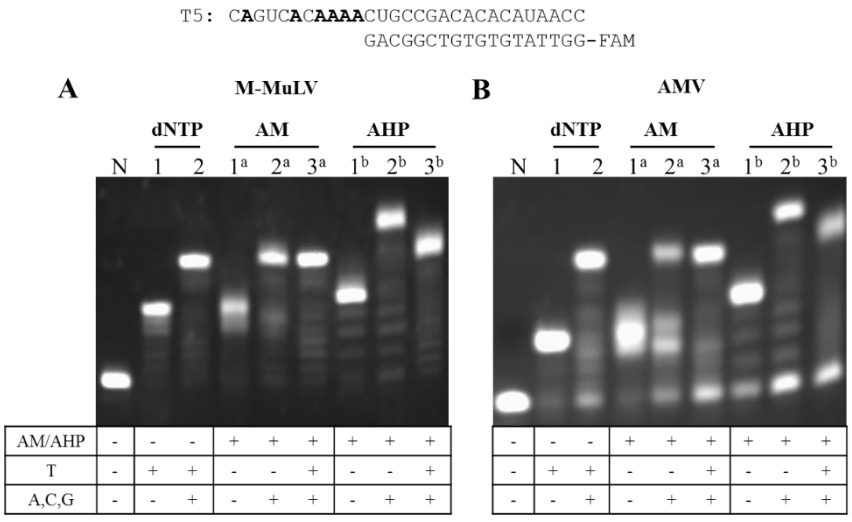


Figure 4.44. Reverse transcription with azide dUTPs, primer P3 and RNA template T5 using M-MuLV (RNase H$^-$, **A**) and AMV (**B**) reverse transcriptases. 3.2 nmol of each triphosphate was used unless otherwise stated. Lane N, reaction without triphosphates; lane 1/1^a/1^b, dTTP or modified dUTP; lane 2/2^a/2^b, dTTP or modified dUTP + dATP + dCTP + dGTP; lane 3^a/3^b, dTTP (1.6 nmol) + modified dUTP (1.6 nmol) + dATP + dCTP + dGTP. (20% PAGE)

4.7 Conclusions

Primer extension, reverse transcription and PCR reactions were successfully used to prepare azide-modified DNA strands using AM and AHP dUTPs. PCR with AHP dUTP was more efficient than AM dUTP and gave correct products with perfect sequence fidelity beyond 500 bp in length. The AHP-modified product was more thermal-stable than the unmodified product, whereas the AM modification decreased DNA duplex stability. SPAAC post-labelling of AHP-modified primer extension products with Cy3-BCN was more efficient than AM dU due to the longer linker between the azide and uracil base. A variety of BCN functionalised fluorophores were used to label the AHP-modified primer extension and PCR products; Cy3-BCN labelling was more efficient than Cy5-BCN and FAM-BCN. A new dual labelling strategy using different combinations of Cy3-BCN and FAM-BCN was developed to label the long AHP dU PCR products.

AHP-modified single strands were successfully generated from PCR amplicons *via* λ-exonuclease digestion and streptavidin magnetic bead separation. From the reaction using different ratios of AHP dUTP to dTTP, ssDNA can be labelled with single fluorophores to yield fluorescent probes with different fluorescent intensities. Importantly, two fluorophores can be used in a dual-labelling strategy, giving a variety of novel probes with distinct fluorescent spectra. These probes could be potentially used in FISH detection to target multiple gene sequences and will be studied in the future. In the context of this work, the successful *in vitro* incorporation of AHP dUTP by Klenow polymerase suggests that the corresponding nucleoside might be a candidate for incorporation into and imaging of genomic DNA in a cellular context.[157]

Chapter 5 Imaging DNA synthesis in cells

5 Imaging DNA synthesis in cells

5.1 Background and aims

Nascent DNA and RNA synthesis has been monitored in proliferating cells using ethynyl modified nucleosides by fluorescence staining *via* the copper-catalysed azide-alkyne cycloaddition reaction (CuAAC, Chapter 1, Section 1.5).[38, 97, 153, 155] Ethynyl nucleoside (EdU, EU and EdC) are converted to their triphosphates by various kinases in cells and are incorporated by cellular DNA/RNA polymerase enzymes into genomic DNA and RNA. After cell fixation these can be reacted with azide derivatives of fluorophores and visualised. The high concentrations of copper catalyst used in the CuAAC labelling reactions are toxic to cells and causes DNA degradation, restricting its applications in live cells.[235-237] To avoid this problem, the inverse electron demand Diels-Alder reaction (IEDDA) using vinyl deoxyuridine (VdU) and the strain-promoted azide-alkyne cycloaddition reaction (SPAAC) using azidomethyl deoxyuridine (AM dU) have recently been found to be efficient copper-free alternatives.[156, 157] Functionalised genomic DNA can be fluorescently labelled with tetrazine or cyclooctyne modified fluorophores using these reactions. This provides potential tools to visualise DNA and RNA in living systems. Until now, nucleosides modified with *trans*-cyclooctene (a more reactive dienophile than a vinyl group for the IEDDA reaction) have not been reported for use in DNA imaging in cells. In this chapter the utility of *trans*-cyclooctene modified nucleosides for the above application is explored. Modified nucleosides containing azides or fluorophores are also examined to expand the cellular DNA imaging toolkit. The long term aim is to provide diverse methodologies to aid the understanding of DNA replication, transcription and related processes in cells and organisms.

5.2 Synthesis of modified nucleosides

5.2.1 Synthesis of azide-modified nucleosides

Compared to alkynes, azides can be functionalised using more varied chemistries, *i.e.* the Staudinger, CuAAC and SPAAC reactions.[75, 107, 115] 5-Azidomethyl-2′-deoxyuridine **4.1** (AM dU, Figure 5.1) has been previously utilised in different cell lines to prepare modified genomic DNA which was subsequently reacted with fluorophores tagged with bicyclo[6.1.0]nonyne (BCN) *via* the SPAAC reaction.[157] Therefore, AM dU **4.1** was used as a comparator for the new methodology described in this chapter. The azido-RNA nucleosides 5-azidomethyl uridine **5.1** (AM U) and 5-azidopropyl uridine **5.2** (APr U), were also studied to investigate their utility in cellular RNA imaging (Figure 5.1).[110, 238] This work was undertaken because successful *in vitro* incorporation of 5-azidopropyl uridine-5′-triphosphate (APr UTP) into RNA has been reported by Srivatsan *et al.*,[110] and the nucleoside APr U is likely to be converted to its triphosphate and incorporated into RNA by cellular enzymes. All three compounds in Figure 5.1 were synthesised by Dr Marta Gerowska in the Brown group and were purified by reversed-phase high performance liquid chromatography (RP-HPLC).

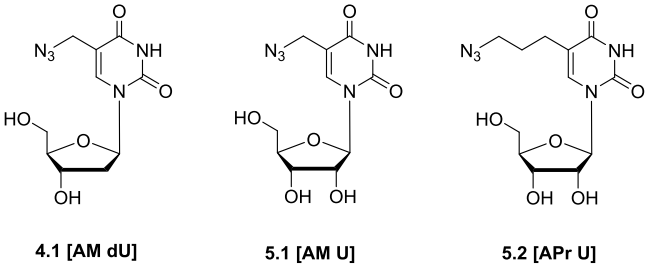
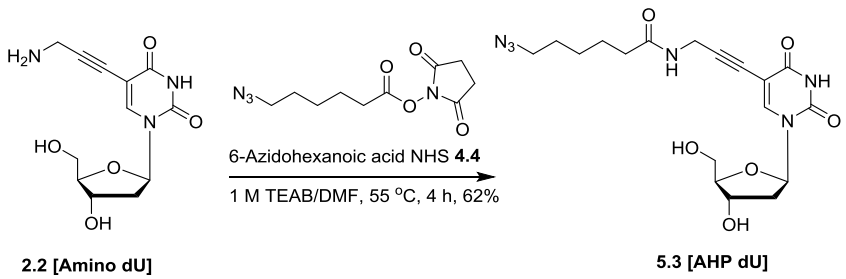


Figure 5.1. Chemical structures of AM dU, AM U and APr U.

As shown in Chapter 4, 5-azidohexanamidopropargyl-2′-deoxyuridine-5′-triphosphate **4.5** (AHP dUTP) is compatible with enzymatic nucleic acid copying and is a better substrate for KOD DNA polymerase than AM dUTP **4.2**; the respective longer chain azide-modified DNA products also demonstrated more efficient fluorescent labelling *via* the SPAAC reaction. Due to these favourable properties we reasoned that AHP dU nucleoside would be more efficiently utilised by cellular DNA polymerase enzymes and fluorescently labelled than the previously reported AM dU.[157] Therefore, 5-azidohexanamidopropargyl-2′-deoxyuridine **5.3** (AHP dU) was synthesised and studied with cellular DNA imaging applications in mind. The synthesis involved labelling 5-aminopropargyl-2′-deoxyuridine **2.2** (amino dU) with 6-azidohexanoic acid NHS active ester **4.4** (Scheme 5.1). These nucleosides with the azide group attached to the uracil base *via* different linkages enable studies to determine their suitability in cellular nucleic acid imaging.



Scheme 5.1. Synthesis of AHP dU.

5.2.2 Synthesis of *trans*-cyclooctene-modified nucleosides

Two *trans*-cyclooctene (TCO) modified nucleosides, which correspond to the modified triphosphate TCO-S dUTP **3.1** and TCO-L dUTP **3.2** investigated in Chapter 3, were also prepared from amino dU **2.2**. These TCO dUTPs demonstrated efficient enzymatic incorporation using different DNA and RNA templates in primer extension, PCR and

reverse transcription. If the corresponding nucleosides are compatible with cellular enzymes, they could be used to visualise genomic DNA by fluorescence staining *via* the IEDDA reaction.[131] Commercially available *trans*-cyclooctene NHS carbonate **3.3** and ester **3.4** were coupled to amino dU **2.2** to afford TCO-S dU **5.4** with a short linker and TCO-L dU **5.5** with a longer tetraethylene glycol linker (Figure 5.2). These two nucleosides were synthesised for subsequent evaluation for genomic DNA labelling in cells.

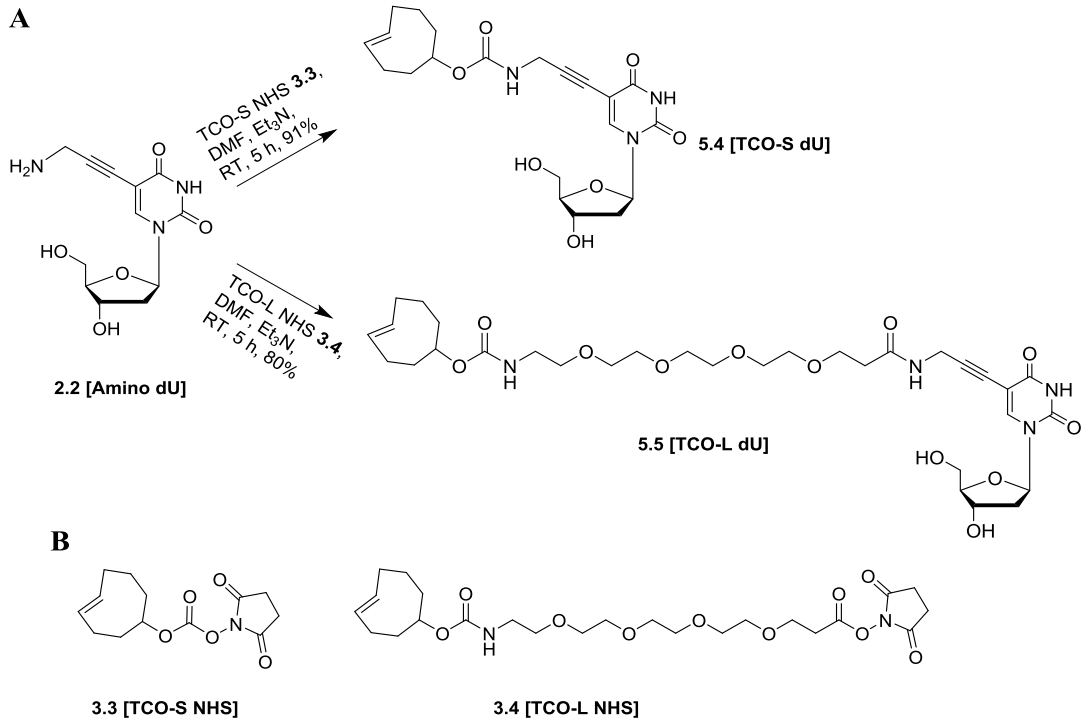
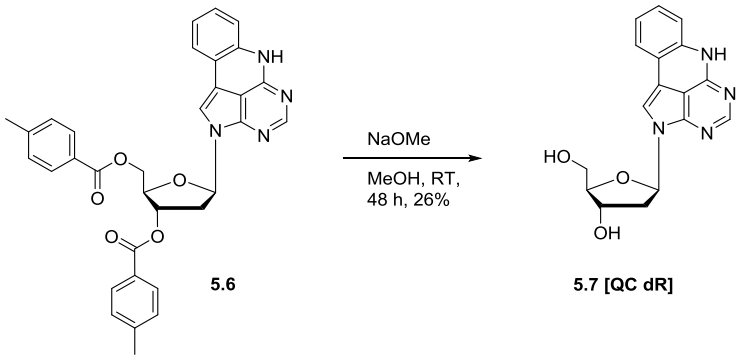


Figure 5.2. **A.** Synthesis of TCO-modified deoxyuridines. **B.** Structures of TCO-S NHS carbonate and TCO-L NHS ester.

5.2.3 Synthesis of a fluorescent nucleobase analogue

If a compound is both a fluorescent molecule and a base analogue, the replicating DNA can be detected directly without an extra fluorescent labelling step.[50] Fluorescent base analogues are structurally related to and can base pair with natural nucleobases.[50] In

fluorescent probes containing this type of analogue, the fluorescent dye is attached rigidly and positioned close to the centre of the duplex. This enables higher resolution determination of distances (e.g. by FRET) than commonly used fluorophores (those that are not involved in base pairing) or more conventional fluorescently-modified bases which have fluorophores attached by long linkers. An example of a fluorescent base analogue is quadracyclic adenine **5.7** (QC dR, $\lambda_{ex,water}$ = 337 nm, $\lambda_{em,water}$ = 456 nm) which pairs specifically with thymidine to form a stable A:T base pair analogue.[48, 239] Wilhelmsson and co-workers showed that this analogue has a higher quantum yield in DNA duplexes than the free monomer **5.7**,[48] and therefore it is likely to give a good signal in cellular DNA imaging applications. This quadracyclic adenine analogue has been suggested to have better cell membrane permeability due to its higher lipophilicity compared to natural adenine.[239] The nucleoside **5.7** was therefore synthesised for testing in cellular DNA visualisation (Scheme 5.2). The starting material, 3′,5′-toluoyl quadracyclic adenine-2′-deoxyriboside **5.6** was previously synthesised in the Brown group.[48] The protecting groups were removed using sodium methoxide to afford compound **5.7**.



Scheme 5.2. Synthesis of quadracyclic adenine analogue.[48]

5.3 Cellular DNA imaging using modified nucleosides

The above azide and cyclooctene modified nucleosides (**4.1**, **5.1** to **5.5**) and a fluorescent analogue (**5.7**) were used to label DNA and RNA in proliferating cells. Commercially available ethynyl dU **5.8** (EdU, Figure 5.3.A) and AM dU **4.1** were used as controls as there are previous reports for their use in cellular/nuclear DNA labelling.[38, 157] Adenocarcinoma human alveolar basal epithelial cell line A549 was used for this application.[240, 241] This cell line is hypotriploid with 66 chromosomes, and grows as an adherent monolayer. It has been demonstrated with good compatibility for cellular DNA imaging application using modified nucleosides.[157] Cells (5,000 cells) were grown for 24 h on the coverslips then treated with fresh nutrient carrying modified nucleosides for a further 24 h. The cells were fixed using 3.7% formaldehyde and treated with 0.5% Triton® X-100 to improve the membrane permeability for subsequent fluorescent staining. Different chemical reactions (CuAAC, SPAAC and IEDDA) were used to stain the cells with a variety of fluorophores. After that, 4′,6-diamidino-2-phenylindole **5.9** (DAPI, Figure 5.3.B) was used to stain the cell to provide a positive control for nuclear imaging and analysed by confocal fluorescence microscopy. The above protocol is the procedure suggested by the supplier of the EdU staining kit (Invitrogen). An alternative protocol described by Neef and Luedtke, where the staining step precedes Triton® X-100 treatment, was also tested but failed to give positive results for either nucleoside.[157] The fluorescent analogue QC dR **5.7** was tested in collaboration with Dr Afaf El-Sagheer and did not give positive results.

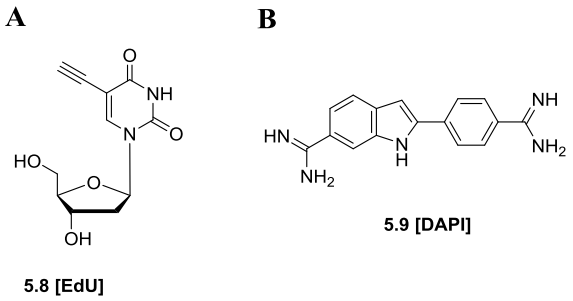


Figure 5.3. Chemical structures of ethynyl dU (**A**) and 4′,6-diamidino-2-phenylindole (DAPI, **B**).

Unfortunately, only the EdU control sample gave an intense fluorescent image of nuclear DNA using Cy3-azide and AlexaFluor-594-azide *via* the CuAAC labelling (Figure 5.4 and Figure 5.5). All the other modified DNA and RNA nucleosides including the AM dU control gave negative results in the fluorescent labelling experiments, with no fluorescence observed in the nucleus. The AM dU only showed fluorescent labelling in the cytoplasm, which is in contrast with the previously reported studies (Figure 5.6).[157] Representative results from experiments with azide and *trans*-cyclooctene modified nucleosides are shown in Figures 5.6 to 5.8.

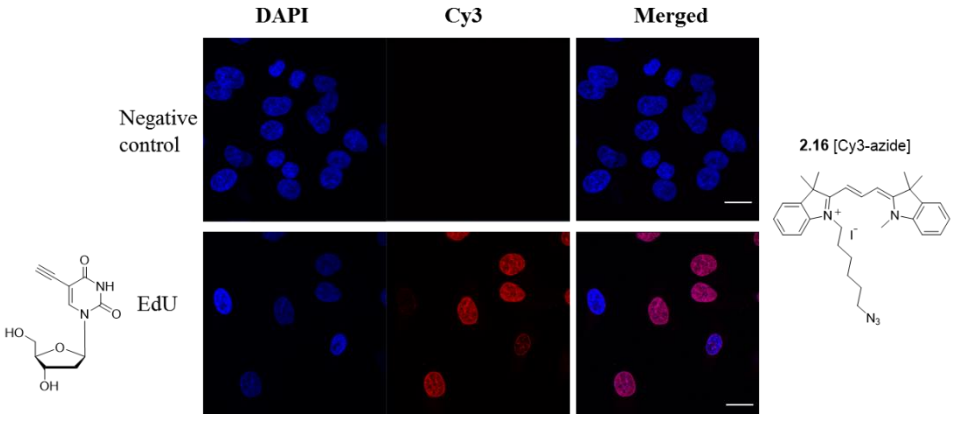


Figure 5.4. Cells treated without modified nucleosides (negative control) or with EdU (10 μM), and stained with Cy3-azide using the CuAAC reaction. 20 micron scale bar shown.

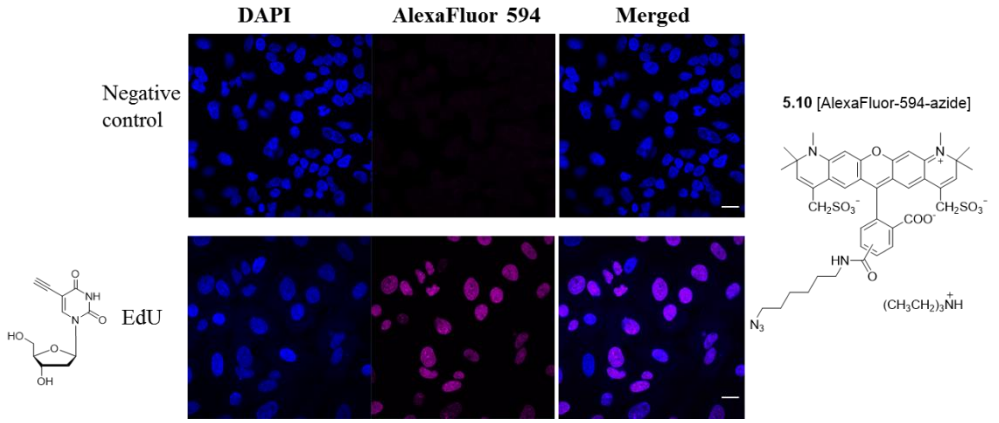


Figure 5.5. Cells treated without modified nucleosides (negative control) or with EdU (10 μM), and stained with AlexaFluor-594-azide using the CuAAC reaction. 20 micron scale bar shown.

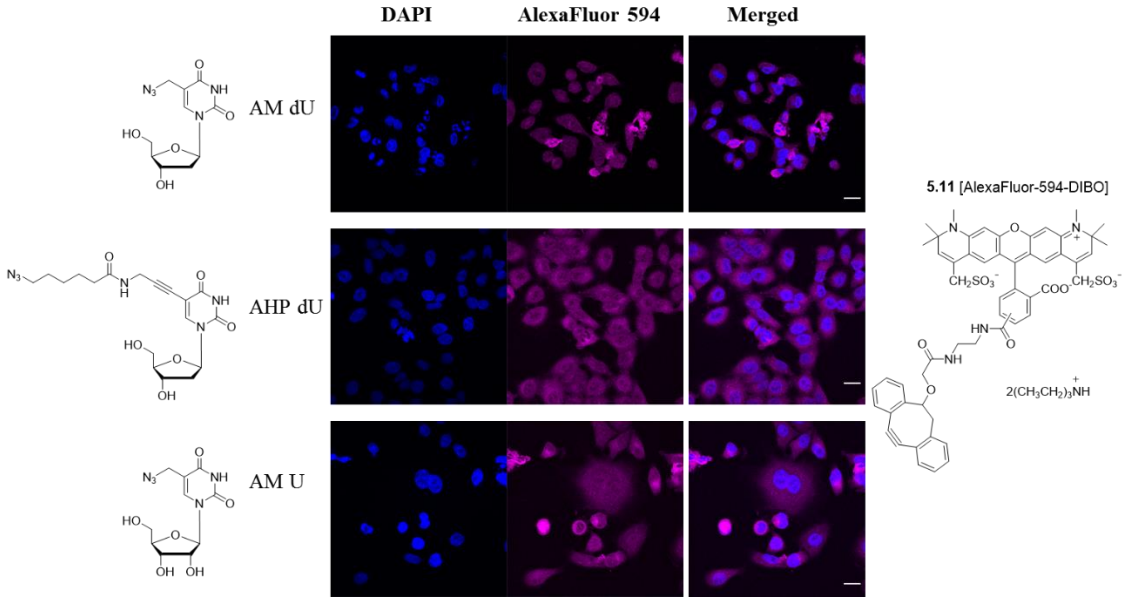


Figure 5.6. Cells treated with different azido deoxyuridines and uridine (100 μM), and stained with AlexFluor-594-DIBO using the SPAAC reaction. 20 micron scale bar shown.

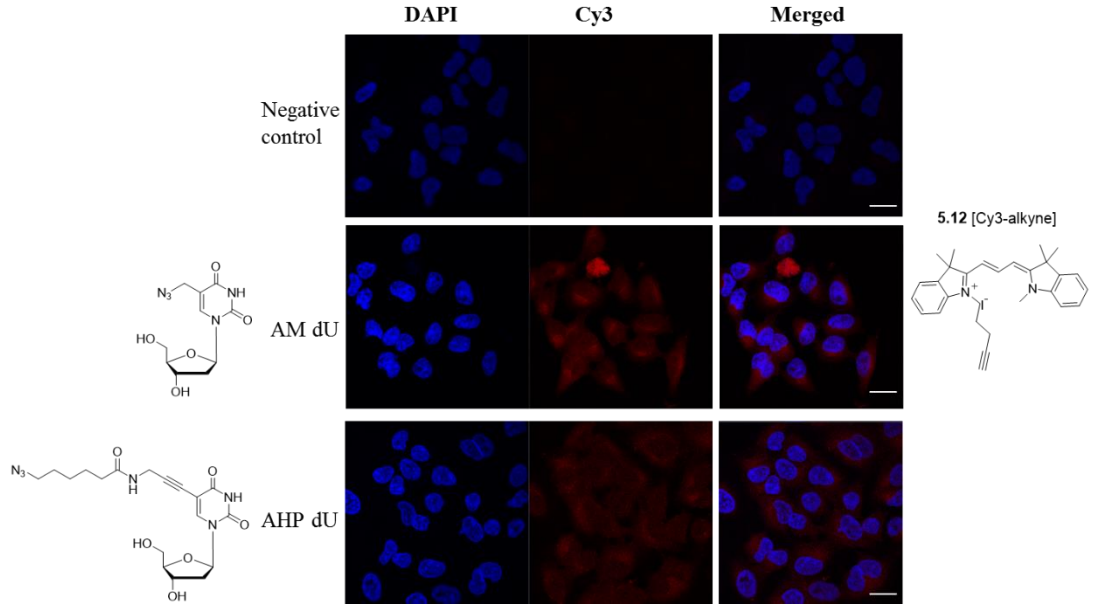


Figure 5.7. Cells treated without modified nucleosides (negative control) or with different azido deoxyuridines (100 μM), and stained with Cy3-alkyne using the CuAAC reaction. 20 micron scale bar shown.

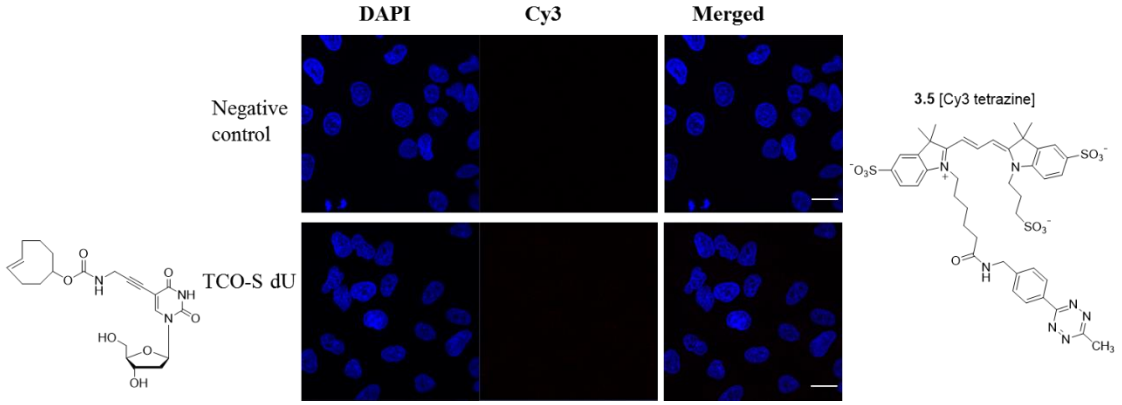


Figure 5.8. Cells treated without modified nucleosides (negative control) or with TCO-S dU (100 μM) and stained with Cy3-tetrazine using the IEDDA reaction. 20 micron scale bar shown.

The putative azido DNA and RNA samples were initially stained with commercially available AlexaFluor-594-dibenzocyclooctyne (AlexaFluor-594-DIBO) *via* the SPAAC reaction (Figure 5.6). The reason for the negative results might be that DIBO reacts with thiols in peptides and proteins that are present in the cell (Figure 5.9).[130, 242]

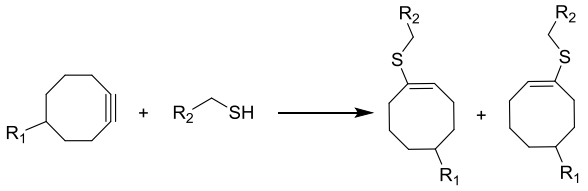


Figure 5.9. General scheme for cyclooctyne reaction with thiol.[130]

To test the above assumption, Cy3-alkyne (provided by Dr Marta Gerowska in the Brown group) was used to label DNA (presumably containing AM dU and AHP dU) *via* the CuAAC reaction (Figure 5.7). Weaker cytoplasm staining was observed compared to the SPAAC staining method (Figure 5.6). Salic and Mitchison reported that cellular DNA imaging using 5-azido-2′-deoxyuridine gave a stronger background signal than when DNA containing EdU was labelled with fluorophores using the CuAAC reaction.[38] The above results suggest that the side-reaction between cyclooctyne and thiols might not be the main reason for the intense cytoplasm staining, and that the modified nucleosides might not be phosphorylated efficiently by the cellular enzymes. Poor transportation of modified nucleosides into the cell and/or inefficient transformation from nucleosides to monophosphates, diphosphates and triphosphates have been reported as major problems for the *in vivo* applications.[243] In addition to this, the failure of the TCO nucleosides (Figure 5.8) might be caused by the conversion from *trans*-cyclooctene to the less reactive *cis*-cyclooctene form, mediated by the thiols in the cytoplasm.[244] The TCO group has been reported to be stable in PBS solution at 37 °C for one week, while in commercial mouse serum, the half-life is only 0.83 h.[245]

5.4 Synthesis and study of protected nucleoside monophosphates

In order to be effective for imaging cellular nucleic acids after delivery into cells, modified nucleosides must be converted into their monophosphates, diphosphates and

triphosphates by cellular kinases.[147, 149] The monophosphate conversion step is well-studied and is the rate limiting step with low yield for some modified nucleosides. Circumventing this step by delivering nucleoside monophosphates into cells would improve their conversion to the active triphosphates. However, under physiological conditions nucleoside monophosphates are negatively charged, and have poor cell membrane permeability. They are also susceptible to dephosphorylation.[150] To improve cell permeability in the antiviral drug field, lipophilic protecting groups have been utilised to produce masked nucleoside monophosphates.[150, 151] Lipophilic monophosphates are delivered into the cell more efficiently than their unprotected charged counterparts. McGuigan and co-workers used this strategy to improve the cell delivery and the uptake of masked phosphate nucleotides in studies on antiviral drugs.[246-248] After entering the cell, the protecting groups are rapidly cleaved by cellular enzymes to afford free nucleoside monophosphates (Figure 5.10).[151] These are converted to diphosphates and then to the active triphosphates. Using this strategy, high levels of antiviral nucleoside triphosphates were produced inside the cells.[249, 250] This approach was therefore applied to our cell imaging investigations for some of the modified nucleosides discussed in previous section.

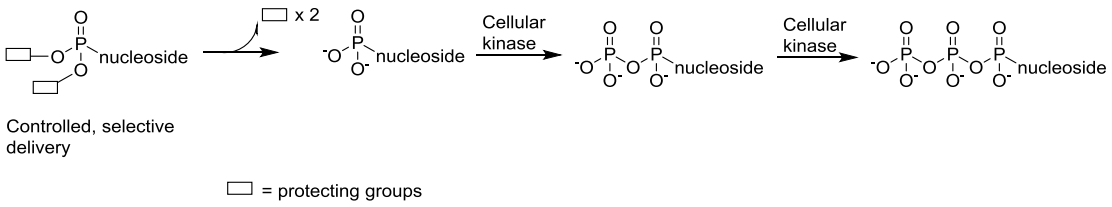


Figure 5.10. Transformation of masked nucleoside monophosphates into active triphosphates in cells.[151]

5.4.1 Synthesis of protected nucleoside monophosphates

Large numbers of phosphate protecting groups have been designed and tested.[250-252] The analysis of structure-activity relationships (SAR, Figure 5.11.A) shows that minor changes to the protecting groups greatly affect their properties. Aryloxy phosphoramidates with an L-alanine ester modification on the phosphorus centre are highly active in terms of deprotection in cells, producing high concentrations of corresponding triphosphates.[249-251] The proposed deprotection mechanism starts with the enzyme-mediated hydrolysis of the ester group from compound **I** to intermediate **II** (Figure 5.11.B).[150] A subsequent spontaneous intracellular cyclisation takes place to afford a five-membered cyclic anhydride **III**. The efficiency of this step is related to the structure of the nucleoside.[253] Further hydrolysis of intermediate **III** (not enzyme-catalysed) affords the intermediate phosphoalaninate **IV**, which has two positions that can be attacked by a water molecule. The last step to produce the free monophosphate nucleoside **V** is proposed to be an enzymatic reaction.

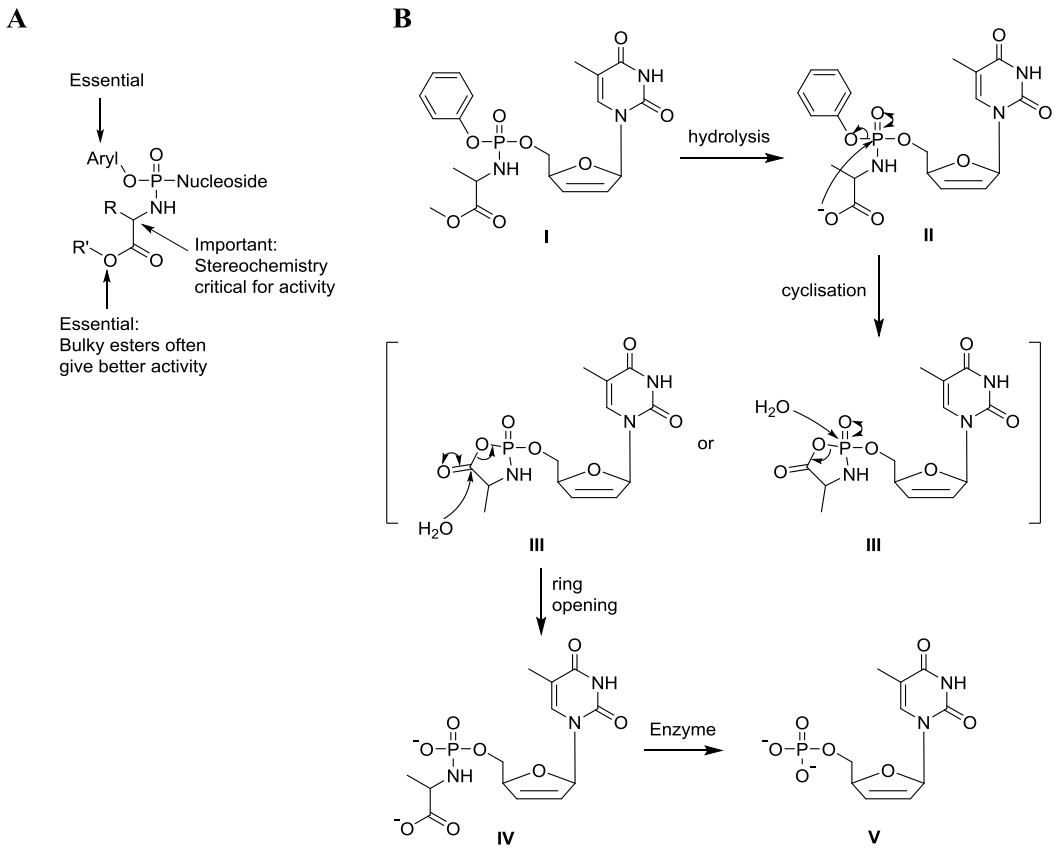
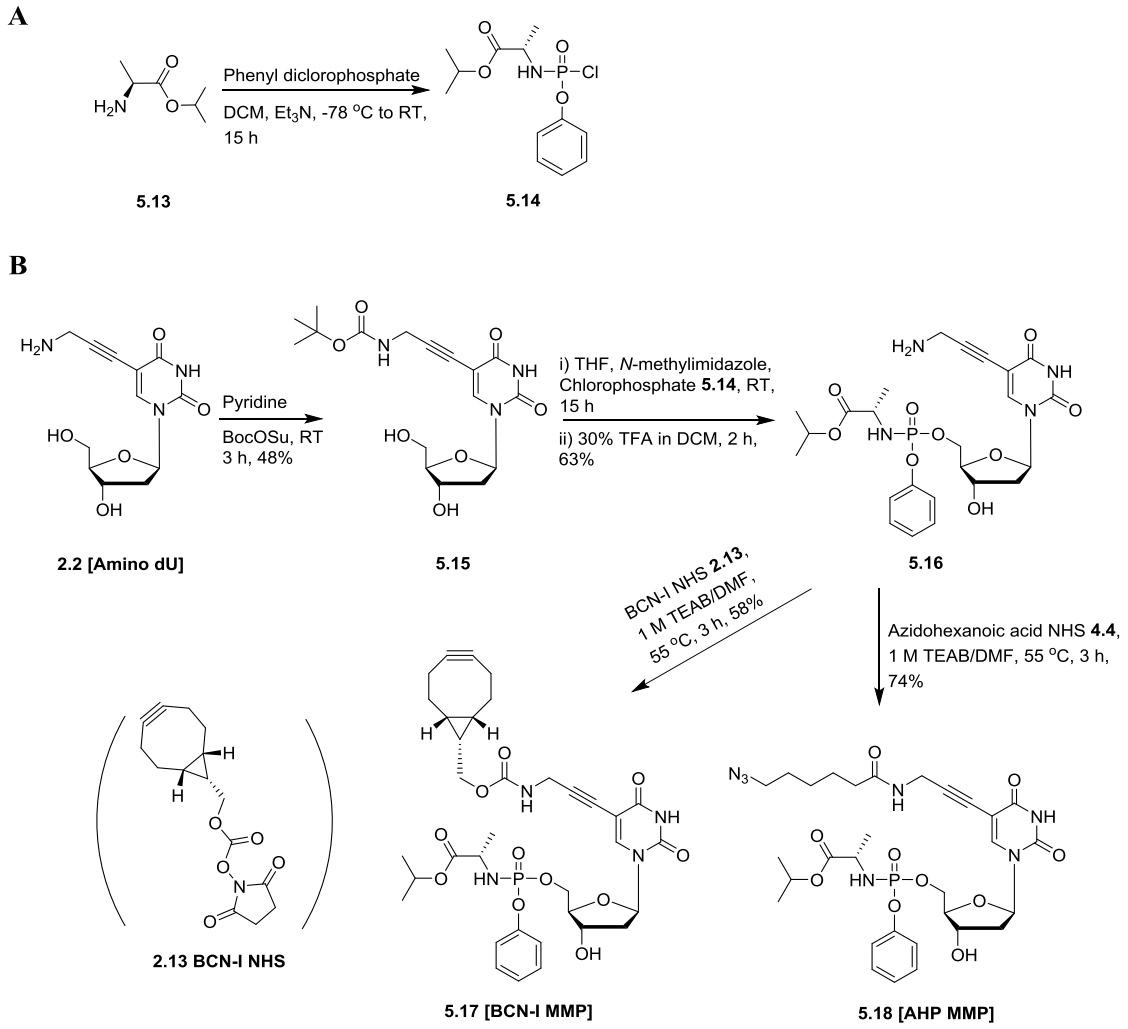


Figure 5.11. **A.** Structure-activity relationships for protecting group strategy in masked nucleoside monophosphates. **B.** Proposed deprotection mechanism inside cells.[150]

Taking the above studies into account we decided to evaluate AHP dU and BCN dU aryloxyphosphoramidate analogues with an L-alanine ester protecting group in cellular DNA imaging, since both of their corresponding triphosphates demonstrated efficiently *in vitro* enzymatic incorporation (Chapter 3 and Chapter 4). The synthesis of the protected monophosphates was based on literature methods.[249, 254] Isopropyl L-alanine ester **5.13** was reacted with phenyl dichlorophosphate to produce the protected chlorophosphate **5.14** (Scheme 5.3.A). The ^{31}P NMR spectra showed that two phosphate diastereoisomers were formed in a near 1:1 ratio. During the phosphorylation step, the reactive amino dU **2.2** had to be protected (Scheme 5.3.B). This type of masked phosphate has been reported to be stable to trifluoroacetic acid (TFA).[250] Therefore, the *t*-butyloxycarbonyl (Boc)

group, which can be deprotected under acidic conditions, was used to protect the amino group of amino dU **2.2**. Under the catalysis of *N*-methylimidazole, the nucleoside **5.15** was reacted with phenyl isopropyl L-alanine chlorophosphate **5.14** to form the masked monophosphate (MMP). The amino MMP **5.16** was obtained after acid deprotection of the Boc group. The BCN-I NHS carbonate **2.13** and azide NHS ester **4.4** were then coupled to the free amine to give compound BCN-I MMP **5.17** and AHP MMP **5.18**.



Scheme 5.3. Synthesis protocol for phenyl isopropyl L-alanine protected chlorophosphate (**A**) and masked nucleoside monophosphates (**B**).

5.4.2 Cell imaging using masked nucleoside monophosphates

The two masked monophosphates BCN-I MMP **5.17** and AHP MMP **5.18** were used in cellular imaging of DNA using the protocol previously described (Section 5.3). BCN-I MMP **5.17** was tested in collaboration with Dr Afaf El-Sagheer and did not give positive results. The DNA in cells incubated with AHP MMP **5.18** was initially labelled with Cy3-alkyne **5.12** using the CuAAC reaction to avoid the side reactions that are potentially associated with the SPAAC labelling. Cytoplasmic staining was observed (Figure 5.12) similar to previous results (Figure 5.7). Bright fluorescent loci were clearly observed in the cytoplasm (Figure 5.12, expanded region). This might indicate the locus of endosomes, where the protecting group of AHP MMP was removed, or where the further conversion of AHP monophosphate to diphosphate and triphosphate was occurring. However, the AHP nucleotide was not successfully incorporated into nuclear DNA, as determined by the lack of fluorescence in the nucleus. Further optimisation by varying cell incubation and staining conditions (incubation time and media), or by using other different cell lines, for example HeLa cells, might improve cell imaging.

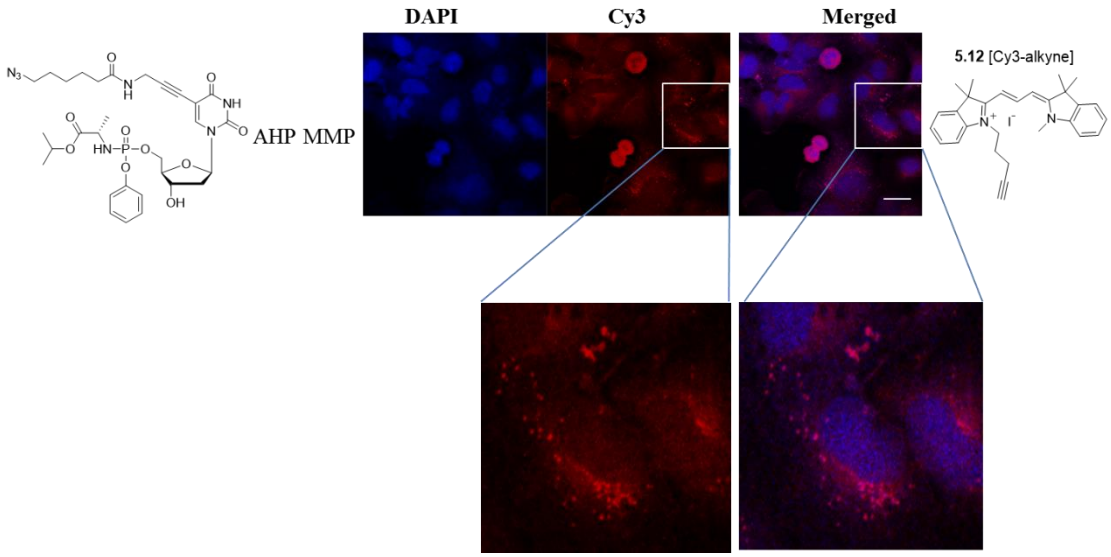


Figure 5.12. Cells treated with media containing AHP-modified masked nucleoside monophosphate (100 μM) and stained with Cy3-alkyne using the CuAAC reaction. 20 micron scale bar shown.

5.5 Conclusions

Several *trans*-cyclooctene and azide modified deoxyuridine nucleosides were prepared to visualise replicating DNA *in cellulo* using the CuAAC, SPAAC and IEDDA reactions. 5-Ethynyl dU (EdU) and 5-azidomethyl dU were used as controls. However, only EdU gave positive nuclear DNA fluorescent imaging. The SPAAC and CuAAC labelling of azido nucleosides showed only strong cytoplasmic fluorescent staining. The reason might be that the modified nucleotides were not produced from the corresponding nucleosides in cells or were not delivered into nucleus.

To improve the delivery and conversion of modified nucleosides to their triphosphates, phenyl isopropyl L-alanine protected azide (AHP) and BCN-I masked monophosphates were synthesised and studied in cellular DNA imaging. In this strategy the corresponding monophosphate can be directly produced after deprotection by cellular enzymes. Cells treated with AHP MMP exhibited accumulation of fluorescent signals in the cytoplasm (fluorescent loci). This may indicate the sites where AHP MMP is converted to free monophosphate or where it is converted to di- or tri-phosphate before delivery into the nucleus, and/or trapping in endosomes. The cellular nucleic acid imaging method described here using modified nucleosides and masked nucleoside monophosphates is a preliminary study. Further optimisation and more systematic studies will be carried out in the future, for example to investigate the cytotoxicity of these modified nucleosides and nucleotides, and to test them using different cell lines.

6 Conclusions and future work

New methods have been developed for the fluorescent labelling of oligonucleotides to expand their use in nucleic acid detection and diagnostics. Amino dUTP **2.11** was synthesised and labelled with a variety of NHS active esters and carbonates to give the corresponding modified triphosphates. In total, three cyclooctyne dUTPs (DIBO dUTP **2.5**, BCN-I dUTP **2.12** and BCN-II dUTP **2.13**), two *trans*-cyclooctene dUTPs (TCO-S dUTP **3.1** and TCO-L dUTP **3.2**) and one azide dUTP (AHP dUTP **4.5**) were prepared as well as a shorter linker AM dUTP **4.2** to compare the effect of linkage length on incorporation into DNA and subsequent labelling. The enzymatic incorporation of these novel triphosphates was examined in primer extension, PCR and reverse transcription reactions, which are used to generate functionalised DNA.

Cyclooctyne-modified dUTPs (DIBO, BCN-I and BCN-II dUTPs) were incorporated successfully by primer extension using Klenow, Gotaq, KOD and Therminator™ II DNA polymerases provided that the template did not contain several contiguous adenine bases. When incorporated against consecutive dA stretches in the template, the bulky cycloalkynes blocked replication and amplification. Fully extended products could only be obtained for BCN-I dUTP, which is smaller than DIBO and has a shorter linker than BCN-II. For PCR, a mixture of BCN-I dUTP and dTTP afforded BCN-functionalised PCR amplicons. These cyclooctyne modified primer extension and PCR products were efficiently labelled with Cy3-azide *via* the SPAAC reaction.

Compared to the above cyclooctyne dUTPs, *trans*-cyclooctene dUTPs demonstrated better enzymatic incorporation efficiency. In primer extension and reverse transcription assays, fully extended TCO dU modified products were obtained even using a challenging template containing four clustered adenines. TCO-S dUTP with a shorter linker was a

better substrate than TCO-L dUTP for all polymerases except Gotaq. Efficient PCR amplification with TCO dUTPs (*i.e.* without natural dTTP) was observed using a short template, however PCR amplicons of *ca.* 500 bp were only obtained when mixtures of dTTP and TCO dUTP were used. TCO-S dU maintained the stability of DNA duplexes better than TCO-L dU, and this effect was more pronounced with an increased number of modified sites in the amplicons. Efficient fluorescent labelling was achieved for TCO functionalised primer extension and PCR products *via* the IEDDA reaction. As anticipated, TCO-L dU products with longer linker were labelled more efficiently than TCO-S dU products due to relief of steric hindrance.

An alternative and complementary strategy to cyclooctyne-modified DNA labelling was also explored, whereby azide-modified DNA strands were prepared and coupled to BCN containing fluorophores or reporter groups. Azide (AM and AHP) dUTPs were more efficiently incorporated into DNA than cyclooctyne and cyclooctene dUTPs for primer extension, reverse transcription and PCR. Promisingly, 100% AHP dU-containing PCR products of *ca.* 500 bp can be obtained with 100% sequence fidelity. λ-exonuclease digestion and streptavidin magnetic separation were successfully used to prepare AHP-modified single-stranded probes from these PCR products. Three novel BCN modified fluorophores (Cy3-BCN, FAM-BCN and Cy5-BCN) were then synthesised. Fluorescent labelling of AHP dU primer extension products demonstrated that Cy3-BCN reacted with AHP dU products *via* the SPAAC reaction more efficiently than Cy5-BCN and FAM-BCN. Efficient fluorescent labelling of PCR products and single-stranded DNA was achieved. The densely labelled single-stranded probes exhibited intense fluorescent signals, which are suitable for use in genomic fluorescence *in situ* hybridisation (FISH) assays. To enlarge the number of DNA probes for detection of genomic DNA in a cellular context, a dual labelling system was investigated by

combining two of the following: FAM-BCN, Cy3-BCN, Rhodamine-B-BCN and Cy5-BCN dyes, in the labelling reactions. Future studies need to focus on the application of these probes in techniques such as FISH to detect gene sequences of interest for disease diagnostics or gene mapping.

For *in cellulo* applications, the anionic charge of nucleoside triphosphates prevents them from passing through the outer cellular membrane. Therefore five azide-modified deoxyuridines, two *trans*-cyclooctene-modified deoxyuridines and a fluorescent quadracyclic adenine were synthesised. However, only intense cytoplasm staining was observed without nuclear DNA labelling using the CuAAC, SPAAC or IEDDA reaction. The reason for this might be that the formation of monophosphate from the modified nucleoside by cellular kinases is slow and rate determining, resulting in very low cellular concentrations of modified triphosphates. To bypass this step, lipophilic phenyl isopropyl L-alanine protecting groups were used to protect nucleoside 5′-monophosphates. These 'masked' nucleosides are known to have better cell permeability than the charged nucleoside monophosphates. Aminopropargyl nucleoside masked monophosphates (MMP) were prepared and labelled with azide and cyclooctyne NHS esters to afford AHP and BCN-I MMPs. Whilst incubation of cells with AHP MMP still gave cytoplasmic staining, brighter foci were also observed, which are unexplained. Further optimisation of incubation and staining conditions, and/or the use of different cell lines will be required to improve cellular DNA labelling; the eventual aim is to expand the toolkit for monitoring nascent cellular DNA synthesis *via* non-toxic labelling strategy, which could be used for live cell DNA an d RNA imaging and tracking.

7 Experimental

7.1 General information

All reagents were purchased from Sigma-Aldrich, Alfa Aesar, Acros Organics, Jena Bioscience, SynAffix, Invitrogen and Fisher Scientific, and used without purification with the exception of dichloromethane (DCM), tetrahydrofuran (THF), triethylamine (TEA) and *N*, *N*-diisopropylethylamine (DIPEA), which were purified by distillation or dried over activated molecular sieves. DIBO PNP carbonate, 5-azidomethyl-2′-deoxyuridine, 6-azidohexanoic acid, Cy3 NHS ester, Cy5 NHS ester and 5′,3′-di-*O*-*p*-toluoyl-quadracyclic adenine analogue-2′-deoxyriboside were previously synthesised in the Brown group. All the molecular sieves used in the experiments were 3 Å (beads, 4 – 8 mesh) from Sigma-Aldrich. Thin layer chromatography (TLC) was performed using Merck Kieselgel 60 F24 silica gel plates (0.22 mm thickness, aluminium backed) and the compounds were visualised by irradiation at 254 nm or by staining with anisaldehyde unless otherwise stated. Column chromatography was carried out under pressure using Merck Kieselgel Si 60 (40 – 63 micron) silica.

NMR spectra were recorded on a Bruker AV300, AV400, AVIII400 or AVII500 spectrometer. Chemical shifts are given in ppm, and *J* values are quoted in Hz. ^{1}H NMR spectra were internally referenced to the appropriate undeuterated solvent signal; ^{13}C NMR spectra were internally referenced to the appropriate deuterated solvent signal; assignment of the compounds was aided by COSY (^{1}H-^{1}H), HSQC-DEPT and HMBC (^{1}H-^{13}C) experiments. Asterisk mark is used in NMR characterisation to differentiate diastereoisomers.

Low-resolution mass spectra (LRMS) were recorded using electrospray ionisation (ESI) on a Fisons VG platform mass spectrometer or a Micromass platform 1 spectrometer in HPLC grade acetonitrile or methanol. High-resolution mass spectra (HRMS) of chemical compouds were recorded using a Bruker 9.4T Fourier transform ion cyclotron resonance mass spectrometer. HR-MS of deoxynucleoside triphosphates were recorded in deionised water on a Bruker microTOF mass spectrometer (ES^- mode).

RP-HPLC columns:

Column S: Phenomenex Luna 10μ C8 100Å, 10 × 250 mm

Column L: Phenomenex Gemini-NX 10μ C18 110A AXIA, 250 × 21.2 mm.

RP-HPLC buffer systems:

TEAB buffer system I: eluent A, 0.1 M TEAB buffer (pH 7.5); eluent B, 40% acetonitrile in 0.1 M TEAB buffer (pH 7.5)

TEAB buffer system II: eluent A, 0.1 M TEAB buffer (pH 7.5); eluent B, 50% acetonitrile in 0.1 M TEAB buffer (pH 7.5)

TEAB buffer system III: eluent A, 0.1 M TEAB buffer (pH 7.5); eluent B, 60% acetonitrile in 0.1 M TEAB buffer (pH 7.5)

Ammonium acetate system: eluent A, 0.1 M ammonium acetate (pH 7.0); eluent B, 50% acetonitrile in 0.1 M ammonium acetate (pH 7.0)

Quantification of triphosphates and oligonucleotides:

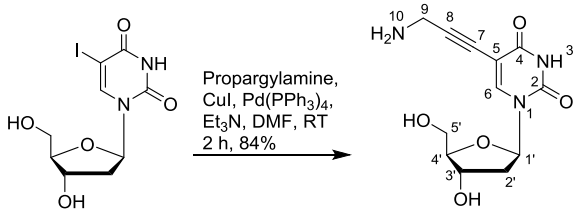
The concentration c of nucleoside triphosphates and oligonucleotides was calculated by Lambert-Beer equation A (aborbance) = OD (optical density) = εcl (extinction coefficient ε, cuvette length l = 1 cm) at RT in deionised water.

Ethanol precipitation:

The nucleotide solution (1 × volume) was mixed with 3 M NaCl for triphosphates or 3 M sodium acetate (pH 5.3) for oligonucleotides (0.1 × volume) and ethanol (3 × volume). Solution was left on dry ice for 20 min and at -20 ºC for 10 min. Samples were centrifuged with a speed of 15,000 × g at 4 ºC for 30 min.

7.2 Chemical synthesis

5-(3-Aminoprop-1-yn-1-yl)-2′-deoxyuridine (2.2, Amino dU)[255]



2.1 2.2

5-Iodo-2′-deoxyuridine **2.1** (5.00 g, 14.1 mmol, 1.0 eq) was co-evaporated with anhydrous dimethylformamide (DMF, 3 × 20 mL), then dissolved in anhydrous DMF (60 mL). CuI (0.54 g, 2.8 mmol, 0.2 eq), propargylamine (2.8 mL, 44 mmol, 3.1 eq) and TEA (10 mL, 72 mmol, 5.1 eq) were added under argon. The reaction was stirred at RT for 10 min before $Pd(PPh_3)_4$ (1.65 g, 1.43 mmol, 0.1 eq) was added and the reaction continued to stir for 2 h. The solvent was removed *in vacuo* and the residue was purified by column chromatography on silica (0% to 20% MeOH in DCM). Compound **2.2** (3.32 g,

11.8 mmol) was obtained in 84% yield as a yellow solid. M_w = 281.10 g/mol (Chemical Formula: $C_{12}H_{15}N_3O_5$)

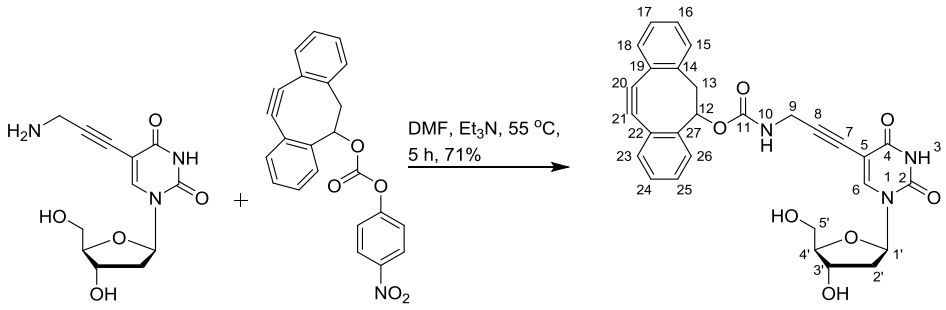
R_f: 0.25 [DCM/MeOH 1:1 *v/v*]

LRMS: [ESI$^-$, MeOH] *m/z*: 280.2 ([M-H]$^-$, 100)

^{1}H NMR (300 MHz, DMSO-d_6) δ 8.13 (s, 1 H, H-6), 6.12 (app t, *J* = 6.7 Hz, 1 H, H-1′), 5.03 (br. s, 4 H, H-10, 5′-OH, 3′-OH), 4.23 (app td, *J* = 4.5, 3.4 Hz, 1 H, H-3′), 3.79 (app q, *J* = 3.4 Hz, 1 H, H-4′), 3.61 (dd, *J* = 11.6, 3.4, 1 H, H-5′a), 3.56 (dd, *J* = 11.6, 3.4 Hz, 1 H, H-5′b), 3.46 (s, 2 H, H-9), 2.11 (dd, *J* = 6.7, 4.5 Hz, 2 H, H-2′)

^{13}C NMR (75 MHz, DMSO-d_6) δ 161.7 (C4), 149.4 (C2), 143.0 (C6), 109.5 (C5), 98.7 (C8), 87.5 (C4′), 84.6 (C1′), 73.9 (C7), 70.2 (C3′), 61.0 (C5′), 40.0 (C2′), 31.4 (C9)

5-(3-((1,2:5,6-dibenzocyclooct-7-yn-3-yl)oxycarbonylamino)prop-1-yn-1-yl)-2′-deoxyuridine (2.4, DIBO dU)



2.2 **2.3** **2.4**

Amino dU **2.2** (0.35 g, 1.2 mmol, 1.2 eq) was co-evaporated with anhydrous DMF (3 × 2 mL), then dissolved in anhydrous DMF (10 mL). TEA (1.0 mL, 7.2 mmol, 7.5 eq) and DIBO *p*-nitrophenyl (PNP) carbonate[128] **2.3** (0.37 g, 0.96 mmol, 1.0 eq) were added under argon. The reaction was stirred at 55 °C for 5 h. The solvent was removed *in vacuo* and the residue was purified by Biotage® SNAP KP-Sil-50 cartridge on Biotage Isolera Four

equipment (0% to 7% MeOH in DCM). Compound **2.4** (0.36 g, 0.68 mmol) was obtained in 71% yield as a yellow foam. M_w = 527.17 g/mol (Chemical Formula: $C_{29}H_{25}N_3O_7$)

R_f: 0.21 [DCM/MeOH 9:1 *v/v*]

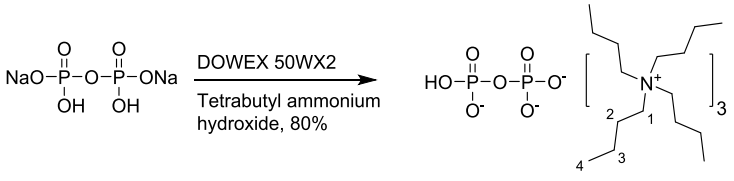
LRMS: [ESI$^+$, MeCN] *m/z*: 550.0 ([M+Na]$^+$, 100).

HRMS: [ESI$^+$, MeCN] Calculated mass [M+H]$^+$: 550.1585; found *m/z*: 550.1585.

^{1}H NMR (400 MHz, DMSO-d_6) δ 11.62 (s, 1 H, H-3), 8.18 – 8.12 (m, 2 H, H-6, H-10), 7.56 (d, *J* = 8.1 Hz, 1 H, ArH), 7.51 – 7.32 (m, 7 H, ArH), 6.11 (app t, *J* = 6.6 Hz, 1 H, H-1′), 5.40 – 5.29 (m, 1 H, H-12), 5.23 (d, *J* = 4.0 Hz, 1 H, OH-3′), 5.06 (t, *J* = 5.1 Hz, 1 H, OH-5′), 4.25 – 4.19 (m, 1 H, H-3′), 4.04 (d, *J* = 5.6 Hz, 2 H, H-9), 3.79 (app. q, *J* = 3.0 Hz, 1 H, H-4′), 3.63 – 3.50 (m, 2 H, H-5′), 3.23 – 3.15 (m, 1 H, H-13a), 2.78 (dd, *J* = 15.2, 3.5 Hz, 1 H, H-13b), 2.07 – 2.18 (m, 2 H, H-2′)

^{13}C NMR (100 MHz, DMSO-d_6) δ 161.6 (C4), 155.0 (C11), 152.1 (ArC), 150.8 (ArC), 149.4 (C2), 143.7 (C6), 130.3 (ArC), 128.4 (2 × ArC), 127.4 (ArC), 127.3 (ArC), 126.1 (ArC), 125.8 (ArC), 123.9 (ArC), 122.9 (ArC), 120.3 (ArC), 112.6 (C20), 109.8 (C5), 98.0 (C8), 89.5 (C21), 87.6 (C4′), 84.7 (C1′), 75.8 (C12), 74.7 (C7), 70.2 (C3′), 61.0 (C5′), 45.5 (C13), 40.0 (C2′), 30.8 (C9)

Tris(tetrabutylammonium) hydrogen pyrophosphate (2.7, Pyrophosphate)[173]



2.6 **2.7**

Disodium pyrophosphate dibasic **2.6** (1.11 g, 5.00 mmol, 1.0 eq) was dissolved in deionised water (20 mL) and passed through Dowex 50WX2 resin (acid form). Deionised water (70 mL) was used to wash the resin until the pH reached 7. The eluent was collected in a 250 mL flask containing 0.2 M tetrabutylammonium hydroxide (25 mL). This was

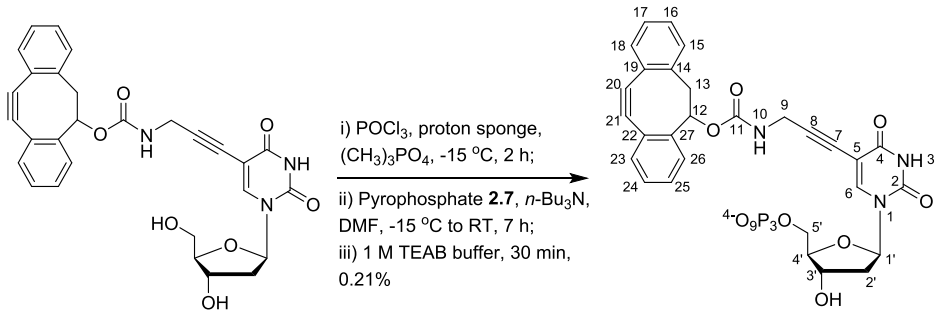
followed by adding 1.0 M tetrabutylammonium hydroxide (5 mL) until the solution reached pH 7. The solvent was removed *in vacuo* and the resulting solid was left under vacuum for 2 days. Compound **2.7** (3.6 g, 4.0 mmol) was obtained in 80% yield as a white solid and used without further purification. The compound is hygroscopic and was aliquoted inside a glovebox. M_w = 902.36 g/mol (Chemical Formula: $C_{48}H_{109}N_3O_7P_2$)

^{1}H NMR (400 MHz, D_2O) δ 3.25 – 3.11 (m, 8 H, H-1), 1.70 – 1.56 (m, 8 H, H-2), 1.34 (app sxt, *J* = 7.3 Hz, 8 H, H-3), 0.93 (t, *J* = 7.3 Hz, 12 H, H-4)

^{13}C NMR (101 MHz, D_2O) δ 58.1 (C1), 23.1 (C2), 19.1 (C3), 12.8 (C4)

^{31}P NMR (162 MHz, D_2O) δ -8.6

5-(3-((1,2:5,6-dibenzocyclooct-7-yn-3-yl)oxycarbonylamino)prop-1-yn-1-yl)-2′-deoxyuridine-5′-*O*-triphosphate (2.5, DIBO dUTP)

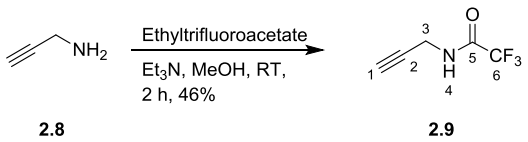


DIBO dU **2.4** (80 mg, 0.15 mmol, 1.0 eq) was co-evaporated with anhydrous pyridine (3 × 2 mL) and dried over KOH in a desiccator individually with proton sponge (49 mg, 0.23 mmol, 1.5 eq), activated molecular sieves and aliquoted pyrophosphate **2.7** (0.80 g, 0.89 mmol, 5.9 eq). Trimethylphosphate and tributylamine were dried over molecular sieves for 15 h. Trimethylphosphate (1.2 mL) and proton sponge were added to DIBO dU **2.4** in a flask under argon at -15 °C. Phosphorus oxychloride (28 µL, 0.30 mmol, 2.0 eq)

was added dropwise and the reaction was stirred at -15 °C for 2 h. A mixture of pyrophosphate **2.7** in DMF (440 µL) and tributylamine (360 µL) was added into the reaction mixture. The reaction was slowly warmed up to RT and stirred for 7 h. 1 M TEAB buffer (5 mL, pH 7.5) was added to quench the reaction. The solvent was removed *in vacuo*, then the residue was dissolved in deionised water (20 mL) and washed with DCM (3 × 20 mL). The water layer was dried *in vacuo* and the residue was re-dissolved in deionised water (1 mL). Ethanol precipitation was carried out by adding 3 M NaCl (100 µL) and ethanol (3 mL). The precipitate was purified by RP-HPLC (Column S with TEAB buffer system III, 3.5% to 60% buffer B in 20 min and monitored by UV absorption at 305 nm). Compound **2.5** (8.3 OD_{290}, 0.32 µmol, ε_{290} = 26000 $M^{-1}cm^{-1}$, calculated from DIBO dU measurement) was obtained in 0.21% yield as a TEA salt. M_w = 767.47 g/mol (Chemical Formula: $C_{29}H_{28}N_3O_{16}P_3$)

HRMS: [ESI^-, H_2O] calculated mass: 766.0604 $[M-H]^-$; found *m/z*: 766.0792 $[M-H]^-$

2,2,2-Trifluoro-*N*-(prop-2-yn-1-yl)acetamide (2.9, TFA-aminopropyne)[256]



2.8 **2.9**

Propargylamine **2.8** (0.60 mL, 9.4 mmol, 1.0 eq) was added to anhydrous MeOH (7 mL) with molecular sieves, then cooled to 0 ºC and stirred for 10 min. TEA (2.0 mL, 14 mmol, 1.5 eq) and ethyltrifluoroacetate (1.6 mL, 14 mmol, 1.4 eq) were added dropwise at 0 ºC. The reaction mixture was then warmed up to RT and stirred for 2 h. Solvent was removed *in vacuo* and the residue was purified by column chromatography on silica (DCM).

Compound **2.9** (0.66 g, 4.3 mmol) was obtained in 46% yield as brown oil. M_w = 152.2 g/mol (Chemical Formula: $C_5H_4F_3NO$)

R_f: 0.81 [DCM/MeOH 95:5 *v/v*], stain: ninhydrin and phosphomolybdic acid (the product is volatile, no strong heat).

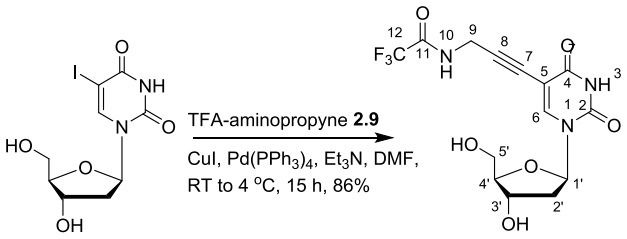
LRMS: [ESI^+, MeCN] *m/z*: 152.2 ($[M+H]^+$, 13); 173.2 ($[M+Na]^+$, 25); 203.2 ($[M+2Na]^+$, 100). [ESI^-, MeCN] 150.1 ($[M-H]^-$, 100).

^{1}H NMR (300 MHz, $CDCl_3$) δ 6.55 (br. s, 1 H, H-4), 4.17 (dd, *J* = 5.5, 2.6 Hz, 2 H, H-3), 2.35 (t, *J* = 2.6 Hz, 1 H, H-1)

^{13}C NMR (100 MHz, $CDCl_3$) δ 156.8 (q, *J* = 38.0 Hz, C5), 115.5 (q, *J* = 288.0 Hz, C6), 78.6 (C2), 73.4 (C1), 29.7 (C3)

^{19}F NMR (282 MHz, $CDCl_3$) δ -76.1 (s)

5-(3-(2,2,2-Trifluoroacetamido)prop-1-yn-1-yl)-2′-deoxyuridine (2.10, TFA-amino dU)[59]



5-Iodo-2′-deoxyuridine **2.1** (0.50 g, 1.4 mmol, 1.0 eq) was co-evaporated with anhydrous DMF (3 × 2 mL) and left under vacuum for 2 h. The compound was dissolved in DMF (7 mL) with molecular sieves and stirred at RT for 10 min. CuI (54 mg, 0.28 mol, 0.2 eq), TEA (2.0 mL, 14 mmol, 10.0 eq), TFA-amidopropyne **2.9** (0.43 g, 2.8 mmol, 2.0 eq) and $Pd(PPh_3)_4$ (0.16 g, 0.14 mmol, 0.1 eq) were added under argon, then stirred at RT for 3 h.

Another portion of TFA-amidopropyne **2.9** (0.20 g, 1.3 mmol, 0.9 eq) was added and the reaction was stirred at 4 ºC for 15 h. The solvent was removed *in vacuo* and the residue was purified by column chromatography on silica (EtOAc with 0.5% TEA). Compound **2.10** (0.44 g, 1.2 mmol) was obtained in 86% yield as a yellow foam. M_w = 377.08 g/mol (Chemical Formula: $C_{14}H_{14}F_3N_3O_6$)

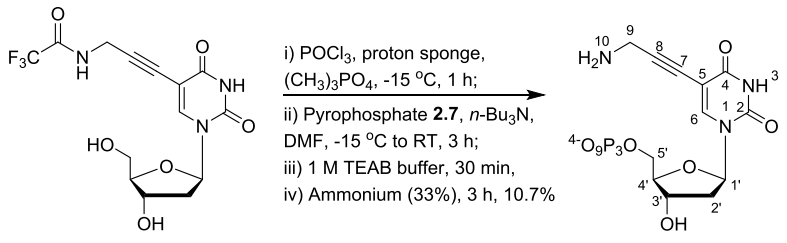
R_f: 0.45 [DCM/MeOH 8:2 *v/v*]

LRMS: [ESI$^+$, MeOH] *m/z*: 400.1 ([M+H]$^+$, 100); 777.2 ([2M+Na]$^+$, 74).

^{1}H NMR (300 MHz, DMSO-d_6) δ 11.62 (s, 1 H, H-3), 10.05 (t, *J* = 5.1 Hz, 1 H, H-10), 8.19 (s, 1 H, H-6), 6.11 (app t, *J* = 6.5 Hz, 1 H, H-1′), 5.23 (d, *J* = 4.2 Hz, 1 H, OH-3′), 5.07 (t, *J* = 5.0 Hz, 1 H, OH-5′), 4.28 – 4.19 (m, 3 H, H-3′, H-9), 3.80 (app q, *J* = 3.3 Hz, 1 H, H-4′), 3.66 – 3.51 (m, 2 H, H-5′), 2.12 (dd, *J* = 6.5, 4.9 Hz, 2 H, H-2′)

^{13}C NMR (75 MHz, DMSO-d_6) δ 161.5 (C4), 149.4 (C2), 144.1 (C6), 97.6 (C8), 87.6 (C4′), 84.8 (C1′), 75.4 (C7), 70.2 (C3′), 61.0 (C5′), 40.3 (C2′), 29.4 (C9)

5-(3-Aminoprop-1-yn-1-yl)-2′-deoxyuridine-5′-*O*-triphosphate (2.11, Amino dUTP)[59]



2.10 **2.11**

TFA-amino dU **2.10** (150 mg, 0.398 mmol, 1.0 eq) was dried over KOH in the same desiccator individually with proton sponge (128 mg, 0.597 mmol, 1.5 eq), activated molecular sieves and aliquoted pyrophosphate **2.7** (2.0 g, 2.2 mmol, 5.5 eq). Trimethylphosphate and tributylamine was dried over molecular sieves for 15 h.

Trimethylphosphate (2.8 mL) and proton sponge were added to the TFA-amino dU **2.10** in a flask under argon at -15 °C. Phosphorus oxychloride (45 µL, 0.48 mmol, 1.2 eq) was added dropwise and stirred at -15 °C for 1 h. A mixture of pyrophosphate **2.7** in DMF (2.2 mL) and tributylamine (400 µL) was added. The reaction mixture was slowly warmed up to RT and stirred for 3 h. 1 M TEAB buffer (7.5 mL, pH 7.5) was added to quench the reaction and the solvent was removed *in vacuo*. A reverse (ISOLUTE) FLASH C18 cartridge eluted with water and acetonitrile was used to purify the triphosphate. RP-HPLC using column L and TEAB buffer system III (3.5% to 60% eluent B in 50 min, monitored by UV absorption at 305 nm) was used to further purify the triphosphate. The intermediate was left in 33% ammonia for 3 h to deprotect the TFA protecting group. The solvent was removed *in vacuo* and the residue was purified by RP-HPLC purification (column L and TEAB buffer system I, 3.5% to 50% eluent B in 50 min). Compound **2.11** (453 OD_{290}, 42.7 µmol, ε_{290} = 10600 $M^{-1}cm^{-1}$) was obtained in 10.7% yield as a TEA salt. M_w = 521.20 g/mol (Chemical Formula: $C_{12}H_{18}N_3O_{14}P_3$)

R_f: 0.09 [*n*-propanol/ammonium hydroxide/H_2O 3:1:1 *v/v/v*]

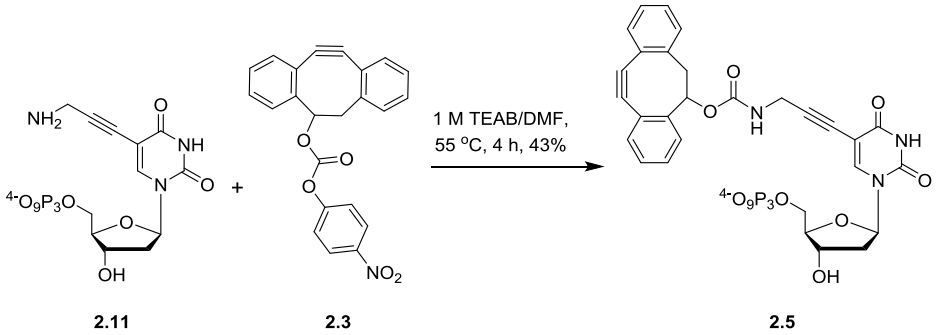
HRMS: [ESI^-, H_2O] Calculated mass: 519.9923 $[M-H]^-$, found *m/z*: 519.9780 $[M-H]^-$.

^{1}H NMR (400 MHz, D_2O) δ 8.22 (s, 1 H, H-6), 6.25 (app t, *J* = 6.1 Hz, 1 H, H-1′), 4.60 (app dt, *J* = 6.1, 5.0 Hz, 1 H, H-3′), 4.24 – 4.16 (m, 2 H, H-5′), 4.14 – 4.07 (m, 1 H, H-4′), 3.80 (s, 2 H, H-9), 3.11 (q, *J* = 7.3 Hz, $Et_2NH^+C\boldsymbol{H}_2CH_3$),2.36 (ddd, *J* = 13.9, 6.1, 5.0 Hz, 1 H, H-2′a), 2.27 (dt, *J* = 13.9, 6.1 Hz, 1 H, H-2′a) 1.21 (t, *J* = 7.3 Hz, $Et_2NH^+CH_2C\boldsymbol{H}_3$)

^{13}C NMR (100 MHz, D_2O) δ 145.1 (C6), 98.6 (C5), 85.5 (C4′), 85.3 (C1′), 69.6 (C3′), 64.7 (C5′), 46.5 ($Et_2NH^+\mathbf{C}H_2CH_3$), 39.5 (C2′), 30.0 (C9), 8.43 ($Et_2NH^+CH_2\mathbf{C}H_3$)

^{31}P NMR (121 MHz, D_2O) δ -5.5 (d, *J* = 20 Hz, P_γ), -10.5 (d, *J* = 20 Hz, P_α), -21.4 (t, *J* = 20 Hz, P_β)

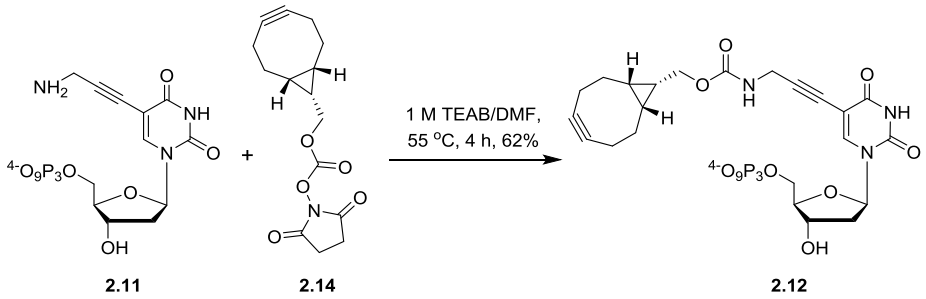
5-(3-((1,2:5,6-dibenzocyclooct-7-yn-3-yl)oxycarbonylamino)prop-1-yn-1-yl)-2′-deoxyuridine-5′-*O*-triphosphate (2.5, DIBO dUTP)



The freeze-dried amino dUTP **2.11** (16 OD_{290}, 1.5 µmol, 1.0 eq) was dissolved in deionised water (80 µL) and 1 M TEAB buffer (80 µL, pH 7.5). The DIBO PNP carbonate **2.3** (2 mg, 5.2 µmol, 3.5 eq) in DMF (160 µL) was added to the triphosphate solution. The mixture was kept at 55 ºC for 4 h and the solvent was removed *in vacuo.* The residue was purified by RP-HPLC (column S with TEAB buffer system III, 3.5% to 60% eluent B in 20 min and monitored by UV absorption at 305 nm). Compound **2.5** (17 OD_{290}, 0.65 µmol) was obtained in 43% yield as a TEA salt. M_w = 767.47 g/mol (Chemical Formula: $C_{29}H_{28}N_3O_{16}P_3$)

HRMS: [ESI^-, H_2O] calculated mass: 766.0604 $[M-H]^-$; found *m/z*: 766.0792 ($[M-H]^-$, 100)

5-(3-(((1R,8S,9s)-bicyclo[6.1.0]non-4-yn-9-yl)methyloxycarbonylamino)prop-1-yn-1-yl)-2′-deoxyuridine-5′-*O*-triphosphate (2.12, BCN-I dUTP)

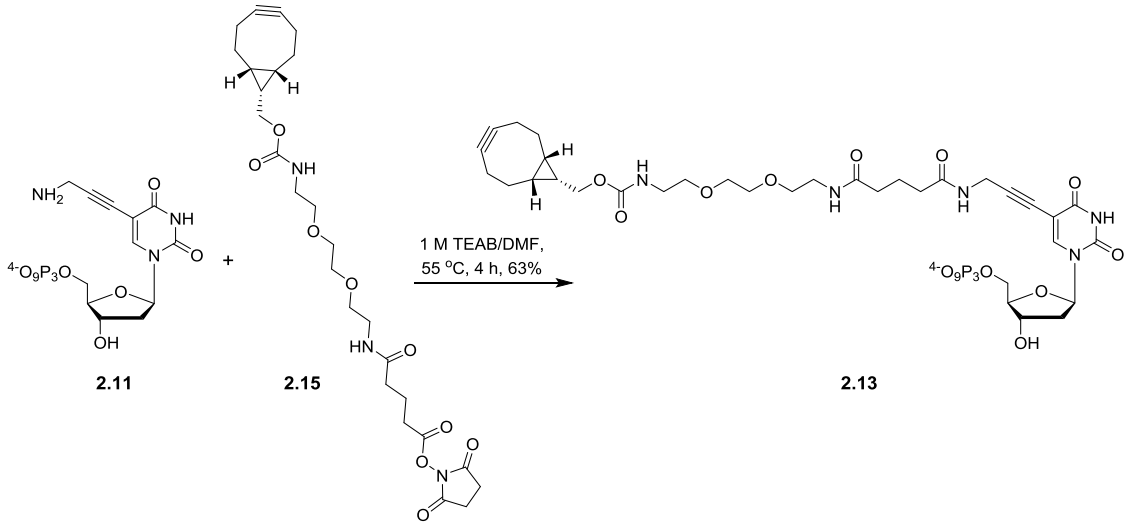


2.11 2.14 2.12

The freeze-dried amino dUTP **2.11** (11 OD_{290}, 1.0 µmol, 1.0 eq) was dissolved in deionised water (200 µL) and 1 M TEAB buffer (200 µL, pH 7.5). The BCN-I carbonate **2.14** (1 mg, 3.4 µmol, 3.7 eq) in DMF (200 µL) was added to the triphosphate solution. The mixture was kept at 55 ºC for 4 h and the solvent was then removed *in vacuo.* The residue was purified by RP-HPLC (column S with TEAB buffer system III, 3.5% to 60% eluent B in 20 min and monitored by UV absorption at 305 nm). Compound **2.12** (6.6 OD_{290}, 0.62 µmol) was obtained in 62% yield as a TEA salt. M_w = 697.42 g/mol (Chemical Formula: $C_{23}H_{30}N_3O_{16}P_3$)

HRMS: [ESI^-, H_2O] calculated mass: 696.0761 $[M-H]^-$; found *m/z*: 696.0742 ($[M-H]^-$, 100)

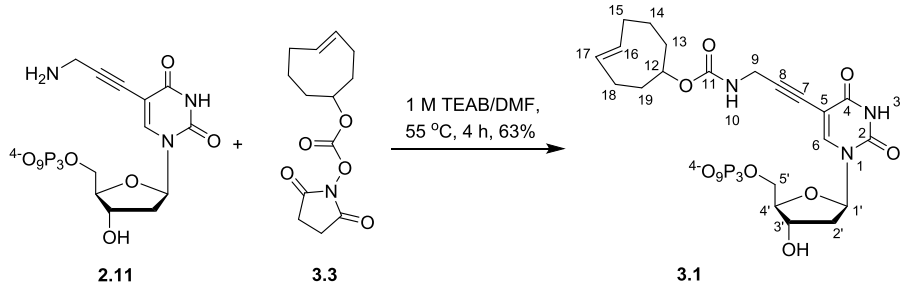
5-(22-((1R,8S,9s)-bicyclo[6.1.0]non-4-yn-9-yl)-5,9,20-trioxo-13,16,21-trioxa-4,10,19-triazadocos-1-yn-1-yl)-2′-deoxyuridine-5′-*O*-triphosphate (2.13, BCN-II dUTP)



The freeze-dried amino dUTP **2.11** (16 OD_{290}, 1.5 µmol, 1.0 eq) was dissolved in deionised water (80 µL) and 1 M TEAB buffer (80 µL, pH 7.5). The BCN-II NHS ester **2.15** (2 mg, 3.7 µmol, 2.5 eq) in DMF (160 µL) was added to the triphosphate solution. The mixture was kept at 55 ºC for 4 h and the solvent was then removed *in vacuo.* The residue was purified by RP-HPLC (column S with TEAB buffer system III, 3.5% to 60% eluent B in 20 min and monitored by UV absorption at 305 nm). Compound **2.13** (10 OD_{290}, 0.94 µmol) was obtained in 63% yield as a TEA salt. M_w = 941.7103 g/mol (Chemical Formula: $C_{34}H_{50}N_5O_{20}P_3$)

HRMS: [ESI^-, H_2O] calculated mass: 940.2184 $[M-H]^-$; found *m/z*: 940.2357 ($[M-H]^-$, 100)

5-(3-(((E)-cyclooct-4-en-1-yl)oxycarbonylamino)prop-1-yn-1-yl)-2′-deoxyuridine-5′-*O*-triphosphate (2.12, TCO-S dUTP)[139]



The freeze-dried amino dUTP **2.11** (20 OD_{290}, 1.9 µmol, 1.0 eq) was dissolved in 1 M TEAB buffer (120 µL, pH 7.5). TCO-S NHS carbonate **3.3** (1.5 mg, 5.6 µmol, 2.9 eq) in DMF (100 µL) was added to the triphosphate solution. The mixture was kept at 55 °C for 4 h and the solvent was removed *in vacuo*. The residue was purified by RP-HPLC (column S with TEAB buffer system III, 3.5% to 60% eluent B in 20 min and monitored by UV absorption at 305 nm). Compound **3.1** (13 OD_{290}, 1.2 µmol) was obtained in 63% yield as a TEA salt. M_w = 673.4 g/mol (Chemical Formula: $C_{21}H_{30}N_3O_{16}P_3$)

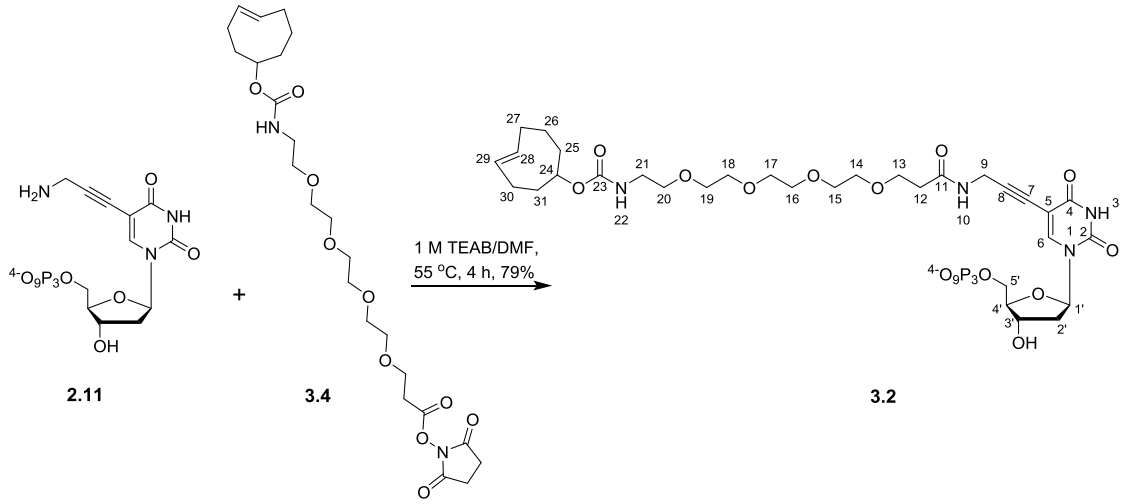
HRMS: [ESI^-, H_2O] Calculated mass: 672.0761 $[M-H]^-$; Found: 672.0801 ($[M-H]^-$, 100)

^{1}H NMR (500 MHz, D_2O) δ 8.07 (s, 1 H, H-6), 6.23 (app t, *J* = 6.5 Hz, 1 H, H-1′), 5.66 (ddd, *J* = 16.2, 9.1, 4.7 Hz, 1 H, H-17), 5.52 (ddd, *J* = 16.2, 10.9, 2.8 Hz, 1 H, H-16), 4.63 – 4.59 (m, 1 H, H-3′), 4.33 – 4.25 (m, 1 H, H-12), 4.24 – 4.12 (m, 3 H, H-4′, H-5′), 4.05 (s, 2 H, H-9), 3.15 (q, *J* = 7.3 Hz, $Et_2NH^+C\boldsymbol{H}_2CH_3$), 2.41 – 2.23 (m, 5 H, H-2′, H-18, H-13a), 2.04 – 1.82 (m, 4 H, H-15, H-13b, H-14a), 1.73 – 1.51 (m, 3 H, H-19, H-14b), 1.23 (t, *J* = 7.3 Hz, $Et_2NH^+CH_2C\boldsymbol{H}_3$)

^{13}C NMR (126 MHz, D_2O) δ 144.6 (C6), 135.7 (C17), 133.4 (C16), 99.1 (C5), 85.53 and 85.49 (C4′, C1′), 82.4 (C12), 70.0 (C3′), 65.0 (C5′), 46.6 ($Et_3NH^+\boldsymbol{C}H_2CH_3$), 40.3 (C15), 38.6 (C2′), 37.8 (C19), 33.7 (C13), 32.0 (C18), 30.7 (C9), 30.5 (C14), 8.2 ($Et_2NH^+CH_2\boldsymbol{C}H_3$)

^{31}P NMR (202 MHz, D_2O) δ -5.1 – -5.4 (m, P_γ), -9.9 – -10.1 (m, P_α), -20.9 – -21.3 (m, P_β)

5-(21-((E)-cyclooct-4-en-1-yloxy)-5,21-dioxo-8,11,14,17-tetraoxa-4,20-diazahenicos-1-yn-1-yl)-2′-deoxyuridine-5′-*O*-triphosphate (3.2, TCO-L dUTP)



The freeze-dried amino dUTP **2.11** (20 OD_{290}, 1.9 µmol, 1.0 eq) was dissolved in 1 M TEAB buffer (120 µL, pH 7.5). TCO-L NHS ester **3.4** (2 mg, 3.9 µmol, 2.1 eq) in DMF (100 µL) was added to the triphosphate solution. The mixture was kept at 55 °C for 4 h and the solvent was removed *in vacuo*. The residue was purified by RP-HPLC (column S with TEAB buffer system III, 3.5% to 60% eluent B in 20 min and monitored by UV absorption at 305 nm). Compound **3.2** (16 OD_{290}, 1.5 µmol) was obtained in 79% yield as a TEA salt. M_w = 920.226 g/mol (Chemical Formula: $C_{32}H_{51}N_4O_{21}P_3$).

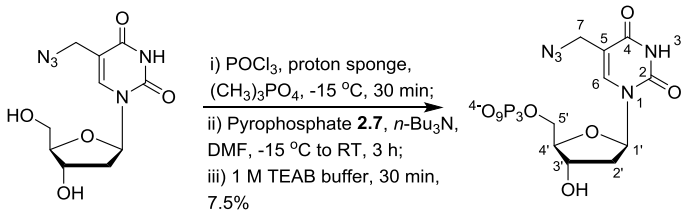
HRMS: [ESI^-, H_2O] Calculated mass: 919.2180 $[M-H]^-$; Found *m/z*: 919.2295 ($[M-H]^-$, 100).

^{1}H NMR (500 MHz , D_2O) δ 8.13 (s, 1 H, H-6), 6.22 (app t, *J* = 6.5 Hz, 1 H, H-1′), 5.69 – 5.57 (m, 1 H, H-29), 5.56 – 5.45 (m, 1 H, H-28), 4.65 – 4.58 (m, 1 H, H-3′), 4.27 – 4.12 (m, 6 H, H-24, H-5′, H-4′, H-9), 3.77 (t, *J* = 6.0 Hz, 2 H, H-13), 3.66 – 3.59 (m, 12 H, H-14 to H-19), 3.54 (t, *J* = 5.2 Hz, 2 H, H-20), 3.24 (t, *J* = 5.2 Hz, 2 H, H-21), 3.16 (q, *J* = 7.3 Hz, $Et_2NH^+C\boldsymbol{H}_2CH_3$), 2.54 (t, *J* = 6.0 Hz, 2 H, H-12), 2.40 – 2.32 (m, 2 H, H-2′), 2.32 – 2.22 (m, 3 H, H-30, H-25a), 2.00 – 1.83 (m, 4 H, H-27, H-25b, H-26a), 1.67 – 1.51 (m, 3 H, H-26b, H-31), 1.23 (t, *J* = 7.3 Hz, $Et_2NH^+CH_2C\boldsymbol{H}_3$)

^{13}C NMR (126 MHz , D_2O) δ 173.8 (C11), 158.3 (C4), 144.9 (C6), 135.6 (C29), 133.3 (C28), 99.1 (C5), 89.7 (C8), 85.65 and 85.59 (C4′, C1′), 81.8 (C24), 73.5 (C7), 70.1 (C3′), 69.59, 69.57, 69.53, 69.50 and 69.4 (C14, C15, C16, C17, C18, C19, C20), 66.6 (C13), 65.0 (C5′), 46.6 ($Et_3NH^+\boldsymbol{C}H_2CH_3$), 40.4 (C27), 39.9 (C21), 38.7 (C2′), 37.8 (C31), 35.9 (C12), 33.7 (C25), 32.0 (C30), 30.6 (C26), 29.6 (C9), 8.2 ($Et_2NH^+CH_2\boldsymbol{C}H_3$)

^{31}P NMR (162 MHz, D_2O) δ -8.0 (br. s, P_γ), -11.4 (d, J = 20 Hz, P_α), -22.6 (br. s, P_β)

5-Azidomethyl-2′-deoxyuridine-5′-*O*-triphosphate (4.2, AM dUTP)



4.1 **4.2**

5-Azidomethyl-2′-deoxyuridine[257] **4.1** (100 mg, 0.353 mmol, 1.0 eq) was dried over KOH in a desiccator individually with proton sponge (113 mg, 0.528 mmol, 1.5 eq), activated molecular sieves and pyrophosphate **2.7** (1.5 g, 1.7 mmol, 4.8 eq). Trimethylphosphate and tributylamine was dried over molecular sieves for 15 h. Trimethylphosphate (2.5 mL) and proton sponge were added to 5-azidomethyl-deoxyuridine **4.1** in a flask under argon at -15 °C. Phosphorus oxychloride (40 µL, 0.43 mmol, 1.2 eq) was added dropwise and stirred for 30 min. A mixture of pyrophosphate **2.7** in DMF (1.6 mL) and tributylamine (339 µL) was added, then the reaction was slowly warmed up to RT and stirred for 3 h. 1 M TEAB buffer (7.5 mL, pH 7.5) was added to quench the reaction and the solvent was removed *in vacuo*. The residue was dissolved in deionised water (3.0 mL) and mixed with 3 M NaCl aqueous solution (300 µL) and ethanol (9.0 mL) to precipitate the triphosphate. To obtain better yields, the precipitation was repeated three times on the supernatant (the supernatant was evaporated and the residue was re-dissolved in water each time) until

further precipitation did not yield any product. The precipitate was dissolved in water and purified by RP-HPLC (column L with TEAB buffer system I, 3.5% to 30% eluent B in 50 min and monitored by UV absorption at 290 nm). Compound **4.2** (258 OD_{265}, 26.5 µmol, ε_{265} = 9720 $M^{-1}cm^{-1}$, calculated from the nucleoside measurement) was obtained in 7.5% yield as a TEA salt. M_w = 522.991 g/mol (Chemical Formula: $C_{10}H_{16}N_5O_{14}P_3$)

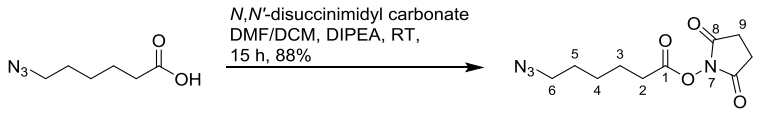
HRMS: [ESI^-, H_2O] Calculated mass: 521.9828 $[M-H]^-$; found *m/z*: 521.9951 ($[M-H]^-$, 100).

^{1}H NMR (400 MHz , D_2O) δ 8.08 (s, 1 H, H-6), 6.33 (app t, *J* = 6.4 Hz, 1 H, H-1′), 4.72 – 4.68 (m, 1 H, H-3′), 4.33 – 4.25 (m, 2 H, H-5′), 4.22 (s, 2 H, H-7), 4.21 – 4.17 (m, 1 H, H-4′), 3.20 (q, *J* = 7.3 Hz, $Et_2NH^+C\boldsymbol{H}_2CH_3$), 2.44 – 2.38 (m, 2 H, H-2′), 1.28 (t, *J* = 7.3 Hz, $Et_2NH^+CH_2C\boldsymbol{H}_3$)

^{13}C NMR (126 MHz, D_2O) δ 178.5 (C4), 152.0 (C2), 141.0 (C6), 109.6 (C5), 85.7 (C4′), 85.2 (C1′), 70.0 (C3′), 65.0 (C5′), 47.1 (C7), 46.6 ($Et_2NH^+\boldsymbol{C}H_2CH_3$), 38.7 (C2′), 8.2 ($Et_2NH^+CH_2\boldsymbol{C}H_3$)

^{31}P NMR (162 MHz, D_2O) δ -6.5 (d, *J* = 20 Hz, P_γ), -11.5 (d, *J* = 20 Hz, P_α), -22.6 (app t, *J* = 20 Hz, P_β)

2,5-Dioxopyrrolidin-1-yl-6-azidohexanoate (4.4, 6-Azido NHS ester)[211]



4.3 **4.4**

6-Azidohexanonic acid **4.3** (840 mg, 5.34 mmol, 1.0 eq), molecular sieves and *N*, *N*′-disuccinimidyl carbonate (1.50 g, 5.86 mmol, 1.1 eq) were dissolved in DCM/DMF(1:2 *v/v*, 12 mL) in a flask. DIPEA (1.8 mL, 10 mmol, 1.9 eq) was added and

the reaction mixture was stirred at RT for 15 h. The solvent was removed *in vacuo*, then the residue was re-dissolved in EtOAc (20 mL) and washed with water (20 mL). The organic layer was dried over anhydrous sodium sulphate and filtered. The solvent was removed *in vacuo* and the residue was purified by column chromatography on silica (0% to 6% EtOAc in DCM). Compound **4.4** (1.2 g, 4.7 mmol) was obtained in 88% yield as clear oil. M_w = 254.25 g/mol (Chemical Formula: $C_{10}H_{14}N_4O_4$)

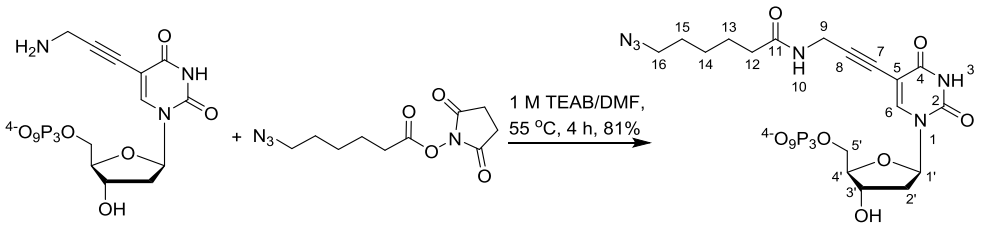
R_f: 0.7 [DCM/EtOAc 7:3 *v/v*] Stain: Mary's reagent (4,4′-*Bis*(dimethylamino)benzhydrol 0.4 g in acetone 100 mL)

LRMS: [ESI^+, MeCN] *m/z*: 277.1 ([$M+Na]^+$, 100).

^{1}H NMR (400 MHz, $CDCl_3$) δ 3.29 (t, *J* = 6.7 Hz, 2 H, H-6), 2.81 (s, 4 H, H-9), 2.62 (t, *J* = 7.4 Hz, 2 H, H-2), 1.78 (app quin, *J* = 7.4 Hz, 2 H, H-3), 1.70 – 1.58 (m, 2 H, H-5), 1.56 – 1.43 (m, 2 H, H-4)

^{13}C NMR (101 MHz, $CDCl_3$) δ 168.8 (C8), 168.0 (C1), 50.7 (C6), 30.4 (C2), 28.0 (C5), 25.5 (C4), 25.2 (C9), 23.8 (C3)

5-(3-(6-azidohexanamido)prop-1-yn-1-yl)-2′-deoxyuridine-5′-*O*-triphosphate (4.5, AHP dUTP)



2.11 **4.4** **4.5**

The freeze-dried amino dUTP **2.11** (111 OD_{290}, 10.5 μmol, 1.0 eq) was dissolved in 1 M TEAB buffer (300 μL, pH 7.5). The 6-azido NHS ester **4.4** (6 mg, 24 μmol, 2.3 eq) in DMF (300 μL) was added to the triphosphate solution. The mixture was kept at 55 °C for

4 h and the solvent was removed *in vacuo.* The residue was purified by RP-HPLC (Column L with TEAB buffer system II, 15% to 60% buffer B in 50 min and monitored by UV absorption at 305 nm). Compound **4.5** (90 OD_{290}, 8.5 μmol) was obtained in 81% yield as a TEA salt. M_w = 660.36 g/mol (Chemical Formula: $C_{18}H_{27}N_6O_{15}P_3$).

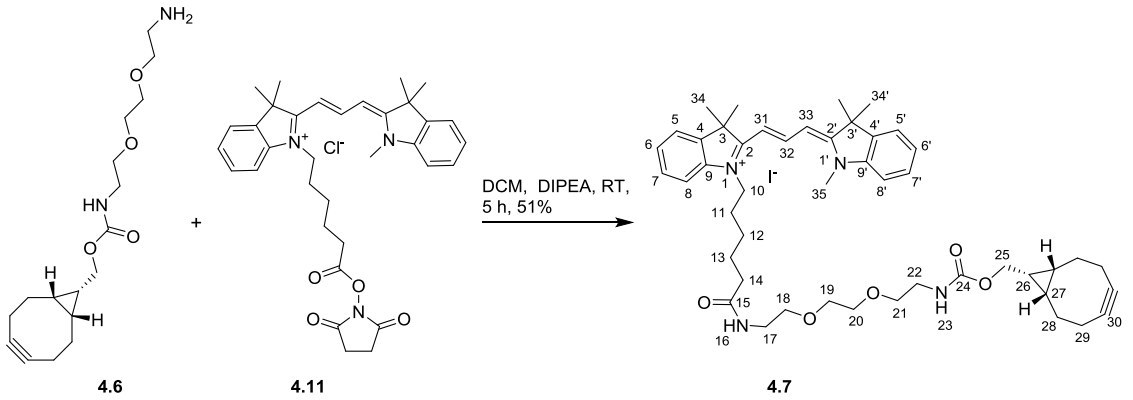
HRMS: [ESI^-, H_2O] calculated mass: 659.0669 $[M-H]^-$; found *m/z*: 659.0632 ($[M-H]^-$, 100).

^{1}H NMR (500 MHz, D_2O) δ 8.08 (s, 1 H, H-6), 6.24 (app t, *J* = 6.6 Hz, 1 H, H-1′), 4.66 – 4.57 (m, 1 H, H-3′), 4.28 – 4.10 (m, 5 H, H-4′, H-5′, H-9), 3.26 (t, *J* = 6.9 Hz, 2 H, H-16), 3.15 (q, *J* = 7.3 Hz, $Et_2N^+C\boldsymbol{H}_2CH_3$), 2.39 – 2.29 (m, 2 H, H-2′), 2.26 (t, *J* = 7.3 Hz, 2 H, H-12), 1.65 – 1.49 (m, 4 H, H-13, H-15), 1.34 (tt, *J* = 9.1, 6.3 Hz, 2 H, H-14), 1.23 (t, *J* = 7.3 Hz, $Et_2N^+CH_2C\boldsymbol{H}_3$)

^{13}C NMR (126 MHz, D_2O) δ 177.2 (C11), 144.9 (C6), 99.5 (C5), 89.7 (C8), 85.9 and 85.8 (C1′, C4′), 75.0 (C7), 70.3 (C3′), 65.4 (C5′), 51.4 (C16), 47.0 ($Et_3N^+\boldsymbol{C}H_2CH_3$), 39.1 (C2′), 35.8 (C12), 30.0 (C9), 28.1 (C15), 26.3 (C14), 25.2 (C13), 8.6 ($Et_2N^+CH_2\boldsymbol{C}H_3$)

^{31}P NMR (202 MHz, D_2O) δ -5.2 (d, *J* = 20 Hz, P_γ), -10.1 (d, *J* = 20 Hz, P_α), -21.3 (app t, *J* = 20 Hz, P_β)

1-(19-((1R,8S,9s)-bicyclo[6.1.0]non-4-yn-9-yl)-6,17-dioxo-10,13,18-trioxa-7,16-diazanonadecan-1-yl)-3,3-dimethyl-2-((E)-3-((Z)-1,3,3-trimethylindolin-2-ylidene)prop-1-en-1-yl)-*3H*-indol-1-ium iodide (4.7, Cy3-BCN)



4.6 4.11 4.7

Cy3 NHS ester[258] **4.11** (30 mg, 0.051 mmol, 1.0 eq) was dissolved in anhydrous DCM (1.5 mL), then anhydrous DIPEA (30 µL, 0.17 mmol, 3.3 eq) and BCN amine **4.6** (25 mg, 0.077 mmol, 1.5 eq) in DCM (0.5 mL) were added. The reaction mixture was stirred at RT for 5 h, diluted with DCM (8 mL) and washed with aqueous saturated potassium iodide (10 mL). The organic layer was dried over anhydrous sodium sulfate, filtered and evaporated. The residue was dissolved in DCM (2.0 mL) and added to cold diethyl ether (20 mL) to afford a red precipitate. The precipitate was filtered off and purified by column chromatography on silica (0% to 5% MeOH in DCM). Compound **4.7** (23 mg, 0.026 mmol) was obtained in 51% yield as a red salt. M_w = 890.9 g/mol (Chemical Formula: $C_{47}H_{63}IN_4O_5$).

R_f: 0.37 [DCM/MeOH 9:1 *v/v*]

LRMS: [ESI$^+$, MeCN], *m/z*: 763.1 ([M-I]$^+$, 100); [ESI$^-$, MeCN] *m/z*: 126.8 (I$^-$, 100)

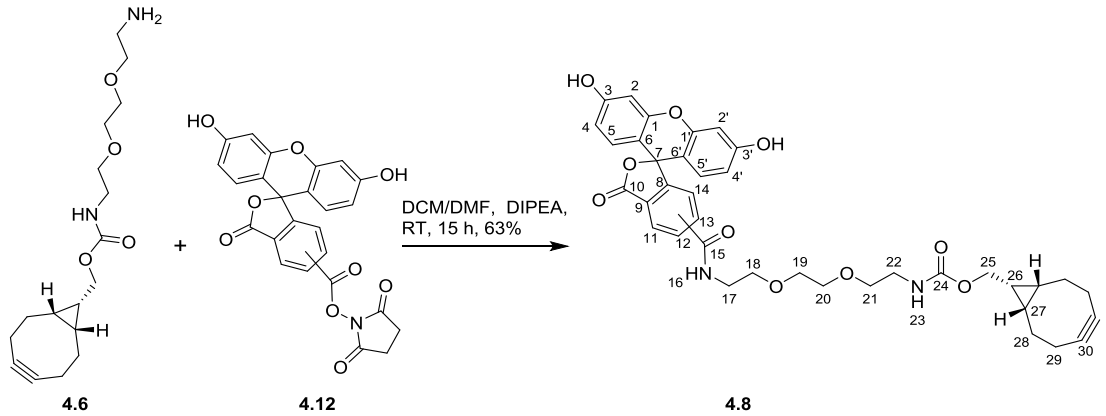
HRMS: [ESI$^+$, MeCN] Calculated mass [M-I]$^+$: 763.4793; found *m/z*: 763.4784.

^{1}H NMR (500 MHz ,$CDCl_3$) δ 8.44 (app t, *J* = 13.4 Hz, 1 H, H-32), 7.44 – 7.38 (m, 3 H, H-7, H-7′, H-31), 7.36 (d, *J* = 7.3 Hz, 2 H, H-5, H-5′), 7.30 – 7.22 (m, H-6, H-6′, H-33,

overlapped with solvent residue), 7.13 (2 × d, *J* = 7.9 Hz, 2 H, H-8, H-8′), 6.88 (br. s, H 16, 1 H), 5.36 (br. s, H-23, 1 H), 4.18 (t, *J* = 7.9 Hz, 2 H, H-10), 4.14 (d, *J* = 8.2 Hz, 2 H, H-25), 3.83 (s, 3 H, H-35), 3.65 – 6.62 (m, 4 H, H-19, H-20), 3.60 (t, *J* = 5.6 Hz, 2 H, H-18), 3.58 (t, *J* = 5.0 Hz, 2 H, H-21), 3.46 (app q, *J* = 5.6 Hz, 2 H, H-17), 3.39 (app q, *J* = 5.0 Hz, 2 H, H-22), 2.37 (t, *J* = 7.1 Hz, 2 H, H-14), 2.32 – 2.16 (m, 6 H, H-29, H-28a), 1.94 – 1.85 (m, 2 H, H-11), 1.84 – 1.75 (m, 2 H, H-13), 1.73 (s, 6 H, H-34), 1.71 (s, 6 H, H-34′), 1.70 – 1.65 (m, 2 H, H-12), 1.63 – 1.52 (m, 2 H, H-28b), 1.40 – 1.31 (m, 1 H, H-26), 0.99 – 0.87 (m, 2 H, H-27)

^{13}C NMR (126 MHz , $CDCl_3$) δ = 174.6 and 174.1 (C2, C2′), 174.0 (C15), 157.3 (C24), 151.2 (C32), 143.2 and 142.2 (C9, C9′), 141.0 and 140.9 (C4, C4′), 129.4 and 129.3 (C7, C7′), 125.8 and 125.7 (C6, C6′), 122.5 and 122.4 (C5, C5′), 111.4 and 111.2 (C8, C8′), 105.6 (C31), 104.9 (C33), 99.3 (C30), 70.8, 70.6, 70.5 and 70.2 (C18, C19, C20, C21), 63.1 (C25), 49.4 and 49.2 (C3, C3′), 45.3 (C10), 41.2 (C22), 39.4 (C17), 36.9 (C14), 33.4 (C35), 29.5 (C28), 28.6 (2 × C34, 2 × C34′), 28.5 (C11), 27.7 (C12), 25.9 (C13), 21.9 (C29), 20.5 (C27), 18.3 (C26)

5(6)-(14-(1R,8S,9s)-bicyclo[6.1.0]non-4-yn-9-yl)-1,12-dioxo-5,8,13-trioxa-2,11-diazatetradecan-1-yl)-3',6'-dihydroxy-3-oxo-*3H*-spiro[2-benzofuran-1,9'-xanthene] (4.8, FAM-BCN)



4.6 4.12 4.8

5(6)-FAM NHS ester **4.12** (30 mg, 0.063 mmol, 1.0 eq) was dissolved in anhydrous DCM/DMF (2.5 mL, 4:1, *v/v*), then anhydrous DIPEA (30 µL, 0.17 mmol, 2.7 eq) and

BCN amine **4.6** (31 mg, 0.096 mmol, 1.5 eq) were added. The reaction mixture was stirred at RT for 15 h. The solvent was removed *in vacuo* and the residue was purified by column chromatography on silica (0% to 15% MeOH in DCM). Compound **4.8** (27 mg, 0.040 mmol) was obtained in 63% yield as a yellow solid. M_w = 682.73 g/mol (Chemical Formula: $C_{38}H_{38}N_2O_{10}$). Carboxyfluorescein is pH dependent and exists as different forms.[217] The structure was shown as the dianionic form in Chapter 4, which has better fluorescent property.

R_f: 0.36 [DCM/MeOH 9:1 *v/v*]

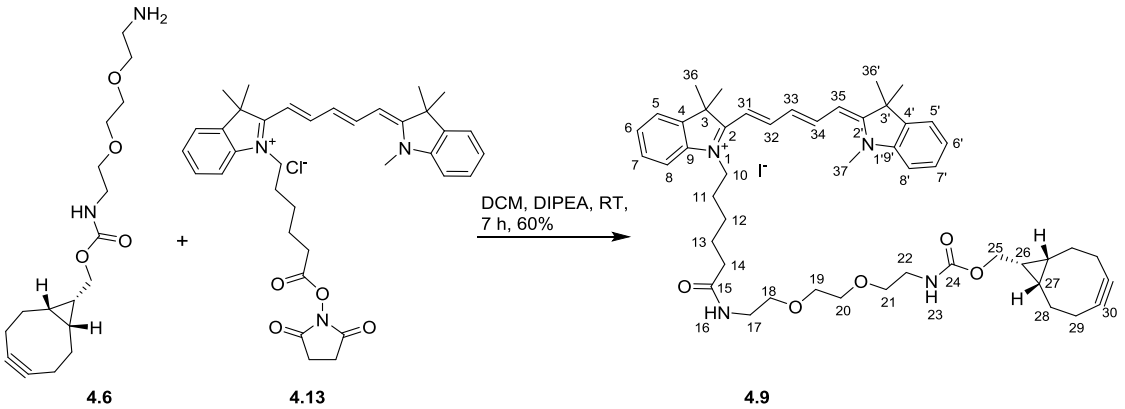
LRMS: [ESI$^+$, MeOH]: *m/z*: 683.1([M+H]$^+$, 100); [ESI$^-$, MeOH] *m/z*: 681.1 ([M-H]$^-$, 100).

HRMS: [ESI$^-$, MeOH] Calculated mass [M-H]$^-$: 683.2600; found *m/z*: 683.2598.

^{1}H NMR (500 MHz, CD_3OD) δ 8.13 – 8.09 (m, 1 H, Ar-H), 8.04 (d, *J* = 8.0 Hz, 1 H, Ar H), 7.61 (s, 1 H, Ar-H), 6.65 (d, *J* = 2.4 Hz, 2 H, H-2, H-2′), 6.59 (d, *J* = 8.8 Hz, 2 H, H-5, H-5′), 6.51 (dd, *J* = 8.8, 2.4 Hz, 2 H, H-4, H-4′), 4.03 (d, *J* = 8.0 Hz, 2 H, H-25), 3.55 (t, *J* = 5.5 Hz, 2 H, H-18), 3.54 – 3.49 (m, 4 H, H-19, H-20), 3.47 (t, *J* = 5.5 Hz, 2 H, H-17), 3.40 (t, *J* = 5.5 Hz, 2 H, H-21), 3.13 (t, *J* = 5.5 Hz, 2 H, H-22), 2.23 – 2.05 (m, 6 H, H-29, H-28a), 1.58 – 1.46 (m, 2 H, H-28b), 1.31 – 1.22 (m, 1 H, H-26), 0.91 – 0.80 (m, 2 H, H-27).

^{13}C NMR (126 MHz, CD_3OD) δ 169.2 (C10), 166.9 (C15), 160.6 (C1, C1′), 157.8 (C24), 152.8 (C3, C3′), 140.7 (C9), 129.0 (ArC, C5, C5′), 124.9 (ArC), 122.9 (ArC), 112.6 (C4, C4′), 109.7 (C6, C6′), 102.2 (C2, C2′), 98.1 (C30), 69.82 and 69.78 (C19, C20), 69.6 (C21), 68.9 (C18), 62.3 (C25), 40.2 (C22), 39.6 (C17), 28.7 (C28), 20.5 (C29), 20.0 (C27), 17.5 (C26).

1-(19-((1R,8S,9s)-bicyclo[6.1.0]non-4-yn-9-yl)-6,17-dioxo-10,13,18-trioxa-7,16-diazanonadecan-1-yl)-3,3-dimethyl-2-((1E,3E)-5-((Z)-1,3,3-trimethylindolin-2-ylidene)penta-1,3-dien-1-yl)-*3H*-indol-1-ium iodide (4.9, Cy5-BCN)



Cy5 NHS ester[258] (30 mg, 0.049 mmol, 1.0 eq) was dissolved in anhydrous DCM (2.0 mL), then anhydrous DIPEA (30 µL, 0.17 mmol, 3.5 eq) and BCN amine **4.6** (25 mg, 0.077 mmol, 1.6 eq) were added. The reaction mixture was stirred at RT for 7 h, diluted with DCM (7 mL) and washed with saturated aqueous potassium iodide (10 mL). The organic layer was dried over sodium sulphate, filtered and evaporated. The crude solid was purified by column chromatography on silica (0% to 10% MeOH in DCM). Compound **4.9** (27 mg, 0.030 mmol) was obtained in 60% yield as a dark blue salt. M_w = 916.40 g/mol (Chemical Formula: $C_{49}H_{65}IN_4O_5$).

R_f: 0.26 [DCM/MeOH 95:5 *v/v*]

LRMS: [ESI$^+$, MeCN] *m/z*: 789.5 ([M-I]$^+$, 100)

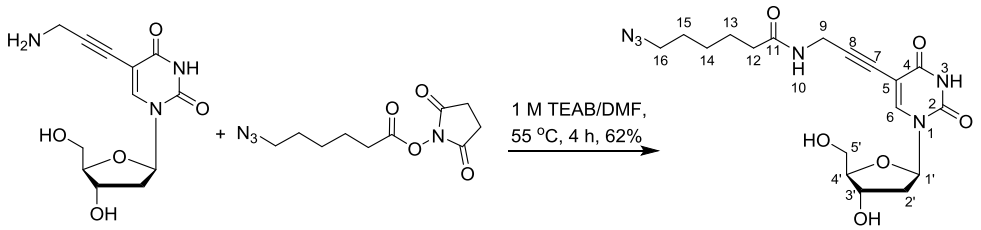
HRMS: [ESI$^+$, MeCN] Calculated mass [M-I]$^+$: 789.4950; found *m/z*: 789.4940.

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.11 – 7.90 (m, 2 H, H-32, H-34), 7.42 – 7.32 (m, 4 H, H-5, H-5′, H-7, H-7′), 7.26 – 7.18 (m, 2 H, H-6, H-6′), 7.11 (d, *J* = 8.0 Hz, 1 H, H-8), 7.09 (d, *J* = 8.0 Hz, 1 H, H-8′), 7.02 (t, *J* = 12.5 Hz, 1 H, H-33), 6.80 (br. s, H-16), 6.46 (d, *J* = 13.7 Hz, 1 H, H-31), 6.36 (d, *J* = 13.6 Hz, 1 H, H-35), 5.34 (br. s, H-23), 4.13 (d,

J = 8.0 Hz, 2 H, H-25), 4.09 (t, *J* = 7.3 Hz, 2 H, H-10), 3.66 (s, 3 H, H-37), 3.63 (s, 4 H, H-19, H-20), 3.59 (t, *J* = 5.6, 2 H, H-18), 3.57 (t, *J* = 5.4, 2 H, H-21), 3.46 (app q, *J* = 5.6 Hz, 2 H, H-17), 3.38 (app q, *J* = 5.4 Hz, 2 H, H-22), 2.34 (t, *J* = 7.3, 2 H, H-14), 2.29 – 2.16 (m, 6 H, H-29, H-28a), 1.89 – 1.80 (m, 2 H, H-11), 1.80 – 1.74 (m, 2 H, H-13), 1.74 (s, 6 H, H-36), 1.72 (s, 6 H, H-36′), 1.62 – 1.52 (m, 4 H, H-12, H-28b), 1.39 – 1.31 (m, 1 H, H-26), 0.97 -0.86 (m, 2 H, H-27).

^{13}C NMR (126 MHz, $CDCl_3$) δ 172.3 (C2), 172.0 (C2′), 171.8 (C15), 155.9 (C24), 152.5 (C32), 151.9 (C34), 141.8 and 140.9 (C9, C9′), 140.2 and 139.8 (C4, C4′), 127.8 and 127.6 (C7, C7′), 125.8 (C33), 124.3 and 124.0 (C6, C6′), 121.2 and 121.1 (C5, C5′), 109.8 and 109.3 (C8, C8′), 103.4 (C31), 102.9 (C35), 97.8 (C30), 69.3, 69.2, 69.1 and 68.7 (C18, C19, C20, C21), 61.6 (C25), 48.4 and 48.1 (C3, C3′), 43.5 (C10), 39.8 (C22), 38.0 (C17), 35.2 (C14), 31.08 (C37), 28.0 (C28), 27.1 (2× C36, 2× C36′), 26.2 (C11), 25.5 (C12), 24.2 (C13), 20.4 (C29), 19.1 (C27), 16.8 (C26).

5-(3-(6-azidohexanamido)prop-1-yn-1-yl)-2′-deoxyuridine (5.3, **AHP dU)**



2.2 **4.4** **5.3**

The amino dU **2.2** (6 mg, 21 µmol, 1.0 eq) was dissolved in deionised water (400 µL) and 1 M TEAB buffer (400 µL, pH 7.5). The 6-azido NHS ester **4.4** (6 mg, 24 µmol, 1.1 eq) in DMF (800 µL) was added and was kept at 55 °C for 4 h. The solvent was removed *in vacuo* and the residue was purified by column chromatography on silica (0% to 7% MeOH in DCM). Compound **5.3** (5.5 mg, 13 µmol) was obtained in 62% yield. M_w = 420.18 g/mol (Chemical Formula: $C_{18}H_{24}N_6O_6$).

R_f: 0.28 [DCM/MeOH 9:1 *v/v*]

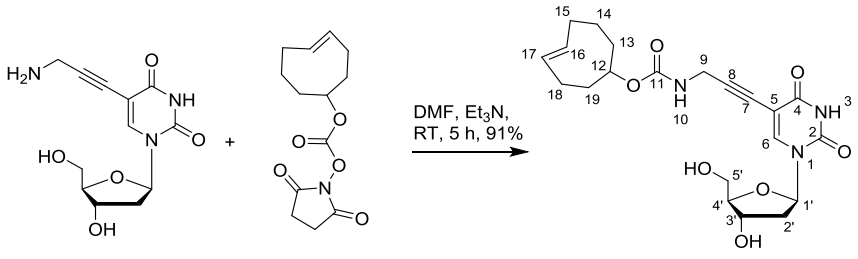
LRMS: [ESI$^+$, MeOH] *m/z*: 421.3 ([M+H]$^+$, 65); 443.2 ([M+Na]$^+$, 100); 863.4 ([2M+Na]$^+$, 18). [ESI$^-$, MeOH]: 419.2 ([M-H]$^-$, 100).

HRMS: [ESI$^+$, MeOH] Calculated mass [M+H]$^+$: 421.1830; found *m/z*: 421.1831.

^{1}H NMR (400 MHz, $(CD_3)_2CO$) δ 10.11 (br. s, 1 H, H-3), 8.16 (s, 1 H, H-6), 7.35 (br. s, 1 H, H-10), 6.14 (t, *J* = 6.7 Hz, 1 H, H-1′), 4.41 - 4.35 (m, 1 H, H-3′), 4.33 (d, *J* = 4.1 Hz, 1 H, 3′-OH), 4.28 (t, *J* = 4.6 Hz, 1 H, 5′-OH), 4.02 (d, *J* = 5.5 Hz, 2 H, H-9), 3.85 (app q, *J* = 2.9 Hz, 1 H, H-4′), 3.71 (ddd, *J* = 11.9, 4.6, 2.9 Hz, 1 H, H-5′a), 3.67 (ddd, *J* = 11.9, 4.6, 2.9 Hz, 1 H, H-5′b), 3.20 (t, *J* = 6.9 Hz, 2 H, H-16), 2.21 – 2.13 (m, 2 H, H-2′), 2.09 (t, *J* = 7.4 Hz, 2 H, H-12), 1.51 (app quin, *J* = 7.4 Hz, 2 H, H-13), 1.47 (app quin, *J* = 6.9 Hz, 2 H, H-15),1.32 – 1.21 (m, 2 H, H-14)

^{13}C NMR (101 MHz, $(CD_3)_2CO$) δ 171.9 (C11), 161.3 (C4), 149.5 (C2), 143.8 (C6), 98.7 (C5), 89.1 (C8), 88.0 (C4′), 85.4 (C1′), 74.3 (C7), 71.1 (C3′), 61.7 (C5′), 51.0 (C16), 40.8 (C2′), 35.3 (C12), 28.9 (C9), 28.4 (C15), 26.1 (C14), 24.9 (C13)

5-(3-(((E)-cyclooct-4-en-1-yl)oxycarbonylamino)prop-1-yn-1-yl)-2′-deoxyuridine (5.4, TCO-S dUTP)[139]



2.2 3.3 5.4

The amino dU **2.2** (24 mg, 85 µmol, 1.5 eq) was dissolved in DMF (500 µL). The TCO-S NHS ester **3.3** (15 mg, 56 µmol, 1.0 eq) in DMF (500 µL) and TEA (100 µL, 717 µmol, 12.8 eq) were added into the solution of amino dU **2.2**. The reaction mixture was stirred at RT for 5 h and the solvent was removed *in vacuo*. The residue was purified by column

chromatography on silica (0% to 6% MeOH in DCM). Compound **5.4** (22 mg, 51 µmol) was obtained in 91% yield. M_w = 433.46 g/mol (Chemical Formula: $C_{21}H_{27}N_3O_7$).

R_f: 0.37 [DCM/MeOH 90:10 *v/v*]

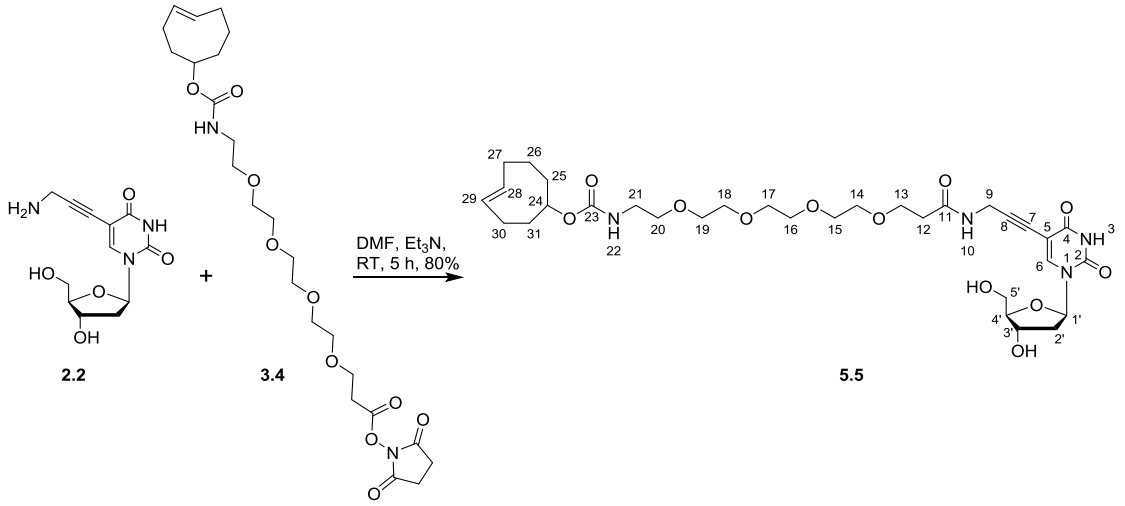
LRMS: [ESI^+, MeCN] *m/z*: 456.1 ($[M+Na]^+$, 100)

HRMS: [ESI^+, MeCN] Calculated mass $[M+H]^+$: 456.1741; found *m/z*: 456.1741.

^{1}H NMR (400 MHz, $(CD_3)_2CO$) δ 8.24 (s, 1 H, H-6), 6.44 (br. s, 1 H, H-10), 6.26 (t, *J* = 6.7 Hz, 1 H, H-1′), 5.67 – 5.54 (m, 1 H, H-17), 5.48 (ddd, *J* = 15.9, 10.8, 3.3 Hz, 1 H, H-16), 4.53 – 4.47 (m, 1 H, H-3′), 4.36 – 4.28 (m, 1 H, H-12), 4.13 – 3.99 (m, 2 H, H-9), 3.97 (d, *J* = 2.8 Hz, 1 H, H-4′), 3.82 (dd, *J* = 11.7, 2.8, 1 H, H-5′a), 3.79 (dd, *J* = 11.7, 2.8, 1 H, H-5′b), 2.36 – 2.22 (m, 5 H, H-18, H-2′, H-13a), 2.01 – 1.82 (m, 4 H, H-15, H-13b, H-14a), 1.81 – 1.51 (m, 3 H, H-19, H-14b)

^{13}C NMR (101 MHz, $(CD_3)_2CO$) δ 161.1 (C4), 155.7 (C11), 149.5 (C2), 143.6 (C6), 134.9 (C17), 132.5 (C16), 98.8 (C5), 89.3 (C8), 88.0 (C4′), 85.4 (C1′), 80.1 (C12), 74.3 (C7), 71.1 (C3′), 61.7 (C5′), 40.9 (C15), 40.7 (C2′), 38.5 (C19), 33.9 (C13), 32.3 (C18), 30.8 and 30.7 (C14, C9)

5-(21-((E)-cyclooct-4-en-1-yloxy)-5,21-dioxo-8,11,14,17-tetraoxa-4,20-diazahenicos-1-yn-1-yl)-2′-deoxyuridine (5.5, TCO-L dU)



The amino dU **2.2** (7 mg, 25 µmol, 1.7 eq) was dissolved in DMF (500 µL). The TCO-L NHS ester **3.4** (8 mg, 15 µmol, 1.0 eq) in DMF (500 µL) and TEA (100 µL, 717 µmol, 47.8 eq) were added into the solution of amino dU **2.2**. The mixture was stirred at RT for 5 h and the solvent was removed *in vacuo*. The residue was purified by column chromatography on silica (0% to 7% MeOH in DCM). Compound **5.5** (8 mg, 12 µmol) was obtained in 80% yield. M_w = 680.75 g/mol (Chemical Formula: $C_{32}H_{48}N_4O_{12}$)

R_f: 0.34 [DCM/MeOH 9:1 *v/v*]

LRMS: [ESI$^+$, MeOH] *m/z*: 703.2 ([M+Na]$^+$, 100).

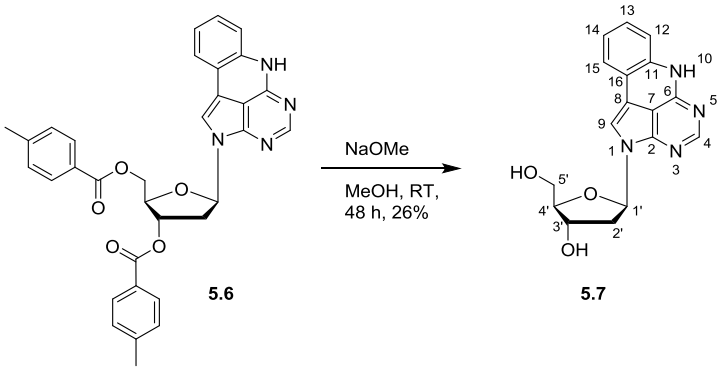
HRMS: [ESI$^+$, MeOH] Calculated mass [M+H]$^+$: 703.3161; found *m/z*: 703.3154.

^{1}H NMR (400 MHz , $(CD_3)_2CO$) δ 10.16 (br. s, 1 H, H-3), 8.34 (s, 1 H, H-6), 7.65 (br. s, 1 H, H-10), 6.27 (t, *J* = 6.6 Hz, 1 H, H-1′), 6.06 (br. s, 1 H, H-22), 5.67 – 5.55 (m, 1 H, H-29), 5.48 (ddd, J = 15.9, 10.8, 3.3 Hz, 1 H, H-28), 4.57 – 4.48 (m, 1 H, H-3′), 4.35 – 4.26 (m, 1 H, H-24), 4.20 – 4.13 (m, 2 H, H-9), 3.97 (app q, *J* = 2.8 Hz, 1 H, H-4′), 3.85 (dd, *J* = 11.7, 2.8 Hz, 1 H, H-5′a), 3.80 (dd, *J* = 11.7, 2.8 Hz, 1 H, H-5′b), 3.73 (t, *J* = 6.1 Hz, 2 H, H-13), 3.65 – 3.55 (m, 12 H, H-14 to 19), 3.52 (t, *J* = 5.6 Hz, 2 H, H-20), 3.27 (app q, *J* = 5.6 Hz, 2 H, H-21), 2.46 (t, *J* = 6.1 Hz, 2 H, H-12), 2.36 – 2.24 (m, 5 H,

H-2′, H-30, H-25a), 2.00 – 1.83 (m, 4 H, H-27, H-25b, H-26a), 1.81 – 1.48 (m, 3H, H-26b, H-31)

^{13}C NMR (101 MHz , $(CD_3)_2CO$) δ 170.3 (C11), 161.1 (C4), 149.5 (C2), 143.9 (C6), 134.9 (C29), 132.5 (C28), 98.7 (C5), 89.0 (C8), 88.0 (C4′), 85.4 (C1′), 79.6 (C24), 74.4 (C7), 71.0 (C3′), 70.32, 70.27, 70.26, 70.12 and 70.04 (C14 to 19), 70.00 (C20), 66.9 (C13), 61.6 (C5′), 41.0 (C27), 40.8 (C2′), 40.5 (C21), 38.5 (C31), 36.3 (C12), 33.9 (C25), 32.3 (C30), 30.8 (C26), 28.9 (C9)

1′-(2,3,5,6-Tetraazaaceanthrylen-2(6*H*)-yl)-2′-deoxyriboside (5.7, Quadracyclic dR)[48]



The 5′,3′-di-*O*-*p*-toluoyl-quadracyclic adenine analogue-2′-deoxyriboside **5.6** (60 mg, 0.11 mmol, 1.0 eq) and NaOMe (54 mg, 1.0 mmol, 9.1 eq) were left under vacuum for 15 h. Anhydrous MeOH (2.5 mL) was added and the reaction was stirred at RT for 48 h. The solvent was removed *in vacuo* and the residue was purified by column chromatography on silica (0% to 5% MeOH in DCM). Compound **5.7** (9 mg, 28 µmol) was obtained in 26% yield. M_w = 324.12 g/mol (Chemical Formula: $C_{17}H_{16}N_4O_3$).

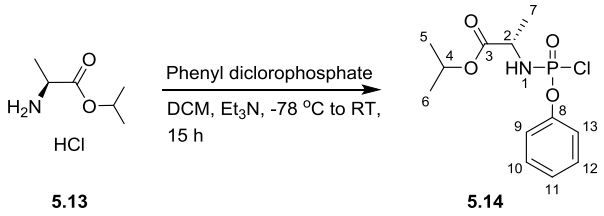
R_f: 0.26 [DCM/MeOH 9:1 *v/v*]

LRMS: [ESI$^+$, MeOH] *m/z*: 325.2 ([M+H]$^+$, 100)

^{1}H NMR (400 MHz, DMSO-d_6) δ 10.63 (s, 1 H, H-10), 8.09 (s, 1 H, H-4), 7.59 (dd, *J* = 7.7, 1.1 Hz, 1 H, Ar-H), 7.44 (s, 1 H, H-9), 7.20 – 7.16 (m, 2 H, Ar-H), 7.04 (ddd,

J = 7.7, 6.8, 1.8 Hz, 1 H, Ar-H), 6.36 (dd, *J* = 8.1, 5.9 Hz, 1 H, H-1′), 5.28 (d, *J* = 4.0 Hz, 1 H, 3′-OH), 5.16 (dd, *J* = 6.6, 4.8 Hz, 1 H, 5′-OH), 4.42 – 4.37 (m, 1 H, H-3′), 3.87 (td, *J* = 4.8, 2.5 Hz, 1 H, H-4′), 3.62 (dt, *J* = 11.7, 4.8 Hz, 1 H, H-5′a), 3.53 (ddd, *J* = 11.7, 6.6, 4.8 Hz, 1 H, H-5′b), 2.65 (ddd, *J* = 13.1, 8.1, 5.8 Hz, 1 H, H-2′a), 2.23 (ddd, *J* = 13.1, 5.9, 2.5 Hz, 1 H, H-2′b)

(Isopropyl-L-alaninyl)(phenyl) chlorophosphate (5.14, Chlorophosphate)[250]

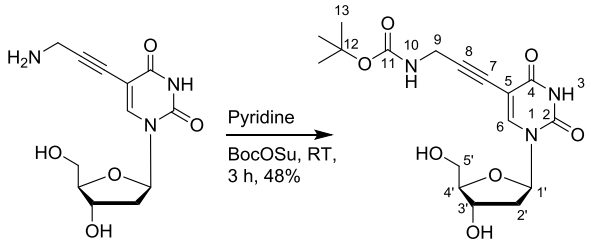


L-Alanine isopropyl ester hydrochloride **5.13** (1.50 g, 8.98 mmol, 1.0 eq) was co-evaporated with anhydrous DCM (2 × 5 mL) and left under vacuum for 15 h. It was dissolved in anhydrous DCM (70 mL) and phenyl dichlorophosphate (1.3 mL, 8.7 mmol, 1.0 eq) was added at -78 ºC. Anhydrous TEA (2.5 mL, 18 mmol, 2.0 eq) was added dropwise over 30 min. The reaction mixture was slowly warmed up to RT and stirred for 15 h. The solvent was removed *in vacuo* and the residue was re-suspended in distilled THF (15 mL) followed by filtering under argon. The filtrate was dried *in vacuo* and left under vacuum for 15 h. Compound **5.14** (2.35 g) was obtained as a yellow oil and used without further purification. M_w = 305.69 g/mol (Chemical Formula: $C_{12}H_{17}ClNO_4P$)

R_f: two isomer: 0.58, 0.66 [DCM/MeOH 9:1 *v/v*] Stain: $KMnO_4$

^{31}P (121 Hz) δ 8.8, 8.4

5-(3-((*t*-Butoxycarbonyl)amino)prop-1-yn-1-yl)-2′-deoxyuridine (5.15, Boc-amino dU)[259]



2.2 **5.15**

Amino dU **2.2** (0.60 g, 2.1 mmol, 1.0 eq) was co-evaporated with anhydrous MeOH (3 × 2 mL), left under vacuum for 15 h and dissolved in anhydrous pyridine (10 mL). *N*-(*t*-Butoxycarbonyloxy) succinimide (BocOSu, 0.46 g, 2.1 mmol, 1.0 eq) was dissolved in pyridine (3.0 mL) and added into the solution containing amino dU **2.2** dropwise at RT. The reaction was stirred at RT for 3 h and the solvent was removed *in vacuo*. The residue was purified by column chromatography on silica (0% to 5% MeOH in DCM with 0.5% TEA). Compound **5.15** (recalculated 0.399 g, 1.0 mmol) was obtained in 48% yield as a white foam (0.493 g), containing 0.9 eq TEA. M_w = 381.38 g/mol (Chemical Formula: $C_{17}H_{23}N_3O_7$).

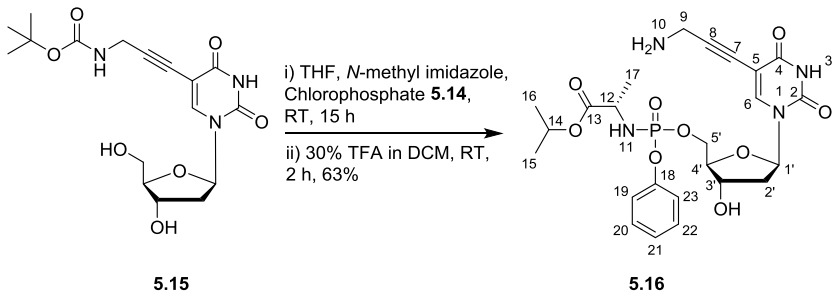
R_f: 0.54 [DCM/MeOH 8:2 *v/v*]

LRMS: [ESI^+, MeOH] *m/z*: 404.2 ($[M+Na]^+$, 100).

^{1}H NMR (400 MHz, DMSO-d_6) δ 8.06 (s, 1 H, H-6), 7.22 (t, *J* = 5.6 Hz, 1 H, H-10), 6.11 (t, *J* = 6.7 Hz, 1 H, H-1′), 4.15 (td, *J* = 4.3, 3.4 Hz, 1 H, H-3′), 3.86 (d, *J* = 5.6 Hz, 2 H, H-9), 3.72 (app q, *J* = 3.4 Hz, 1 H, H-4′), 3.53 (dd, *J* = 11.7, 3.4 Hz, 1 H, H-5′a), 3.48 (dd, *J* = 11.7, 3.4 Hz, 1 H, H-5′b), 2.04 (dd, *J* = 6.7, 4.3 Hz, 2 H, H-2′), 1.40 (s, 9 H, H-13)

^{13}CNMR (100 MHz, DMSO-d_6) δ 161.6 (C4), 155.2 (C11), 149.4 (C2), 143.5 (C6), 98.2 (C5), 90.1 (C8), 87.6 (C4′), 84.7 (C1′), 78.2 (C12), 74.1 (C7), 70.2 (C3′), 61.0 (C5′), 40.0 (C2′), 30.1 (C9), 28.2 (C13)

5-(3-Aminoprop-1-yn-1-yl)-2′-deoxyuridine-5′-*O*-((isopropyl-L-alaninyl)(phenyl))phosphate (5.16, Amino dU masked monophosphate)



5.15 5.16

Chlorophosphate **5.14** and Boc-amino dU **5.15** were dried under vacuum for 15 h in separate flasks. Boc-amino dU **5.15** (0.32 g, 0.8 mmol, 1.0 eq) was dissolved in anhydrous THF (10 mL). *N*-methyl imidazole (0.42 mL, 5.3 mmol, 6.6 eq) and chlorophosphate **5.14** (1.30 g, 4.29 mmol, 5.4 eq) in anhydrous THF (4.2 mL) were added at -78 ºC. The reaction mixture was slowly warmed up to RT and stirred for 15 h. The solvent was removed *in vacuo*, then the residue was re-dissolved in EtOAc (10 mL) and washed with water (10 mL). The organic layer was dried over $MgSO_4$ and filtered, then the solvent was removed *in vacuo*. Anhydrous DCM (3.0 mL) and TCA (1.5 mL) were added to the residue to deprotect the Boc group. The mixture was stirred at RT for 2 h. The solvent was then removed *in vacuo* and the residue was purified by column chromatography on silica (0% to 20% MeOH in DCM). Compound **5.16** (re-calculated 0.28 g, 0.50 mmol, two diastereoisomers) was obtained in 63% yield as a yellow oil (0.40 g), containing 3.1 eq *N*-methyl imidazole. M_w = 550.18 g/mol (Chemical Formula: $C_{24}H_{31}N_4O_9P$).

R_f: 0.46 [DCM/MeOH 8:2 *v/v*]

LRMS: [ESI$^+$, MeOH] *m/z*: 551.2 ([M+H$^+$], 61) 573.2 ([M+Na$^+$], 100).

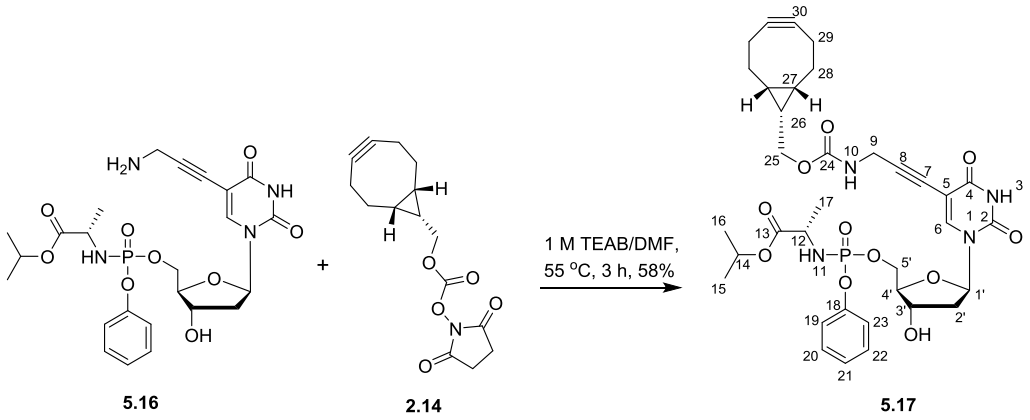
HRMS: [ESI$^+$, MeOH] Calculated mass [M+H]$^+$: 551.1901; found *m/z*: 551.1895.

^{1}H NMR (400 MHz, DMSO-d_6) δ 11.74 (br. s, 1 H, H-3), 7.97 (s, 0.4 H, H-6*) and 7.95 (s, 0.6 H, H-6*), 7.42 – 7.32 (m, 2 H, ArH), 7.26 – 7.14 (m, 3 H, ArH), 6.14 – 6.07 (m, 1 H, H-1′), 6.05 – 5.96 (m, 1 H, H-11), 4.92 – 4.80 (m, 1 H, H-14), 4.29 – 4.15 (m, 3 H, H-3′, H-5′), 4.04 – 3.94 (m, 1 H, H-4′), 3.89 (s, 2 H, H-9), 3.74 – 3.82 (m, 1 H, H-12), 2.20 – 2.06 (m, 2 H, H-2′), 1.28 – 1.20 (m, 3 H, H-17), 1.19 – 1.12 (m, 6 H, H-15, H-16)

^{13}C NMR (100 MHz, DMSO-d_6) δ 161.2 (C4), 144.6 (C6), 129.6 (2 × ArC), 124.6 (ArC), 120.5 (ArC), 120.1 (ArC), 112.8 (C5), 97.4 (C8), 85.6 (C1′), 85.2 (C4′), 70.0 (C3′), 68.0 (C14), 65.9 (C5′), 48.6 (C12), 39.1 (C2′), 29.1 (C9), 21.3 (C15,16), 19.7 (C17)

^{31}P (121 Hz, $CDCl_3$) δ 3.7, 3.5

5-(3-(((1R,8S,9s)-bicyclo[6.1.0]non-4-yn-9-yl)methyloxycarbonylamido)prop-1-yn-1-yl)-2′-dexoyuridine-5′-*O*-((isopropyl-L-alaninyl)(phenyl))phosphate (5.17, BCN-I dU masked monophosphate)



5.16 2.14 5.17

Amino dU masked monophosphate **5.16** (46 mg, 84 µmol, 2.2 eq) was mixed with BCN-I NHS ester **2.14** (11 mg, 38 µmol, 1.0 eq) in DMF (1.0 mL) and 1 M TEAB buffer (300 µL, pH 7.5). Reaction mixture was left at 55 °C for 3 h, then diluted with DCM (10 mL) and washed with deionised water (2 × 10 mL). The organic layer was dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the residue was purified by column chromatography on silica (0% to 5% MeOH in DCM). Compound **5.17** was obtained in

58% yield (16 mg, 22 µmol, two diastereoisomers) as a solid. M_w = 726.27 g/mol (Chemical Formula: $C_{35}H_{43}N_4O_{11}P$).

R_f: 0.24 [DCM/MeOH 95:5 *v/v*]

LRMS: [ESI^+, MeCN] *m/z*: 749.2 ($[M+Na]^+$, 100).

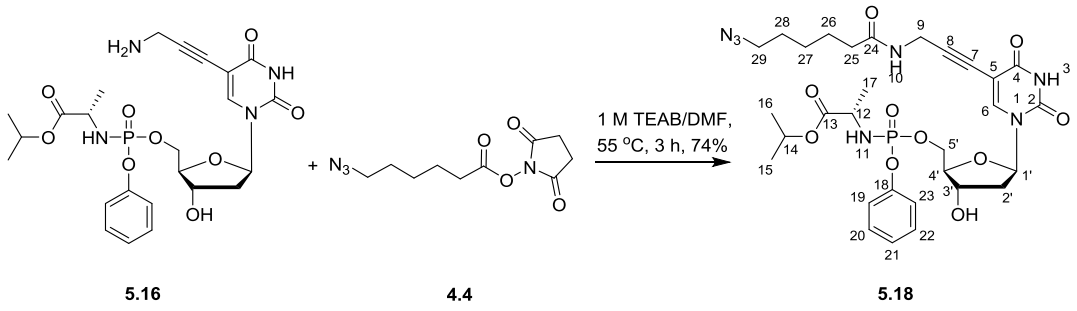
HRMS: [ESI^+, MeCN] Calculated mass $[M+H]^+$: 749.2558; found *m/z*: 749.2556.

^{1}H NMR (400 MHz, $CDCl_3$) δ 9.07 (s, 1 H, H-3), 7.93 (s, 0.4 H, H-6*) and 7.86 (s, 0.6 H, H-6*), 7.38 – 7.29 (m, 2 H, ArH), 7.29 – 7.23 (m, ArH, overlapped with solvent residue), 7.22 – 7.12 (m, 1 H, ArH), 6.24 – 6.17 (m, 1 H, H-1′), 5.69 – 5.53 (m, 1 H, H-10), 5.02 (app spt, *J* = 6.2 Hz) and 5.01 (app spt, *J* = 6.2 Hz) (overlapping peaks, 1 H, H-14), 4.54 – 4.30 (m, 3 H, H-3′, H-5′), 4.26 – 3.96 (m, 7 H, H-4′, H-9, H-12, H-25, H-11), 2.50 – 1.79 (m, 8 H, H-2′, H-29, H-28a), 1.62 – 1.48 (m, 2 H, H-28b), 1.39 (d, *J* = 6.3 Hz, 3 H, H-17), 1.36 – 1.31 (m, 1 H, H-26), 1.28 – 1.20 (m, 6 H, H-15, H-16), 0.97 – 0.87 (m, 2 H, H-27)

^{13}C NMR (100 MHz, $CDCl_3$) δ 149.0 (C2), 143.1 (C6), 129.8 (2 × ArC), 125.4 (ArC), 120.2 (ArC), 120.1 (ArC), 99.5 (C5), 98.8 (C30), 85.4 and 85.3 (C1′, C4′), 70.5 (C3′), 69.7 (C14), 65.7 (C5′), 63.2 (C25), 50.7 (C12), 40.3 (C2′), 31.7 (C9), 29.0 (C28), 21.7 (C15), 21.6 (C16), 21.4 (C29), 21.0 (C26), 20.7 (C17), 20.2 (C27)

^{31}P (121 Hz, $CDCl_3$) δ 4.0

5-(3-(6-Azidohexanamido)prop-1-yn-1-yl)-2′-deoxyuridine-5′-*O*-((isopropyl-L-alaninyl)(phenyl))phosphate (5.18, AHP dU masked monophosphate)



5.16 **4.4** **5.18**

Amino dU masked monophosphate **5.16** (53 mg, 96 µmol, 2.0 eq) was mixed with 6-azido NHS ester **4.4** (12 mg, 47 µmol, 1.0 eq) in DMF (1.0 mL) and 1 M TEAB buffer (300 µL, pH 7.5). Reaction mixture was left at 55 °C for 3 h, then diluted with DCM (10 mL) and washed with deionised water (2 × 10 mL). The organic layer was dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the residue was purified by column chromatography on silica (0% to 5% MeOH in DCM). Compound **5.18** was obtained in 74% yield (24 mg, 35 µmol, two diastereoisomers). M_w = 689.26 g/mol (Chemical Formula: $C_{30}H_{40}N_7O_{10}P$).

R_f: 0.37 [DCM/MeOH 9:1 *v/v*]

LRMS: [ESI$^+$, MeCN] *m/z*: 712.2 ([M+Na]$^+$, 100)

HRMS: [ESI$^+$, MeCN] Calculated mass [M+H]$^+$: 690.2647; found *m/z*: 690.2648

1H NMR (400 MHz, $CDCl_3$) δ 9.66 (s, 0.4 H, H-3*), 9.62 (s, 0.6 H, H-3*), 7.96 (s, 0.4 H, H-6*), 7.93 (s, 0.6 H, H-6*), 7.37 – 7.29 (m, 2 H, ArH), 7.26 – 7.21 (m, 2 H, ArH), 7.21 – 7.15 (m, 1 H, ArH), 6.97 (br. s, 1 H, H-10), 6.26 – 6.19 (m, 1 H, H-1′), 5.01 (app spt, *J* = 5.9 Hz, 1 H, H-14), 4.55 – 4.26 (m, 4 H, H-3′, H-5′, H-11), 4.26 – 3.93 (m, 4 H, H-9, H-12, H-4′), 3.23 (t, *J* = 6.8 Hz, 2 H, H-29), 2.50 – 2.33 (m, 1 H, H-2′a), 2.21 – 1.93 (m, 3 H, H-25, H-2′b), 1.71 – 1.50 (m, 4 H, H-26, H-28), 1.39 (d, *J* = 6.9 Hz, 3 H, H-17), 1.37 – 1.29 (m, 2 H, H-27), 1.28 – 1.16 (m, 6 H, H-15, H-16)

^{13}C NMR (100 MHz, $CDCl_3$) δ 172.7 (C24), 161.9 (C4), 149.2 (C2), 143.2 (C6), 129.9 (2 × ArC), 125.4 (ArC), 120.2 (ArC), 120.1 (ArC), 99.4 (C5), 89.7 (C8), 85.6 (C1′), 85.3 (C4′), 74.7 (C7), 70.4 (C3′), 69.8 (C14), 65.8 (C5′), 51.2 (C29), 50.7 (C12), 40.6 (C2′), 35.7 (C25), 30.2 (C9), 28.6 (C28), 26.3 (C27), 25.0 (C26), 21.7 and 21.6 (C15, C16), 20.8 (C17)

^{31}P (121 Hz, $CDCl_3$) δ 3.5

7.3 General method for oligonucleotide synthesis and purification

Standard and modified DNA phosphoramidites, solid supports, and additional reagents were purchased from Link Technologies and Applied Biosystems Ltd. All oligonucleotides were synthesised on an Applied Biosystems 394 automated DNA/RNA synthesiser using a standard 1.0 μmol scale phosphoramidite cycle of acid-catalysed detritylation, coupling, capping, and iodine oxidation. Stepwise coupling efficiencies and overall yields were determined by the automated trityl cation conductivity monitoring facility and in all cases were >98.0%. All β-cyanoethyl phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use. The coupling time for normal A, G, C, and T monomers was 40 s. Cleavage of oligonucleotides from the solid support and deprotection was achieved by exposure to concentrated aqueous ammonia solution for 60 min at RT followed by heating at 55 °C in a sealed tube for 5 h.

The building blocks for the RNA analogues were prepared using 2′-*t*-butyldimethylsilyl protected RNA phosphoramidite monomers with *t*-butylphenoxyacetyl protection of the A, G and C nucleobases and unprotected U (Sigma-Aldrich). A solution of 0.3 M benzylthiotetrazole in acetonitrile (Link Technologies) was used as the coupling agent, *t*-butylphenoxyacetic anhydride was employed as the capping agent and 0.1 M iodine as

the oxidising agent (Sigma-Aldrich). All RNA phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use, and the coupling time for all monomers was 10 min. Stepwise coupling efficiencies were >96%. Cleavage of RNA oligonucleotides from the solid support and deprotection were achieved by exposure to concentrated aqueous ammonia/ethanol (3:1 *v/v*) at RT for 2 h followed by heating at 55 °C in a sealed tube for 45 min. RNA oligonucleotide solutions were concentrated to ~ 0.2 mL *in vacuo* and lyophilised. The residues were dissolved in DMSO (300 μL) and triethylamine trihydrofluoride (300 μL) was added, after which the reaction mixtures were kept at 65 °C for 2.5 h. Sodium acetate (3 M, 50 μL) and butanol (3 mL) were added with vortexing and the samples were kept at -80°C for 30 min then centrifuged with a speed of 15,000 × g at 4 °C for 10 min. The supernatant was decanted and the precipitate was washed with ethanol (2 × 0.75 mL) then dried under vacuum.

The fully deprotected DNA was purified by RP-HPLC on a Gilson system using HPLC column S and Ammonium acetate system (0% to 35% buffer B in 20 min). Elution was monitored by UV absorption at 295 nm. After HPLC purification, oligonucleotides were desalted using NAP-10 columns (GE Healthcare). For HPLC purification of RNA, TEAB buffer system II was used (0% to 35% buffer B in 20 min). The fractions from HPLC were evaporated without need for additional desalting.

All oligonucleotides were characterised by HPLC-MS using an Acquity UPLC system with a BEH C18 1.7 μm column (Waters) and a Bruker micro-TOF mass spectrometer (ESI^- mode). A gradient of methanol in TEA and hexafluoroisopropanol (HFIP) was employed, increasing from 0% to 70% buffer B over 8 min, with a flow rate of 0.2 mL/min (buffer A: 8.6 mM TEA, 200 mM HFIP in 5% methanol/water; buffer B: 20% buffer A

in methanol). Raw data were processed and deconvoluted using the Data Analysis function of the Bruker Daltronics Compass^{TM} 1.3 software package.

Modified oligonucleotide synthesis

The 5′-terminal 6-FAM, Cy5, phosphate and biotin modifications were introduced by using the corresponding phosphoramidite on the solid phase oligonucleotide synthesiser. The 3′-end BCN modifications were prepared by labelling the amino-functionalised oligonucleotides, which were synthesised on the amino C7 dU resin. The labelling reaction was carried out in 0.5 M Na_2CO_3/$NaHCO_3$ (100 μL, pH 8.7) with 1 mg of BCN-I NHS carbonate **2.14** in DMF (100 μL) at 55 °C for 5 h. The solvent was removed *in vacuo* and the residue was purified by RP-HPLC as described above and followed by denaturing PAGE purification.

Table 7.1. Mass spectrometry characterisations of synthesised oligonucleotides.

Code	Sequences(5′-3′)	Mass Found (Calc.)
T1	CAGTCACTGTACTGCCGACACACATAACC (DNA template)	8775(8776)
T2	CAGTCACAAAACTGCCGACACACATAACC (DNA template)	8777(8778)
P3	FAM-GGTTATGTGTGTCGGCAG (primer)	6138(6138)
T4	CAGUCACUGUACUGCCGACACACAUAACC (RNA template)	9169(9169)
T5	CAGUCACAAAACUGCCGACACACAUAACC (RNA template)	9199(9199)
P6	GCATTCGAGCAACGTAAG (PCR primer)	5532(5533)
P7	GGTTATGTGTGTCGGCAG (PCR primer)	5601(5602)
$P6_p$	Phosphate-GCATTCGAGCAACGTAAG (primer for λ-exonuclease digestion)	5612(5612)
$P7_p$	Phosphate-GGTTATGTGTGTCGGCAG (primer for λ-exonuclease digestion)	5681(5680)
$P6_b$	Biotin-GCATTCGAGCAACGTAAG (primer for streptavidin binding)	5937(5937)
T8	GGTTATGTGTGTCGGCAGTATTGTCAGTGTGAATTCCAGAGTGTGAGATTGTGTGCTGGCGATCTTACGTTGCTCGAATGC (PCR template)	25201(25199)
P9	GTTTGGCTTTAGAGGCTGGAG (primer for T11)	6548(6548)
P10	ACTGCAATACGAATAATGGCTAC (primer for T11)	7040(7041)
$P9_p$	Phosphate-GTTTGGCTTTAGAGGCTGGAG (primer for λ-exonuclease digestion)	6627(6627)
$P10_p$	Phosphate-ACTGCAATACGAATAATGGCTAC (primer for λ-exonuclease digestion)	7119(7120)
Cy5-T_{10}-BCN	Cy5-TTTTTTTTTT-BCN	3900(3897)

FAM = 6-carboxyfluorescein; BCN = Bicyclo[6.1.0]nonyne

7.4 Primer extension protocol

In a 20 µL reaction (1 × buffer), 66 pmol of primer (P3), 132 pmol of template (T1 or T2), 3.2 nmol in total of modified dUTP/dTTP in addition to 3.2 nmol each of dATP, dCTP and dGTP were mixed with the polymerase enzyme (1 unit for Gotaq, Klenow and Therminator™ II, 0.5 unit for KOD) and stock buffer solution (5 × or 10 × solution supplied with the enzymes). In the case of KOD polymerase, 1 mM $MgCl_2$ was added separately to the reaction mixture. The reactions were heated in a heating block or a BIO-RAD T100™ thermal cycler for 1.5 h (Therminator™ II, KOD and Gotaq at 60 or

72 °C, Klenow at 37 °C) and formamide (20 µL) was added before analysis by 20% denaturing PAGE at a constant 20 W of power in 1 × Tri-Borate-EDTA (TBE) buffer. Gel images were recorded using Syngene G: BOX with a transilluminator (302 nm) and the gene-snap imaging software. For the mass spectrometry analysis, two same reactions were carried out, followed by NAP column desalting and lyophilising or by ethanol precipitation. The residue was dissolved in deionised water (15 µL) and analysed on a Bruker microTOF mass spectrometer (ES^- mode).

KOD DNA polymerase was purchased from Merck Millipore; DNA polymerase I large (Klenow) fragment and Therminator^{TM} II polymerases were from New England Biolabs; Gotaq polymerase was from Promega.

Kinetics measurements of modified nucleotide incorporation

Primer extension reactions were prepared using above protocol with different concentrations of dTTP or modified dUTP and heated at 60 °C in a BIO-RAD T100^{TM} thermal cycler. The reactions were then quenched after 30, 60, 90, 120, 150, 180 and 210 s by freezing in liquid nitrogen and mixing with formamide/TBE buffer (20 µL). The reaction mixtures were analysed by 20% denaturing PAGE and quantified using gene-snap imaging software. Reaction rates were calculated as the yield of reaction divided by reaction time. The supplier defined polymerase unit was used as polymerase concentration to calculate the relative catalytic constant K_{cat}.

Melting temperature measurements

Melting temperatures were measured on a BIO-RAD CFX96 real-time PCR instrument by adding SYBR Green (0.6 µL) to the primer extension reactions. When the primer extension was complete, samples were heated to 95 °C for 30 s, cooled to 30 °C, then the

temperature was increased to 95 °C at a rate of 1 or 0.5 °C/s and held (monitored) at each temperature for 5 s. The T_m measurement for the same condition was carried out only once and presented without errors.

Fluorescent labelling of primer extension products

Two same primer extension reactions (2 × 20 µL) were performed as explained in the previous section using modified dUTP, or unmodified dNTP as a negative control. Ethanol precipitation of the products was carried out after primer extension. The precipitates were re-dissolved in 1 × Gotaq green buffer (40 µL) and half of the solution (20 µL) was treated with 6.4 nmol of functionalised fluorophores. When the reaction was complete, another ethanol precipitation was performed. Formamide (20 µL) was added to the labelled and unlabelled products followed by a 20% denaturing PAGE analysis at a constant 20 W of power. For the mass spectrometry analysis, two same labelling reactions were carried out and analysed using HPLC-MS.

7.5 PCR protocol

For template T8, Gotaq polymerase (1 unit), 5 × Gotaq buffer and SYBR Green (0.6 µL, 3.75 ×) were added to samples containing 10 pmol of two primers (P6, P7), 50 pg of template (T8), 5 nmol in total of modified dUTP/dTTP in addition to 5 nmol each of dATP, dCTP and dGTP. The final reaction volume was 20 µL with 1 × reaction buffer. KOD polymerase (0.25 unit) was also used by adding additional 1 mM $MgCl_2$ solution.

For plasmid template T11, KOD polymerase (0.5 unit), 10 × KOD polymerase buffer 1 (pH 8.0), 25 mM $MgCl_2$ (0.8 µL) and SYBR Green (0.6 µL, 3.75 ×) were added to samples containing 10 pmol of two primers (P9, P10), 300 pg of plasmid template (T11),

10 nmol in total of modified dUTP/dTTP in addition to 10 nmol each of dATP, dCTP and dGTP. The final reaction volume was 20 µL with 1 × reaction buffer and 1 mM $MgCl_2$.

Amplification was performed on a BIO-RAD CFX96 real-time PCR instrument using the following procedure: an initial denaturing at 95 ºC for 2 min, followed by 25 cycles of denaturing at 95 ºC for 30 s, annealing at 51 ºC (T8) or 54 ºC (T11) for 30 s, and elongating at 72 ºC for 30 s. Further extension was carried out at 72 ºC for 5 min. This was followed by melting temperature measurements involving heating the reaction mixture to 95 °C for 30 s, cooling to 30 °C, then increasing the temperature to 95 °C at a rate of 1 or 0.5 °C/s and holding (monitored) at each temperature for 5 s. Samples were analysed by 1.5% or 2% agarose gel electrophoresis with ethidium bromide (200 ng/mL) at constant voltage (126 V) in 1 × TBE buffer.

When the same PCR amplifications were repeated 5 to 10 times, errors of melting temperatures were calculated using the confidence intervals derived from Student's *t* distribution with 95% confidence. In graphs these errors are represented by error bars.

Sequencing of template T11 PCR Products

PCR amplifications (4 × each) using AHP dUTP and T11 were carried out as previously described and purified by agarose gel electrophoresis. Products were extracted using Qiagen gel extraction kit according to the manufacturer's instructions. The solution of AHP-modified PCR amplicons were obtained and quantified using a NanoDrop 2000 UV spectrometer. Additionally, second-round PCR reactions were carried out using 10 nmol of unmodified triphosphates (dNTPs), 1 unit of Gotaq polymerase and 10 ng of the AHP-modified PCR amplicons as templates. The products were purified using agarose gel electrophoresis (1.5%), followed by extraction using Qiagen gel extraction kit. The

AHP-modified PCR amplicons and second-round PCR amplicons were then sent to Beckman Coulter Genomics Inc. for Sanger sequencing.

Fluorescent labelling of modified PCR products

Evaluation of reaction buffer system

AHP-modified PCR amplicons (4 ×) were purified by Qiagen agarose gel purification kit. Aqueous solution (50 µL) was obtained and topped up to 80 µL using deionised water. 10 µL of this solution was aliquoted into 6 individual vials and then labelled with Cy3-BCN (4 nmol) in Gotaq buffer or KOD + $MgCl_2$ buffer (final solution is 1 ×). Reactions were left at RT for 10, 30 and 60 min, then analysed by 1.5% agarose gel containing ethidium bromide.

Fluorescent labelling of PCR amplicons

For cyclooctyne-modified PCR amplicons (T8), the labelling reactions with Cy3-azide (16 nmol) were performed using the same protocol for labelling of primer extension products. For TCO- or AHP-modified or unmodified PCR amplicons (T11), products were extracted from agarose gel using Qiagen gel extraction kit and quantified using a NanoDrop 2000 UV spectrometer. 1.2 µg of each amplicons were diluted with Gotaq green buffer solution (60 µL, 1 ×) and labelled with 10 nmol of functionalised fluorophores at RT for 30 min. Samples were then analysed by 1.5% agarose gel electrophoresis at constant voltage (126 V) in 1 × TBE buffer.

Labelled AHP-modified products were extracted from the agarose gel using Qiagen gel extraction kit and aqueous solutions (50 µL) were obtained. The extracted solutions were diluted with deionised water (150 µL, for Cy3 and Cy5 labelling) or 100 mM TEAB buffer (150 µL, pH 8.0, for FAM labelling) and analysed by Perkin Elmer Luminescence

Spectrometer LS50B. Solutions were then lyophilised and analysed by agarose gel electrophoresis. Gels were imaged first and then stained with ethidium bromide solution (1 μg/mL) for 20 min to enable the unlabelled DNA to be visualised.

7.6 Synthesis of fluorescent single-stranded probes

λ-exonuclease digestion protocol

PCR was carried out with one unmodified primer and one phosphorylated primer using different ratios of AHP dUTP to dTTP. In the asymmetric PCR, the ratio of unmodified primer to phosphorylated-primer was 10:1. The PCR/APCR protocol was the same as described in the PCR protocol, except the PCR cycles were increased to 30. The reaction mixtures were washed with phenol/chloroform/isoamyl alcohol (25:24:1, saturated with Tris-HCl) and precipitated with ethanol to remove the DNA polymerase. The DNA precipitates were re-dissolved in 1 × λ-exonuclease buffer (20 μL) containing 5 units (T8 products) or 10 units (T11 products) of λ-exonuclease and SYBR green (0.6 μL, 3.75 ×). The reactions were heated at 37 °C for 2 h in a BIORAD CFX96 real-time PCR equipment and followed by melting temperature measurement to check the presence/absence of remaining duplex products. If there is un-digested duplexes remaining, an extra 5 units of λ-exonuclease was added and heated at 37 °C for 1 h. When the digestion was complete, the solutions (10 μL) were mixed with the complementary strand and melting temperature experiments were carried out. The single-stranded and re-annealed samples were analysed by 1.5% or 2% agarose gel with ethidium bromide (200 ng/mL) or SYBR gold (0.5 ×) at constant voltage (126 V) in 1 × TBE buffer.

Fluorescent labelling of ssDNA from λ-exonuclease digestion

AHP-modified and unmodified single-stranded products were prepared using above digestion method. Viva spin columns (3000, Sartorius Stedim Biotech) were used to remove extra triphosphates and other small molecules from the ssDNA solutions. Centrifugation (15, 000 × g) was repeated three times and the solution was quantified using a NanoDrop 2000 UV spectrometer. 1.46 μg of T8 single-stranded products or 2.54 μg of T11 single-stranded products were mixed with 1 × Gotaq green buffer (20 μL) with 10 nmol functionalised fluorophores and left at 55 °C for 1 h. When the reaction was complete, ethanol precipitation was used to remove excess free dye, then the precipitates were re-dissolved in 1 × Gotaq green buffer (10 μL). Half of the solution (5 μL) was mixed with the complementary strand and melting temperature experiments were carried out in a BIO-RAD CFX96 real-time PCR equipment. Samples were analysed by 1.5% or 2% agarose gel electrophoresis at constant voltage (126 V) in 1 × TBE buffer. Template T8 products were also analysed by 12% native PAGE at constant voltage (120 V) in 1 × TBE buffer for 15 h. Gels were imaged first, then stained with ethidium bromide (1 μg/mL) or SYBR Gold (1 ×) solution for 20 min to enable the unlabelled DNA to be visualised.

Fluorescent dual-labelling of ssDNA on streptavidin beads

PCR was carried out with one unmodified primer and one biotinylated primer using different ratios of AHP dUTP to dTTP. Dynabeads™ M-280 Streptavidin (10 mg/mL, Thermo Fisher Scientific) was washed three times with 2 × TE buffer (10 mM Tris-HCl, 1 mM EDTA, 2 M NaCl, pH 7.5). AHP-modified products from two PCR reactions (40 μL, T8 products) were mixed with Dynabeads (40 μL, 0.1 mg) in 2 × TE buffer and gently rotated at RT for 30 min. The supernatant was removed and the beads were washed three times with 2 × TE buffer. The beads were re-suspended in 20 mM NaOH (40 μL)

and rotated at RT for 30 min. The supernatant was collected in 40 mM HCl (30 µL) with 100 × TE buffer (1 µL). The beads were washed three times with 2 × TE buffer and re-suspended in 1 × Gotaq green buffer (10 µL) with 5 nmol of each BCN-fluorophore combination. The labelling reactions were heated at 55 °C for 1 h. When the reaction was complete, 70% ethanol was used to wash the beads three times. The beads were re-suspended in formamide (40 µL) and heated at 65 °C for 5 min in a heating block. The formamide solutions were analysed by 8% denaturing PAGE or collected and lyophilised. The residues were re-dissolved in 1 × Gotaq green buffer (15 µL). Half of the solution (7.5 µL) was mixed with the complementary strand (T8, 1.2 eq) and melting temperature experiments were carried out in a BIORAD CFX96 real-time PCR equipment. The labelled ssDNA and re-annealed samples were analysed by 12% native PAGE. Gels were imaged before staining at different channels according to the dye used (Table 7.2) and followed by staining with 1 × SYBR Gold.

Table 7.2. Gel imaging parameters for dual-labelling system

	Light	**Filter**
Stained gel	Transilluminator (302 nm)	EtBr/UV (572 – 625 nm)
FAM Channel	Epi-blue (455 – 487 nm)	Short wavelength (516 – 599 nm)
Cy3 (Rhodamine) Channel	Epi-green (520 – 550 nm)	Long wavelength (611 – 641 nm)
Cy5 Channel	Epi-red (615 – 650 nm)	FRLP (670 – 780 nm)

7.7 Reverse transcription

In a 20 µL reaction (1 × buffer), reverse transcriptase (M-MuLV (RNase H^-) 100 units, or AMV 5 units) and 10 × buffer were added to samples containing 66 pmol of primer (P3), 132 pmol of RNA template (T4 or T5), 3.2 nmol in total of modified dUTP/dTTP in addition to 3.2 nmol each of dATP, dCTP and dGTP. 1 × dithiothreitol (DTT, supplied

with the enzyme) was added to the M-MuLV reactions, while 20 units of RNase inhibitor were added to the AMV reactions. The reaction mixtures were left at 42 ºC for 1 or 15 h. When the reactions were complete, formamide (20 µL) was added and samples were analysed by 20% denaturing PAGE at a constant power of 20 W. For the mass spectrometry analysis, two same reactions were carried out, followed by ethanol precipitation. The precipitates were dissolved in deionised water (15 µL) and analysed by HPLC-MS. M-MuLV reverse transcriptase (RNase H^-), AMV reverse transcriptase and RNase inhibitor were purchased from New England Biolabs.

7.8 Cell biology

Cell cultures

A549 cells were cultivated at 37 °C/5% CO_2 in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum.

Metabolic labelling

Cells were seeded in 24-well plates containing glass coverslips (VWR; thickness 1.5, diameter 13 mm) with 5,000 cells per well and incubated for 24 h. The supernatant was removed, and fresh media containing variable concentrations of each nucleoside (diluted from 1,000 × stock solutions in DMSO) were added. After incubating for another 24 h, the cells were fixed in paraformaldehyde (3.7%) at RT for 15 min, and washed with 1 × phosphate buffered saline (PBS, containing 3% bovine serum albumin (BSA)) once, followed by addition of 0.5% Triton® X-100 in PBS (0.5 mL) to each well and incubation at RT for 20 min.

Cell labelling using CuAAC

Fixed cells were labelled with freshly prepared labelling mixture (200 µL, 10 µM AlexaFluor-594-alkyne or Cy3-azide, 1 mM $CuSO_4$, 2 mM THPTA, and 10 mM sodium ascorbate in 1 × PBS) at RT for 1 h in the dark.

Cell labelling using SPAAC

Fixed cells were labelled with freshly prepared labelling mixture (200 µL, 10 µM AlexaFluor-594-BCN, Cy3-BCN or Cy3-tetrazine in 1 × PBS) at RT for 1 h in the dark.

The labelling solution was then removed and the cells were washed with 1 × PBS (3% BSA) three times. The coverslips were fixed on slides using hard set™ mounting media with DAPI from VECTASHIELD and then visualised on a Zeiss confocal microscope (LSM780, 63x/1.4 Plan-Apochromat (oil) objective, laser 405 for DAPI, laser 543 for Cy3, laser 594 for AlexaFluor-594).

References

1. Großhans, H.; Filipowicz, W., Molecular biology: the expanding world of small RNAs. *Nature* **2008**, *451*, 414-416.

2. Watson, J. D.; Crick, F. H. C., Molecular structure of nucleic acids – a structure for deoxyribose nucleic acid. *Nature* **1953**, *171*, 737-738.

3. Chargaff, E.; Vischer, E.; Doniger, R.; Green, C.; Misani, F., The composition of the desoxypentose nucleic acids of thymus and spleen. *J. Biol. Chem.* **1949**, *177*, 405-416.

4. Chargaff, E.; Lipshitz, R.; Green, C., Composition of the desoxypentose nucleic acids of four genera of sea-urchin. *J. Biol. Chem.* **1952**, *195*, 155-160.

5. Hamilton, P. L.; Arya, D. P., Natural product DNA major groove binders. *Nat. Prod. Rep.* **2012**, *29*, 134-143.

6. Simonsson, S.; Samuelsson, T.; Elias, P., The herpes simplex virus type 1 origin binding protein: specific recognition of phosphates and methyl groups defines the interacting surface for a monomeric DNA binding domain in the major groove of DNA. *J. Biol. Chem.* **1998**, *273*, 24633-24639.

7. Baraldi, P. G.; Bovero, A.; Fruttarolo, F.; Preti, D.; Tabrizi, M. A.; Pavani, M. G.; Romagnoli, R., DNA minor groove binders as potential antitumor and antimicrobial agents. *Med. Res. Rev.* **2004**, *24*, 475-528.

8. Bewley, C. A.; Gronenborn, A. M.; Clore, G. M., Minor groove-binding architectural proteins: structure, function, and DNA recognition. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 105-131.

9. Gao, Y. G.; Robinson, H.; Wang, A. H., High-resolution A-DNA crystal structures of d(AGGGGCCCCT) – an A-DNA model of poly(dG)•poly(dC). *Eur. J. Biochem.* **1999**, *261*, 413-420.

10. Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E., Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 2179-2183.

11. Dunn, D. B., The isolation of 5-methylcytidine from RNA. *Biochim. Biophys. Acta* **1960**, *38*, 176-178.

12. Carell, T.; Brandmayr, C.; Hienzsch, A.; Müller, M.; Pearson, D.; Reiter, V.; Thoma, I.; Thumbs, P.; Wagner, M., Structure and function of noncanonical nucleobases. *Angew. Chem. Int. Ed.* **2012**, *51*, 7110-7131.

13. Christman, J. K.; Sheikhnejad, G.; Marasco, C. J.; Sufrin, J. R., 5-Methyl-2′-deoxycytidine in single-stranded DNA can act in cis to signal *de novo* DNA methylation. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 7347-7351.

14. Ichikawa, E.; Kato, K., Sugar-modified nucleosides in past 10 years, a review. *Curr. Med. Chem.* **2001**, *8*, 385-423.

15. Périgaud, C.; Gosselin, G.; Imbach, J. L., Nucleoside analogues as chemotherapeutic agents: a review. *Nucleos. Nucleot. Nucl.* **1992**, *11*, 903-945.

16. Schabel, F. M., The antiviral activity of 9-*β*-D-arabinofuranosyladenine (ara-A). *Chemotherapy* **1968**, *13*, 321-338.

17. Loakes, D., The applications of universal DNA base analogues. *Nucleic Acids Res.* **2001**, *29*, 2437-2447.

18. Liang, F.; Liu, Y.; Zhang, P., Universal base analogues and their applications in DNA sequencing technology. *RSC Adv.* **2013**, *3*, 14910-14928.

19. Bergstrom, D. E.; Zhang, P.; Toma, P. H.; Andrews, P. C.; Nichols, R., Synthesis, structure, and deoxyribonucleic acid sequencing with a universal nucleoside: 1-(2′-deoxy-*β*-D-ribofuranosyl)-3-nitropyrrole. *J. Am. Chem. Soc.* **1995**, *117*, 1201-1209.

20. Nielsen, P.; Egholm, M.; Berg, R.; Buchardt, O., Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **1991**, *254*, 1497-1500.

21. Nielsen, P. E., Peptide nucleic acid: a versatile tool in genetic diagnostics and molecular biology. *Curr. Opin. Biotechnol.* **2001**, *12*, 16-20.

22. Zipper, H.; Brunner, H.; Bernhagen, J.; Vitzthum, F., Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications. *Nucleic Acids Res.* **2004**, *32*, e103.

23. Giglio, S.; Monis, P. T.; Saint, C. P., Demonstration of preferential binding of SYBR Green I to specific DNA fragments in real-time multiplex PCR. *Nucleic Acids Res.* **2003**, *31*, e136.

24. Pattyn, F.; Speleman, F.; De Paepe, A.; Vandesompele, J., RTPrimerDB: the real-time PCR primer and probe database. *Nucleic Acids Res.* **2003**, *31*, 122-123.

25. Chehab, F. F.; Kan, Y. W., Detection of specific DNA sequences by fluorescence amplification: a color complementation assay. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 9178-9182.

26. Frégeau, C. J.; Fourney, R. M., DNA typing with fluorescently tagged short tandem repeats – a sensitive and accurate approach to human identification. *Biotechniques* **1993**, *15*, 100-119.

27. Wagner, M.; Horn, M.; Daims, H., Fluorescence *in situ* hybridisation for the identification and characterisation of prokaryotes. *Curr. Opin. Microbiol.* **2003**, *6*, 302-309.

28. Karger, A. E.; Harris, J. M.; Gesteland, R. F., Multiwavelength fluorescence detection for DNA sequencing using capillary electrophoresis. *Nucleic Acids Res.* **1991**, *19*, 4955-4962.

29. Smith, L. M.; Sanders, J. Z.; Kaiser, R. J.; Hughes, P.; Dodd, C.; Connell, C. R.; Heiner, C.; Kent, S. B. H.; Hood, L. E., Fluorescence detection in automated DNA sequence analysis. *Nature* **1986**, *321*, 674-679.

30. Sobrino, B.; Brión, M.; Carracedo, A., SNPs in forensic genetics: a review on SNP typing methodologies. *Forensic Sci. Int.* **2005**, *154*, 181-194.

31. Didenko, V. V., DNA probes using fluorescence resonance energy transfer (FRET): designs and applications. *Biotechniques* **2001**, *31*, 1106-1121.

32. Phillips, R. B.; Reed, K. M., Application of fluorescence *in situ* hybridization (FISH) techniques to fish genetics: a review. *Aquaculture* **1996**, *140*, 197-216.

33. Trask, B. J., Fluorescence *in situ* hybridization: applications in cytogenetics and gene mapping. *Trends Genet.* **1991**, *7*, 149-154.

34. Liehr, T.; Starke, H.; Weise, A.; Lehrer, H.; Claussen, U., Multicolor FISH probe sets and their applications. *Histol. Histopathol.* **2004**, *19*, 229-237.

35. Arndtjovin, D. J.; Jovin, T. M., Fluorescence labeling and microscopy of DNA. *Methods Cell Biol.* **1989**, *30*, 417-448.

36. Meyer, P.; Dworkin, J., Applications of fluorescence microscopy to single bacterial cells. *Res. Microbiol.* **2007**, *158*, 187-194.

37. Bick, M. D.; Davidson, R. L., Total substitution of bromodeoxyuridine for thymidine in DNA of a bromodeoxyuridine-dependent cell line. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 2082-2086.

38. Salic, A.; Mitchison, T. J., A chemical method for fast and sensitive detection of DNA synthesis *in vivo*. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 2415-2420.

39. Davies, M. J.; Shah, A.; Bruce, I. J., Synthesis of fluorescently labelled oligonucleotides and nucleic acids. *Chem. Soc. Rev.* **2000**, *29*, 97-107.

40. Caruthers, M. H., Chemical synthesis of DNA and DNA analogues. *Acc. Chem. Res.* **1991**, *24*, 278-284.

41. Beaucage, S. L.; Caruthers, M. H., Deoxynucleoside phosphoramidites – a new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett.* **1981**, *22*, 1859-1862.

42. Reese, C. B., Oligo- and poly-nucleotides: 50 years of chemical synthesis. *Org. Biomol. Chem.* **2005**, *3*, 3851-3868.

43. French, D. J.; Archard, C. L.; Brown, T.; McDowell, D. G., HyBeacon probes: a new tool for DNA sequence detection and allele discrimination. *Mol. Cell. Probes* **2001**, *15*, 363-374.

44. Chen, X.; Sullivan, P. F., Single nucleotide polymorphism genotyping: biochemistry, protocol, cost and throughput. *Pharmacogenomics J.* **2003**, *3*, 77-96.

45. Hall, L. M.; Gerowska, M.; Brown, T., A highly fluorescent DNA toolkit: synthesis and properties of oligonucleotides containing new Cy3, Cy5 and Cy3B monomers. *Nucleic Acids Res.* **2012**, *40*, e108.

46. Sau, S. P.; Hrdlicka, P. J., C2′-Pyrene-functionalized triazole-linked DNA: universal DNA/RNA hybridization probes. *J. Org. Chem.* **2012**, *77*, 5-16.

47. Xiao, Q.; Ranasinghe, R. T.; Tang, A. M. P.; Brown, T., Naphthalenyl- and anthracenyl-ethynyl dT analogues as base discriminating fluorescent nucleosides and intramolecular energy transfer donors in oligonucleotide probes. *Tetrahedron* **2007**, *63*, 3483-3490.

48. Dierckx, A.; Miannay, F. A.; Gaied, N.; Preus, S.; Björck, M.; Brown, T.; Wilhelmsson, L. M., Quadracyclic adenine: a non-perturbing fluorescent adenine analogue. *Chem. Eur. J.* **2012**, *18*, 5987-5997.

49. Ren, R. X. F.; Chaudhuri, N. C.; Paris, P. L.; Rumney, S.; Kool, E. T., Naphthalene, phenanthrene, and pyrene as DNA base analogues: synthesis, structure, and fluorescence in DNA. *J. Am. Chem. Soc.* **1996**, *118*, 7671-7678.

50. Wilhelmsson, L. M., Fluorescent nucleic acid base analogues. *Q. Rev. Biophys.* **2010**, *43*, 159-183.

51. Preus, S.; Börjesson, K.; Kilså, K.; Albinsson, B.; Wilhelmsson, L. M., Characterization of nucleobase analogue FRET acceptor tC^{nitro}. *J. Phys. Chem. B* **2010**, *114*, 1050-1056.

52. Jaffe, E. K.; Volin, M.; Bronson-Mullins, C. R.; Dunbrack, R. L.; Kervinen, J.; Martins, J.; Quinlan, J. F.; Sazinsky, M. H.; Steinhouse, E. M.; Yeung, A. T., An artificial gene for human porphobilinogen synthase allows comparison of an allelic variation implicated in susceptibility to lead poisoning. *J. Biol. Chem.* **2000**, *275*, 2619-2626.

53. Gramlich, P. M. E.; Wirges, C. T.; Manetto, A.; Carell, T., Postsynthetic DNA modification through the copper-catalyzed azide-alkyne cycloaddition reaction. *Angew. Chem. Int. Ed.* **2008**, *47*, 8350-8358.

54. Steitz, T. A., DNA polymerases: structural diversity and common mechanisms. *J. Biol. Chem.* **1999**, *274*, 17395-17398.

55. Steitz, T. A., DNA- and RNA-dependent DNA polymerases. *Curr. Opin. Struct. Biol.* **1993**, *3*, 31-38.

56. Brautigam, C. A.; Steitz, T. A., Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Curr. Opin. Struct. Biol.* **1998**, *8*, 54-63.

57. Joyce, C. M.; Steitz, T. A., Function and structure relationships in DNA polymerases. *Annu. Rev. Biochem* **1994**, *63*, 777-822.

58. Bergen, K.; Betz, K.; Welte, W.; Diederichs, K.; Marx, A., Structures of KOD and 9°N DNA polymerases complexed with primer template duplex. *ChemBioChem* **2013**, *14*, 1058-1062.

59. Jäger, S.; Rasched, G.; Kornreich-Leshem, H.; Engeser, M.; Thum, O.; Famulok, M., A versatile toolbox for variable DNA functionalization at high density. *J. Am. Chem. Soc.* **2005**, *127*, 15071-15082.

60. Kranaster, R.; Marx, A., Engineered DNA polymerases in biotechnology. *ChemBioChem* **2010**, *11*, 2077-2084.

61. Yu, H.; Chao, J.; Patek, D.; Mujumdar, R.; Mujumdar, S.; Waggoner, A. S., Cyanine dye dUTP analogs for enzymatic labeling of DNA probes. *Nucleic Acids Res.* **1994**, *22*, 3226-3232.

62. Bartlett, J. M. S.; Stirling, D., A short history of the polymerase chain reaction. *Methods Mol. Biol.* **2003**, *226*, 3-6.

63. Sanchez, J. A.; Pierce, K. E.; Rice, J. E.; Wangh, L. J., Linear-after-the-exponential (LATE)-PCR: an advanced method of asymmetric PCR and its uses in quantitative real-time analysis. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 1933-1938.

64. Thoresen, L. H.; Jiao, G. S.; Haaland, W. C.; Metzker, M. L.; Burgess, K., Rigid, conjugated, fluoresceinated thymidine triphosphates: syntheses and polymerase mediated incorporation into DNA analogues. *Chem. Eur. J.* **2003**, *9*, 4603-4610.

65. Seela, F.; Feiling, E.; Gross, J.; Hillenkamp, F.; Ramzaeva, N.; Rosemeyer, H.; Zulauf, M., Fluorescent DNA: the development of 7-deazapurine nucleoside triphosphates applicable for sequencing at the single molecule level. *J. Biotechnol.* **2001**, *86*, 269-279.

66. Giller, G.; Tasara, T.; Angerer, B.; Mühlegger, K.; Amacker, M.; Winter, H., Incorporation of reporter molecule-labeled nucleotides by DNA polymerases. I. Chemical synthesis of various reporter group-labeled 2′-deoxyribonucleoside-5′-triphosphates. *Nucleic Acids Res.* **2003**, *31*, 2630-2635.

67. Tasara, T.; Angerer, B.; Damond, M.; Winter, H.; Dörhöfer, S.; Hübscher, U.; Amacker, M., Incorporation of reporter molecule-labeled nucleotides by DNA polymerases. II. High-density labeling of natural DNA. *Nucleic Acids Res.* **2003**, *31*, 2636-2646.

68. Merkel, M.; Peewasan, K.; Arndt, S.; Ploschik, D.; Wagenknecht, H. A., Copper-free postsynthetic labeling of nucleic acids by means of bioorthogonal reactions. *ChemBioChem* **2015**, *16*, 1541-1553.

69. El-Sagheer, A. H.; Brown, T., Click chemistry with DNA. *Chem. Soc. Rev.* **2010**, *39*, 1388-1405.

70. Nampalli, S.; McDougall, M. G.; Lavrenov, K.; Xiao, H. G.; Kumar, S., Utility of thiol-cross-linked fluorescent dye labeled terminators for DNA sequencing. *Bioconjugate Chem.* **2002**, *13*, 468-473.

71. ’t Hoen, P. A. C.; de Kort, F.; van Ommen, G. J. B.; den Dunnen, J. T., Fluorescent labelling of cRNA for microarray applications. *Nucleic Acids Res.* **2003**, *31*, e20.

72. Smith, L. M.; Fung, S.; Hunkapiller, M. W.; Hunkapiller, T. J.; Hood, L. E., The synthesis of oligonucleotides containing an aliphatic amino group at the 5′ terminus: synthesis of fluorescent DNA primers for use in DNA sequence analysis. *Nucleic Acids Res.* **1985**, *13*, 2399-2412.

73. Cox, W. G.; Singer, V. L., Fluorescent DNA hybridization probe preparation using amine modification and reactive dye coupling. *Biotechniques* **2004**, *36*, 114-122.

74. Kolb, H. C.; Finn, M. G.; Sharpless, K. B., Click chemistry: diverse chemical function from a few good reactions. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004-2021.

75. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B., A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective ‘ligation’ of azides and terminal alkynes. *Angew. Chem. Int. Ed.* **2002**, *41*, 2596-2599.

76. Moses, J. E.; Moorhouse, A. D., The growing applications of click chemistry. *Chem. Soc. Rev.* **2007**, *36*, 1249-1262.

77. Kolb, H. C.; Sharpless, K. B., The growing impact of click chemistry on drug discovery. *Drug Discov. Today* **2003**, *8*, 1128-1137.

78. Thirumurugan, P.; Matosiuk, D.; Jozwiak, K., Click chemistry for drug development and diverse chemical-biology applications. *Chem. Rev.* **2013**, *113*, 4905-4979.

79. Liang, L.; Astruc, D., The copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) ‘click’ reaction and its applications. An overview. *Coord. Chem. Rev.* **2011**, *255*, 2933-2945.

80. Amblard, F.; Cho, J. H.; Schinazi, R. F., Cu(I)-catalyzed Huisgen azide-alkyne 1,3-dipolar cycloaddition reaction in nucleoside, nucleotide, and oligonucleotide chemistry. *Chem. Rev.* **2009**, *109*, 4207-4220.

81. Golas, P. L.; Matyjaszewski, K., Marrying click chemistry with polymerization: expanding the scope of polymeric materials. *Chem. Soc. Rev.* **2010**, *39*, 1338-1354.

82. Zhang, L.; Chen, X. G.; Xue, P.; Sun, H. H. Y.; Williams, I. D.; Sharpless, K. B.; Fokin, V. V.; Jia, G. C., Ruthenium-catalyzed cycloaddition of alkynes and organic azides. *J. Am. Chem. Soc.* **2005**, *127*, 15998-15999.

83. Boren, B. C.; Narayan, S.; Rasmussen, L. K.; Zhang, L.; Zhao, H.; Lin, Z.; Jia, G.; Fokin, V. V., Ruthenium-catalyzed azide-alkyne cycloaddition: scope and mechanism. *J. Am. Chem. Soc.* **2008**, *130*, 8923-8930.

84. Montagu, A.; Roy, V.; Balzarini, J.; Snoeck, R.; Andrei, G.; Agrofoglio, L. A., Synthesis of new C5-(1-substituted-1,2,3-triazol-4 or 5-yl)-2′-deoxyuridines and their antiviral evaluation. *Eur. J. Med. Chem.* **2011**, *46*, 778-786.

85. Chemama, M.; Fonvielle, M.; Arthur, M.; Valéry, J. M.; Etheve-Quelquejeu, M., Synthesis of stable aminoacyl-tRNA analogues containing triazole as a bioisoster of esters. *Chem. Eur. J.* **2009**, *15*, 1929-1938.

86. Worrell, B. T.; Malik, J. A.; Fokin, V. V., Direct evidence of a dinuclear copper intermediate in Cu(I)-catalyzed azide-alkyne cycloadditions. *Science* **2013**, *340*, 457-460.

87. Ryazantsev, D. Y.; Tsybulsky, D. A.; Prokhorenko, I. A.; Kvach, M. V.; Martynenko, Y. V.; Philipchenko, P. M.; Shmanai, V. V.; Korshun, V. A.; Zavriev, S. K., Two-dye and one- or two-quencher DNA probes for real-time PCR assay: synthesis and comparison with a TaqMan™ probe. *Anal. Bioanal.Chem.* **2012**, *404*, 59-68.

88. Seela, F.; Sirivolu, V. R., Nucleosides and oligonucleotides with diynyl side chains: base pairing and functionalization of 2′-deoxyuridine derivatives by the copper(I)-catalyzed alkyne-azide ‘click’ cycloaddition. *Helv. Chim. Acta* **2007**, *90*, 535-552.

89. Gierlich, J.; Burley, G. A.; Gramlich, P. M. E.; Hammond, D. M.; Carell, T., Click chemistry as a reliable method for the high-density postsynthetic functionalization of alkyne-modified DNA. *Org. Lett.* **2006**, *8*, 3639-3642.

90. Kumar, P.; Østergaard, M. E.; Baral, B.; Anderson, B. A.; Guenther, D. C.; Kaura, M.; Raible, D. J.; Sharma, P. K.; Hrdlicka, P. J., Synthesis and biophysical properties of C5-functionalized LNA (locked nucleic acid). *J. Org. Chem.* **2014**, *79*, 5047-5061.

91. Gramlich, P. M. E.; Warncke, S.; Gierlich, J.; Carell, T., Click-click-click: single to triple modification of DNA. *Angew. Chem. Int. Ed.* **2008**, *47*, 3442-3444.

92. Seela, F.; Sirivolu, V. R.; Chittepu, P., Modification of DNA with octadiynyl side chains: synthesis, base pairing, and formation of fluorescent coumarin dye conjugates of four nucleobases by the alkyne-azide ‘click’ reaction. *Bioconjugate Chem.* **2008**, *19*, 211-224.

93. Sirivolu, V. R.; Chittepu, P.; Seela, F., DNA with branched internal side chains: synthesis of 5-tripropargylamine-dU and conjugation by an azide-alkyne double click reaction. *ChemBioChem* **2008**, *9*, 2305-2316.

94. Ustinov, A. V.; Korshun, V. A., Oligonucleotides containing arylacetylene residues: synthesis and post-synthetic modification *via* [3+2] cycloaddition. *Russ. Chem. Bull.* **2006**, *55*, 1268-1274.

95. El-Sagheer, A. H.; Sanzone, A. P.; Gao, R.; Tavassoli, A.; Brown, T., Biocompatible artificial DNA linker that is read through by DNA polymerases and is functional in Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 11338-11343.

96. Berndl, S.; Herzig, N.; Kele, P.; Lachmann, D.; Li, X.; Wolfbeis, O. S.; Wagenknecht, H. A., Comparison of a nucleosidic vs non-nucleosidic postsynthetic 'click' modification of DNA with base-labile fluorescent probes. *Bioconjugate Chem.* **2009**, *20*, 558-564.

97. Guan, L.; van der Heijden, G. W.; Bortvin, A.; Greenberg, M. M., Intracellular detection of cytosine incorporation in genomic DNA by using 5-ethynyl-2′-deoxycytidine. *ChemBioChem* **2011**, *12*, 2184-2190.

98. Motea, E. A.; Lee, I.; Berdis, A. J., Development of a 'clickable' non-natural nucleotide to visualize the replication of non-instructional DNA lesions. *Nucleic Acids Res.* **2012**, *40*, 2357-2367.

99. Wenge, U.; Ehrenschwender, T.; Wagenknecht, H.-A., Synthesis of 2′-*O*-propargyl nucleoside triphosphates for enzymatic oligonucleotide preparation and 'click' modification of DNA with nile red as fluorescent probe. *Bioconjugate Chem.* **2013**, *24*, 301-304.

100. Gramlich, P. M. E.; Wirges, C. T.; Gierlich, J.; Carell, T., Synthesis of modified DNA by PCR with alkyne-bearing purines followed by a click reaction. *Org. Lett.* **2008**, *10*, 249-251.

101. Gierlich, J.; Gutsmiedl, K.; Gramlich, P. M. E.; Schmidt, A.; Burley, G. A.; Carell, T., Synthesis of highly modified DNA by a combination of PCR with alkyne-bearing triphosphates and click chemistry. *Chem. Eur. J.* **2007**, *13*, 9486-9494.

102. Gutsmiedl, K.; Fazio, D.; Carell, T., High-density DNA functionalization by a combination of Cu-catalyzed and Cu-free click chemistry. *Chem. Eur. J.* **2010**, *16*, 6877-6883.

103. Ishizuka, T.; Kimoto, M.; Sato, A.; Hirao, I., Site-specific functionalization of RNA molecules by an unnatural base pair transcription system *via* click chemistry. *Chem. Commun.* **2012**, *48*, 10835-10837.

104. Burley, G. A.; Gierlich, J.; Mofid, M. R.; Nir, H.; Tal, S.; Eichen, Y.; Carell, T., Directed DNA metallization. *J. Am. Chem. Soc.* **2006**, *128*, 1398-1399.

105. Motea, E. A.; Lee, I.; Berdis, A. J., A non-natural nucleoside with combined therapeutic and diagnostic activities against leukemia. *ACS Chem. Biol.* **2012**, *7*, 988-998.

106. Obeid, S.; Bußkamp, H.; Welte, W.; Diederichs, K.; Marx, A., Interactions of non-polar and 'click-able' nucleotides in the confines of a DNA polymerase active site. *Chem. Commun.* **2012**, *48*, 8320-8322.

107. Meyer, H. S. J., Über neue organische phosphorverbindungen III. Phosphinmethylenderivate und phosphinimine. *Helv. Chim. Acta* **1919**, *2*, 635-646.

108. Aigner, M.; Hartl, M.; Fauster, K.; Steger, J.; Bister, K.; Micura, R., Chemical synthesis of site-specifically 2′-azido-modified RNA and potential applications for bioconjugation and RNA interference. *ChemBioChem* **2011**, *12*, 47-51.

109. Kiviniemi, A.; Virta, P.; Lönnberg, H., Utilization of intrachain 4′-*C*-azidomethylthymidine for preparation of oligodeoxyribonucleotide conjugates by click chemistry in solution and on a solid support. *Bioconjugate Chem.* **2008**, *19*, 1726-1734.

110. Rao, H.; Sawant, A. A.; Tanpure, A. A.; Srivatsan, S. G., Posttranscriptional chemical functionalization of azide-modified oligoribonucleotides by bioorthogonal click and Staudinger reactions. *Chem. Commun.* **2012**, *48*, 498-500.

111. Weisbrod, S. H.; Marx, A., A nucleoside triphosphate for site-specific labelling of DNA by the Staudinger ligation. *Chem. Commun.* **2007**, 1828-1830.

112. Winz, M. L.; Samanta, A.; Benzinger, D.; Jäschke, A., Site-specific terminal and internal labeling of RNA by poly(A) polymerase tailing and copper-catalyzed or copper-free strain-promoted click chemistry. *Nucleic Acids Res.* **2012**, *40*, e78.

113. Sørensen, R. S.; Okholm, A. H.; Schaffert, D.; Kodal, A. L. B.; Gothelf, K. V.; Kjems, J., Enzymatic ligation of large biomolecules to DNA. *ACS Nano* **2013**, *7*, 8098-8104.

114. Winz, M. L.; Linder, E. C.; André, T.; Becker, J.; Jäschke, A., Nucleotidyl transferase assisted DNA labeling with different click chemistries. *Nucleic Acids Res.* **2015**, *43*, e110.

115. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R., A strain-promoted [3+2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc.* **2004**, *126*, 15046-15047.

116. Chang, P. V.; Prescher, J. A.; Sletten, E. M.; Baskin, J. M.; Miller, I. A.; Agard, N. J.; Lo, A.; Bertozzi, C. R., Copper-free click chemistry in living animals. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 1821-1826.

117. Jewett, J. C.; Sletten, E. M.; Bertozzi, C. R., Rapid Cu-free click chemistry with readily synthesized biarylazacyclooctynones. *J. Am. Chem. Soc.* **2010**, *132*, 3688-3690.

118. Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R., Copper-free click chemistry for dynamic *in vivo* imaging. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 16793-16797.

119. Debets, M. F.; van Berkel, S. S.; Dommerholt, J.; Dirks, A. J.; Rutjes, F. P. J. T.; van Delft, F. L., Bioconjugation with strained alkenes and alkynes. *Acc. Chem. Res.* **2011**, *44*, 805-815.

120. Lang, K.; Davis, L.; Wallace, S.; Mahesh, M.; Cox, D. J.; Blackman, M. L.; Fox, J. M.; Chin, J. W., Genetic encoding of bicyclononynes and *trans*-cyclooctenes for site-specific protein labeling *in vitro* and in live mammalian cells *via* rapid fluorogenic Diels-Alder reactions. *J. Am. Chem. Soc.* **2012**, *134*, 10317-10320.

121. Ning, X.; Guo, J.; Wolfert, M. A.; Boons, G. J., Visualizing metabolically labeled glycoconjugates of living cells by copper-free and fast Huisgen cycloadditions. *Angew. Chem. Int. Ed.* **2008**, *47*, 2253-2255.

122. Dommerholt, J.; Schmidt, S.; Temming, R.; Hendriks, L. J. A.; Rutjes, F.; van Hest, J. C. M.; Lefeber, D. J.; Friedl, P.; van Delft, F. L., Readily accessible bicyclononynes for bioorthogonal labeling and three-dimensional imaging of living cells. *Angew. Chem. Int. Ed.* **2010**, *49*, 9422-9425.

123. Jawalekar, A. M.; Malik, S.; Verkade, J. M. M.; Gibson, B.; Barta, N. S.; Hodges, J. C.; Rowan, A.; van Delft, F. L., Oligonucleotide tagging for copper-free click conjugation. *Molecules* **2013**, *18*, 7346-7363.

124. van Delft, P.; Meeuwenoord, N. J.; Hoogendoorn, S.; Dinkelaar, J.; Overkleeft, H. S.; van der Marel, G. A.; Filippov, D. V., Synthesis of oligoribonucleic acid conjugates using a cyclooctyne phosphoramidite. *Org. Lett.* **2010**, *12*, 5486-5489.

125. Jayaprakash, K. N.; Peng, C. G.; Butler, D.; Varghese, J. P.; Maier, M. A.; Rajeev, K. G.; Manoharan, M., Non-nucleoside building blocks for copper-assisted and copper-free click chemistry for the efficient synthesis of RNA conjugates. *Org. Lett.* **2010**, *12*, 5410-5413.

126. Marks, I. S.; Kang, J. S.; Jones, B. T.; Landmark, K. J.; Cleland, A. J.; Taton, T. A., Strain-promoted 'click' chemistry for terminal labeling of DNA. *Bioconjugate Chem.* **2011**, *22*, 1259-1263.

127. van Delft, P.; van Schie, E.; Meeuwenoord, N. J.; Overkleeft, H. S.; van der Marel, G. A.; Filippov, D. V., Oligonucleotide conjugates by means of copper-free click chemistry – expanding the repertoire of strained cyclooctyne phosphoramidites. *Synthesis Stuttgart* **2011**, 2724-2732.

128. Shelbourne, M.; Chen, X.; Brown, T.; El-Sagheer, A. H., Fast copper-free click DNA ligation by the ring-strain promoted alkyne-azide cycloaddition reaction. *Chem. Commun.* **2011**, *47*, 6257-6259.

129. Shelbourne, M.; Brown, T., Jr.; El-Sagheer, A. H.; Brown, T., Fast and efficient DNA crosslinking and multiple orthogonal labelling by copper-free click chemistry. *Chem. Commun.* **2012**, *48*, 11184-11186.

130. van Geel, R.; Pruijn, G. J. M.; van Delft, F. L.; Boelens, W. C., Preventing thiol-yne addition improves the specificity of strain-promoted azide-alkyne cycloaddition. *Bioconjugate Chem.* **2012**, *23*, 392-398.

131. Blackman, M. L.; Royzen, M.; Fox, J. M., Tetrazine ligation: fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity. *J. Am. Chem. Soc.* **2008**, *130*, 13518-13519.

132. Šečkutė, J.; Yang, J.; Devaraj, N. K., Rapid oligonucleotide-templated fluorogenic tetrazine ligations. *Nucleic Acids Res.* **2013**, *41*, e148.

133. Schoch, J.; Wiessler, M.; Jäschke, A., Post-synthetic modification of DNA by inverse-electron-demand Diels-Alder reaction. *J. Am. Chem. Soc.* **2010**, *132*, 8846-8847.

134. Asare-Okai, P. N.; Agustin, E.; Fabris, D.; Royzen, M., Site-specific fluorescence labelling of RNA using bio-orthogonal reaction of *trans*-cyclooctene and tetrazine. *Chem. Commun.* **2014**, *50*, 7844-7847.

135. Knall, A. C.; Slugovc, C., Inverse electron demand Diels-Alder (iEDDA)-initiated conjugation: a (high) potential click chemistry scheme. *Chem. Soc. Rev.* **2013**, *42*, 5131-5142.

136. Karver, M. R.; Weissleder, R.; Hilderbrand, S. A., Synthesis and evaluation of a series of 1,2,4,5-tetrazines for bioorthogonal conjugation. *Bioconjugate Chem.* **2011**, *22*, 2263-2270.

137. Schoch, J.; Ameta, S.; Jäschke, A., Inverse electron-demand Diels-Alder reactions for the selective and efficient labeling of RNA. *Chem. Commun.* **2011**, *47*, 12536-12537.

138. Schoch, J.; Staudt, M.; Samanta, A.; Wiessler, M.; Jäschke, A., Site-specific one-pot dual labeling of DNA by orthogonal cycloaddition chemistry. *Bioconjugate Chem.* **2012**, *23*, 1382-1386.

139. Wang, K.; Wang, D.; Ji, K.; Chen, W.; Zheng, Y.; Dai, C.; Wang, B., Post-synthesis DNA modifications using a *trans*-cyclooctene click handle. *Org. Biomol. Chem.* **2015**, *13*, 909-915.

140. Schoch, J.; Jäschke, A., Synthesis and enzymatic incorporation of norbornene-modified nucleoside triphosphates for Diels-Alder bioconjugation. *RSC Adv.* **2013**, *3*, 4181-4183.

141. Ameta, S.; Becker, J.; Jäschke, A., RNA-peptide conjugate synthesis by inverse-electron demand Diels-Alder reaction. *Org. Biomol. Chem.* **2014**, *12*, 4701-4707.

142. Liang, Y.; Mackey, J. L.; Lopez, S. A.; Liu, F.; Houk, K. N., Control and design of mutual orthogonality in bioorthogonal cycloadditions. *J. Am. Chem. Soc.* **2012**, *134*, 17904-17907.

143. Karver, M. R.; Weissleder, R.; Hilderbrand, S. A., Bioorthogonal reaction pairs enable simultaneous, selective, multi-target imaging. *Angew. Chem. Int. Ed.* **2012**, *51*, 920-922.

144. Sletten, E. M.; Bertozzi, C. R., Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew. Chem. Int. Ed.* **2009**, *48*, 6974-6998.

145. Prescher, J. A.; Bertozzi, C. R., Chemistry in living systems. *Nat. Chem. Biol.* **2005**, *1*, 13-21.

146. Boutorine, A. S.; Novopashina, D. S.; Krasheninina, O. A.; Nozeret, K.; Venyaminova, A. G., Fluorescent probes for nucleic acid visualization in fixed and live cells. *Molecules* **2013**, *18*, 15357-15397.

147. Balzarini, J.; Herdewijn, P.; De Clercq, E., Differential patterns of intracellular metabolism of 2′,3′-didehydro-2′,3′-dideoxythymidine and 3′-azido-2′,3′-dideoxythymidine, two potent anti-human immunodeficiency virus compounds. *J. Biol. Chem.* **1989**, *264*, 6127-6133.

148. Hitchcock, M. J. M., 2′,3′-Didehydro-2′,3′-dideoxythymidine d4T an anti-HIV agent. *Antiviral Chem. Chemother.* **1991**, *2*, 125-132.

149. Riddler, S. A.; Anderson, R. E.; Mellors, J. W., Antiretroviral activity of stavudine (2′,3′-didehydro-3′-deoxythymidine, d4T). *Antiviral Res.* **1995**, *27*, 189-203.

150. Mehellou, Y.; Balzarini, J.; McGuigan, C., Aryloxy phosphoramidate triesters: a technology for delivering monophosphorylated nucleosides and sugars into cells. *ChemMedChem* **2009**, *4*, 1779-1791.

151. Meier, C., CycloSal phosphates as chemical trojan horses for intracellular nucleotide and glycosylmonophosphate delivery – chemistry meets biology. *Eur. J. Org. Chem.* **2006**, 1081-1102.

152. Gratzner, H. G., Monoclonal antibody to 5-bromo- and 5-iododeoxyuridine: a new reagent for detection of DNA replication. *Science* **1982**, *218*, 474-475.

153. Jao, C. Y.; Salic, A., Exploring RNA transcription and turnover *in vivo* by using click chemistry. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 15779-15784.

154. Zhao, H.; Halicka, H. D.; Li, J.; Biela, E.; Berniak, K.; Dobrucki, J.; Darzynkiewicz, Z., DNA damage signaling, impairment of cell cycle progression, and apoptosis triggered by 5-ethynyl-2′-deoxyuridine incorporated into DNA. *Cytometry Part A* **2013**, *83*, 979-988.

155. Neef, A. B.; Luedtke, N. W., Dynamic metabolic labeling of DNA *in vivo* with arabinosyl nucleosides. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 20404-20409.

156. Rieder, U.; Luedtke, N. W., Alkene-tetrazine ligation for imaging cellular DNA. *Angew. Chem. Int. Ed.* **2014**, *53*, 9168-9172.

157. Neef, A. B.; Luedtke, N. W., An azide-modified nucleoside for metabolic labeling of DNA. *ChemBioChem* **2014**, *15*, 789-793.

158. Hollenstein, M., Nucleoside triphosphates – building blocks for the modification of nucleic acids. *Molecules* **2012**, *17*, 13569-13591.

159. Čapek, P.; Cahová, H.; Pohl, R.; Hocek, M.; Gloeckner, C.; Marx, A., An efficient method for the construction of functionalized DNA bearing amino acid groups through cross-coupling reactions of nucleoside triphosphates followed by primer extension or PCR. *Chem. Eur. J.* **2007**, *13*, 6196-6203.

160. Sato, K.; Sasaki, A.; Matsuda, A., Highly fluorescent 5-(5,6-dimethoxybenzothiazol-2-yl)-2′-deoxyuridine-5′-triphosphate as an efficient substrate for DNA polymerases. *ChemBioChem* **2011**, *12*, 2341-2346.

161. Vaish, N. K.; Fraley, A. W.; Szostak, J. W.; McLaughlin, L. W., Expanding the structural and functional diversity of RNA: analog uridine triphosphates as candidates for *in vitro* selection of nucleic acids. *Nucleic Acids Res.* **2000**, *28*, 3316-3322.

162. Zhu, Z. R.; Chao, J.; Yu, H.; Waggoner, A. S., Directly labeled DNA probes using fluorescent nucleotides with different length linkers. *Nucleic Acids Res.* **1994**, *22*, 3418-3422.

163. Yoshikawa, M.; Kato, T.; Takenishi, T., A novel method for phosphorylation of nucleosides to 5′-nucleotides. *Tetrahedron Lett.* **1967**, *50*, 5065-5068.

164. Ludwig, J.; Eckstein, F., Rapid and efficient synthesis of nucleoside 5′-*O*-(1-thiotriphosphates), 5′-triphosphates and 2′,3′-cyclophosphorothioates using 2-chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one. *J. Org. Chem.* **1989**, *54*, 631-635.

165. Wanninger-Weiß, C.; Wagenknecht, H. A., Synthesis of 5-(2-pyrenyl)-2′-deoxyuridine as a DNA modification for electron-transfer studies: the critical role of the position of the chromophore attachment. *Eur. J. Org. Chem.* **2008**, 64-69.

166. Ding, P.; Wunnicke, D.; Steinhoff, H. J.; Seela, F., Site-directed spin-labeling of DNA by the azide-alkyne 'click' reaction: nanometer distance measurements on 7-deaza-2′-deoxyadenosine and 2′-deoxyuridine nitroxide conjugates spatially separated or linked to a 'dA-dT' base pair. *Chem. Eur. J.* **2010**, *16*, 14385-14396.

167. Latham, J. A.; Johnson, R.; Toole, J. J., The application of a modified nucleotide in aptamer selection: novel thrombin aptamers containing 5-(1-pentynyl)-2′-deoxyuridine. *Nucleic Acids Res.* **1994**, *22*, 2817-2822.

168. Borsenberger, V.; Kukwikila, M.; Howorka, S., Synthesis and enzymatic incorporation of modified deoxyuridine triphosphates. *Org. Biomol. Chem.* **2009**, *7*, 3826-3835.

169. Sagi, J.; Szemzö, A.; Ebinger, K.; Szabolcs, A.; Sagi, G.; Ruff, E.; Ötvös, L., Base-modified oligodeoxynucleotides. I. Effect of 5-alkyl, 5-(1-alkenyl) and 5-(1-alkynyl) substitution of the pyrimidines on duplex stability and hydrophobicity. *Tetrahedron Lett.* **1993**, *34*, 2191-2194.

170. Znosko, B. M.; Barnes, T. W.; Krugh, T. R.; Turner, D. H., NMR studies of DNA single strands and DNA:RNA hybrids with and without 1-propynylation at C5 of oligopyrimidines. *J. Am. Chem. Soc.* **2003**, *125*, 6090-6097.

171. Hwang, G. T.; Seo, Y. J.; Kim, S. J.; Kim, B. H., Fluorescent oligonucleotide incorporating 5-(1-ethynylpyrenyl)-2′-deoxyuridine: sequence-specific fluorescence changes upon duplex formation. *Tetrahedron Lett.* **2004**, *45*, 3543-3546.

172. Sonogashira, K., Development of Pd–Cu catalyzed cross-coupling of terminal acetylenes with sp^2-carbon halides. *J. Organomet. Chem.* **2002**, *653*, 46-49.

173. Koohang, A.; Bailey, J. L.; Coates, R. M.; Erickson, H. K.; Owen, D.; Poulter, C. D., Enantioselective inhibition of squalene synthase by aziridine analogues of presqualene diphosphate. *J. Org. Chem.* **2010**, *75*, 4769-4777.

174. Korhonen, H. J.; Bolt, H. L.; Vicente-Gines, L.; Perks, D. C.; Hodgson, D. R. W., PPN pyrophosphate: a new reagent for the preparation of nucleoside triphosphates. *Phosphorus, Sulfur Silicon Relat. Elem.* **2015**, *190*, 758-762.

175. Lee, S. E.; Sidorov, A.; Gourlain, T.; Mignet, N.; Thorpe, S. J.; Brazier, J. A.; Dickman, M. J.; Hornby, D. P.; Grasby, J. A.; Williams, D. M., Enhancing the catalytic repertoire of nucleic acids: a systematic study of linker length and rigidity. *Nucleic Acids Res.* **2001**, *29*, 1565-1573.

176. Litosh, V. A.; Wu, W.; Stupi, B. P.; Wang, J.; Morris, S. E.; Hersh, M. N.; Metzker, M. L., Improved nucleotide selectivity and termination of 3′-OH unblocked reversible terminators by molecular tuning of 2-nitrobenzyl alkylated HOMedU triphosphates. *Nucleic Acids Res.* **2011**, *39*, e39.

177. Clark, J. M., Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* **1988**, *16*, 9677-9686.

178. Okamoto, A.; Tanaka, K.; Nishiza, K.; Saito, I., Synthesis of an artificial hole-transporting nucleoside triphosphate, $d^{MD}ATP$, and its enzymatic incorporation into DNA. *Bioorg. Med. Chem.* **2004**, *12*, 5875-5880.

179. Balintová, J.; Špaček, J.; Pohl, R.; Brázdová, M.; Havran, L.; Fojta, M.; Hocek, M., Azidophenyl as a click-transformable redox label of DNA suitable for electrochemical detection of DNA-protein interactions. *Chem. Sci.* **2015**, *6*, 575-587.

180. Weisbrod, S. H.; Marx, A., Novel strategies for the site-specific covalent labelling of nucleic acids. *Chem. Commun.* **2008**, 5675-5685.

181. Ménová, P.; Cahová, H.; Plucnara, M.; Havran, L.; Fojta, M.; Hocek, M., Polymerase synthesis of oligonucleotides containing a single chemically modified nucleobase for site-specific redox labelling. *Chem. Commun.* **2013**, *49*, 4652-4654.

182. Cramer, J.; Rangam, G.; Marx, A.; Restle, T., Varied active-site constraints in the Klenow fragment of E. coli DNA polymerase I and the lesion-bypass Dbh DNA polymerase. *ChemBioChem* **2008**, *9*, 1243-1250.

183. Cavanaugh, N. A.; Beard, W. A.; Wilson, S. H., DNA polymerase *β* ribonucleotide discrimination insertion, misinsertion, extension, and coding. *J. Biol. Chem.* **2010**, *285*, 24457-24465.

184. Wang, X.; Chen, X.; Liu, Y.; Zhu, J., Primer extension reaction assays for incorporation of deoxynucleotide analogue into DNA. *Chin. J. Chem.* **2015**, *33*, 192-198.

185. Ruparel, H.; Bi, L.; Li, Z.; Bai, X.; Kim, D. H.; Turro, N. J.; Ju, J., Design and synthesis of a 3′-*O*-allyl photocleavable fluorescent nucleotide as a reversible terminator for DNA sequencing by synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5932-5937.

186. Michaelis, L.; Menten, M. L.; Johnson, K. A.; Goody, R. S., The original Michaelis constant: translation of the 1913 Michaelis-Menten paper. *Biochemistry* **2011**, *50*, 8264-8269.

187. Voet, D., Voet, J. G., *Biochemistry*; John Wiley & sons, Inc., **2011**, pp 487-492, pp 1158-1158.

188. Harvey, B. J.; Perez, C.; Levitus, M., DNA sequence-dependent enhancement of Cy3 fluorescence. *Photochem. Photobiol. Sci.* **2009**, *8*, 1105-1110.

189. Ren, X.; Gerowska, M.; El-Sagheer, A. H.; Brown, T., Enzymatic incorporation and fluorescent labelling of cyclooctyne-modified deoxyuridine triphosphates in DNA. *Bioorg. Med. Chem.* **2014**, *22*, 4384-4390.

190. Krebs, J. F.; Kore, A. R., Novel FRET-based assay to detect reverse transcriptase activity using modified dUTP analogues. *Bioconjugate Chem.* **2008**, *19*, 185-191.

191. Hizi, A.; KamathLoeb, A. S.; Rose, K. D.; Loeb, L. A., Mutagenesis by human immunodeficiency virus reverse transcriptase: incorporation of *O*-6-methyldeoxyguanosine triphosphate. *Mutat. Res. Fund. Mol. Mech. Mut.* **1997**, *374*, 41-50.

192. Varela-Echavarría, A.; Garvey, N.; Preston, B. D.; Dougherty, J. P., Comparison of moloney murine leukemia virus mutation rate with the fidelity of its reverse transcriptase *in vitro*. *J. Biol. Chem.* **1992**, *267*, 24681-24688.

193. Okello, J. B.; Rodriguez, L.; Poinar, D.; Bos, K.; Okwi, A. L.; Bimenya, G. S.; Sewankambo, N. K.; Henry, K. R.; Kuch, M.; Poinar, H. N., Quantitative assessment of the sensitivity of various commercial reverse transcriptases based on armored HIV RNA. *PLoS One* **2010**, *5*, e13931.

194. Seitchik, J. L.; Peeler, J. C.; Taylor, M. T.; Blackman, M. L.; Rhoads, T. W.; Cooley, R. B.; Refakis, C.; Fox, J. M.; Mehl, R. A., Genetically encoded tetrazine amino acid directs rapid site-specific *in vivo* bioorthogonal ligation with *trans*-cyclooctenes. *J. Am. Chem. Soc.* **2012**, *134*, 2898-2901.

195. Liu, D. S.; Tangpeerachaikul, A.; Selvaraj, R.; Taylor, M. T.; Fox, J. M.; Ting, A. Y., Diels-Alder cycloaddition for fluorophore targeting to specific proteins inside living cells. *J. Am. Chem. Soc.* **2012**, *134*, 792-795.

196. Neves, A. A.; Stöckmann, H.; Wainman, Y. A.; Kuo, J. C. H.; Fawcett, S.; Leeper, F. J.; Brindle, K. M., Imaging cell surface glycosylation *in vivo* using ‘double click’ chemistry. *Bioconjugate Chem.* **2013**, *24*, 934-941.

197. Meimetis, L. G.; Carlson, J. C. T.; Giedt, R. J.; Kohler, R. H.; Weissleder, R., Ultrafluorogenic coumarin-tetrazine probes for real-time biological imaging. *Angew. Chem. Int. Ed.* **2014**, *53*, 7531-7534.

198. Dziuba, D.; Pohl, R.; Hocek, M., Polymerase synthesis of DNA labelled with benzylidene cyanoacetamide-based fluorescent molecular rotors: fluorescent light-up probes for DNA-binding proteins. *Chem. Commun.* **2015**, *51*, 4880-4882.

199. Amblar, M.; de Lacoba, M. G.; Corrales, M. A.; López, P., Biochemical analysis of point mutations in the 5′-3′ exonuclease of DNA polymerase I of Streptococcus

pneumoniae – functional and structural implications. *J. Biol. Chem.* **2001**, *276*, 19172-19181.

200. Derbyshire, V.; Grindley, N. D.; Joyce, C. M., The 3′-5′ exonuclease of DNA polymerase I of Escherichia coli: contribution of each amino acid at the active site to the reaction. *The EMBO Journal* **1991**, *10*, 17-24.

201. Højland, T.; Veedu, R. N.; Vester, B.; Wengel, J., Enzymatic synthesis of DNA strands containing α-L-LNA (α-L-configured locked nucleic acid) thymine nucleotides. *Artif DNA PNA XNA* **2012**, *3*, 14-21.

202. Wetmur, J. G., DNA probes: applications of the principles of nucleic acid hybridization. *Crit. Rev. Biochem. Mol. Biol.* **1991**, *26*, 227-259.

203. Sanborn, M. E.; Connolly, B. K.; Gurunathan, K.; Levitus, M., Fluorescence properties and photophysics of the sulfoindocyanine Cy3 linked covalently to DNA. *J. Phys. Chem. B* **2007**, *111*, 11064-11074.

204. Gebinoga, M.; Oehlenschläger, F., Comparison of self-sustained sequence-replication reaction systems. *Eur. J. Biochem.* **1996**, *235*, 256-261.

205. Li, L.; Li, X. C.; Li, L.; Wang, J. X.; Jin, W. R., Ultra-sensitive DNA assay based on single-molecule detection coupled with fluorescent quantum dot-labeling and its application to determination of messenger RNA. *Anal. Chim. Acta* **2011**, *685*, 52-57.

206. Wong, L.; Pearson, H.; Fletcher, A.; Marquis, C. P.; Mahler, S., Comparison of the efficiency of Moloney Murine Leukaemia Virus (M-MuLV) reverse transcriptase, RNase H^--M-MuLV reverse transcriptase and Avian Myeloblastoma Leukaemia Virus (AMV) reverse transcriptase for the amplification of human immunoglobulin genes. *Biotechnol. Tech.* **1998**, *12*, 485-489.

207. Stoecker, K.; Dorninger, C.; Daims, H.; Wagner, M., Double labeling of oligonucleotide probes for fluorescence *in situ* hybridization (DOPE-FISH) improves signal intensity and increases rRNA accessibility. *Appl. Environ. Microbiol.* **2010**, *76*, 922-926.

208. Jewett, J. C.; Bertozzi, C. R., Cu-free click cycloaddition reactions in chemical biology. *Chem. Soc. Rev.* **2010**, *39*, 1272-1279.

209. Rao, H.; Tanpure, A. A.; Sawant, A. A.; Srivatsan, S. G., Enzymatic incorporation of an azide-modified UTP analog into oligoribonucleotides for post-transcriptional chemical functionalization. *Nat. Protoc.* **2012**, *7*, 1097-1112.

210. Yang, M. J.; Cao, J., Principle and applications of multicolor fluorescence *in situ* hybridization technology. *Prog. Biochem. Biophys.* **1998**, *25*, 333-337.

211. Lampkins, A. J.; O'Neil, E. J.; Smith, B. D., Bio-orthogonal phosphatidylserine conjugates for delivery and imaging applications. *J. Org. Chem.* **2008**, *73*, 6053-6058.

212. Collins, M. L.; Irvine, B.; Tyner, D.; Fine, E.; Zayati, C.; Chang, C.; Horn, T.; Ahle, D.; Detmer, J.; Shen, L.; Kolberg, J.; Bushnell, S.; Urdea, M. S.; Ho, D. D., A branched DNA signal amplification assay for quantification of nucleic acid targets below 100 molecules/mL. *Nucleic Acids Res.* **1997**, *25*, 2979-2984.

213. Wang, F.; Flanagan, J.; Su, N.; Wang, L. C.; Bui, S.; Nielson, A.; Wu, X.; Vo, H. T.; Ma, X. J.; Luo, Y., RNAscope – a novel *in situ* RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J. Mol. Diagn.* **2012**, *14*, 22-29.

214. Sanger, F.; Nicklen, S.; Coulson, A. R., DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 5463-5467.

215. Gruber, H. J.; Hahn, C. D.; Kada, G.; Riener, C. K.; Harms, G. S.; Ahrer, W.; Dax, T. G.; Knaus, H. G., Anomalous fluorescence enhancement of Cy3 and Cy3.5 versus anomalous fluorescence loss of Cy5 and Cy7 upon covalent linking to IgG and noncovalent binding to avidin. *Bioconjugate Chem.* **2000**, *11*, 696-704.

216. Kubista, M.; Sjöback, R.; Nygren, J., Quantitative spectral analysis of multicomponent equilibria. *Anal. Chim. Acta* **1995**, *302*, 121-125.

217. Sjöback, R.; Nygren, J.; Kubista, M., Absorption and fluorescence properties of fluorescein. *Spectrochim. Acta, Part A* **1995**, *51*, L7-L21.

218. An, S. F.; Franklin, D.; Fleming, K. A., Generation of digoxigenin-labelled double-stranded and single-stranded probes using the polymerase chain reaction. *Mol. Cell. Probes* **1992**, *6*, 193-200.

219. Svobodová, M.; Pinto, A.; Nadal, P.; O' Sullivan, C. K., Comparison of different methods for generation of single-stranded DNA for SELEX processes. *Anal. Bioanal.Chem.* **2012**, *404*, 835-842.

220. Wirges, C. T.; Gramlich, P. M. E.; Gutsmiedl, K.; Gierlich, J.; Burley, G. A.; Carell, T., Pronounced effect of DNA hybridization on click reaction efficiency. *QSAR Comb. Sci.* **2007**, *26*, 1159-1164.

221. Higuchi, R. G.; Ochman, H., Production of single-stranded DNA templates by exonuclease digestion following the polymerase chain reaction. *Nucleic Acids Res.* **1989**, *17*, 5865-5865.

222. Kovall, R.; Matthews, B. W., Toroidal structure of λ-exonuclease. *Science* **1997**, *277*, 1824-1827.

223. Mitsis, P. G.; Kwagh, J. G., Characterization of the interaction of lambda exonuclease with the ends of DNA. *Nucleic Acids Res.* **1999**, *27*, 3057-3063.

224. de Muro, M., Probe design, production, and applications. In *Medical Biomethods Handbook*, Humana Press, **2005**, pp 13-23.

225. Rogan, P. K.; Cazcarro, P. M.; Knoll, J. H. M., Sequence-based design of single-copy genomic DNA probes for fluorescence *in situ* hybridization. *Genome Res.* **2001**, *11*, 1086-1094.

226. Arvey, A.; Hermann, A.; Hsia, C. C.; Ie, E.; Freund, Y.; McGinnis, W., Minimizing off-target signals in RNA fluorescent *in situ* hybridization. *Nucleic Acids Res.* **2010**, *38*, e115.

227. Pierce, K. E.; Sanchez, J. A.; Rice, J. E.; Wangh, L. J., Linear-after-the-exponential (LATE)-PCR: primer design criteria for high yields of specific single-stranded DNA and improved real-time detection. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 8609-8614.

228. Nakano, S.; Fujimoto, M.; Hara, H.; Sugimoto, N., Nucleic acid duplex stability: influence of base composition on cation effects. *Nucleic Acids Res.* **1999**, *27*, 2957-2965.

229. Wilchek, M.; Bayer, E. A., Introduction to avidin-biotin technology. In *Methods in Enzymology*, Academic Press, **1990**, pp 5-13.

230. Demidov, V. V.; Bukanov, N. O.; Frank-Kamenetskii, D., Duplex DNA capture. *Curr. Issues Mol. Biol.* **2000**, *2*, 31-35.

231. Weber, P. C.; Wendoloski, J. J.; Pantoliano, M. W.; Salemme, F. R., Crystallographic and thermodynamic comparison of natural and synthetic ligands bound to streptavidin. *J. Am. Chem. Soc.* **1992**, *114*, 3197-3200.

232. Shamah, S. M.; Healy, J. M.; Cload, S. T., Complex target SELEX. *Acc. Chem. Res.* **2008**, *41*, 130-138.

233. Ulrich, H.; Magdesian, M. H.; Alves, M. J. M.; Colli, W., *In vitro* selection of RNA aptamers that bind to cell adhesion receptors of Trypanosoma cruzi and inhibit cell invasion. *J. Biol. Chem.* **2002**, *277*, 20756-20762.

234. Li, Y.; Lee, H. J.; Corn, R. M., Fabrication and characterization of RNA aptamer microarrays for the study of protein-aptamer interactions with SPR imaging. *Nucleic Acids Res.* **2006**, *34*, 6416-6424.

235. Chiou, S. H., DNA- and protein-scission activities of ascorbate in the presence of copper ion and a copper-peptide complex. *J. Biochem.* **1983**, *94*, 1259-1267.

236. Imlay, J. A.; Linn, S., DNA damage and oxygen radical toxicity. *Science* **1988**, *240*, 1302-1309.

237. Gaetke, L. M.; Chow, C. K., Copper toxicity, oxidative stress, and antioxidant nutrients. *Toxicology* **2003**, *189*, 147-163.

238. Farkaš, J.; Šorm, F., Nucleic acid components and their analogs. CXXV. Synthesis of 5-[bis(2-chloroethyl)aminomethyl]uridine hydrochloride. *Collect. Czech. Chem. Commun.* **1969**, *34*, 1696-1701.

239. Buhr, C. A.; Matteucci, M. D.; Froehler, B. C., Synthesis of a tetracyclic 2′-deoxyadenosine analog. *Tetrahedron Lett.* **1999**, *40*, 8969-8970.

240. Balis, J. U.; Bumgarner, S. D.; Paciga, J. E.; Paterson, J. F.; Shelley, S. A., Synthesis of lung surfactant-associated glycoproteins by A549 cells: description of an *in vitro* model for human type-II cell dysfunction. *Exp. Lung Res.* **1984**, *6*, 197-213.

241. Giard, D. J.; Aaronson, S. A.; Todaro, G. J.; Arnstein, P.; Kersey, J. H.; Dosik, H.; Parks, W. P., *In vitro* cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. *J. Natl. Cancer Inst.* **1973**, *51*, 1417-1423.

242. Lowe, A. B., Thiol-yne 'click'/coupling chemistry and recent applications in polymer and materials synthesis and modification. *Polymer* **2014**, *55*, 5517-5549.

243. Tomassini, J. E.; Getty, K.; Stahlhut, M. W.; Shim, S.; Bhat, B.; Eldrup, A. B.; Prakash, T. P.; Carroll, S. S.; Flores, O.; MacCoss, M.; McMasters, D. R.; Migliaccio, G.; Olsen, D. B., Inhibitory effect of 2′-substituted nucleosides on hepatitis C virus replication correlates with metabolic properties in replicon cells. *Antimicrob. Agents Chemother.* **2005**, *49*, 2050-2058.

244. Taylor, M. T.; Blackman, M. L.; Dmitrenko, O.; Fox, J. M., Design and synthesis of highly reactive dienophiles for the tetrazine-*trans*-cyclooctene ligation. *J. Am. Chem. Soc.* **2011**, *133*, 9646-9649.

245. Rossin, R.; van den Bosch, S. M.; ten Hoeve, W.; Carvelli, M.; Versteegen, R. M.; Lub, J.; Robillard, M. S., Highly reactive *trans*-cyclooctene tags with improved stability for Diels-Alder chemistry in living systems. *Bioconjugate Chem.* **2013**, *24*, 1210-1217.

246. McGuigan, C.; Pathirana, R. N.; Mahmood, N.; Devine, K. G.; Hay, A. J., Aryl phosphate derivatives of AZT retain activity against HIV1 in cell lines which are resistant to the action of AZT. *Antiviral Res.* **1992**, *17*, 311-321.

247. McGuigan, C.; Pathirana, R. N.; Mahmood, N.; Hay, A. J., Aryl phosphate derivates of AZT inhibit HIV replication in cells where the nucleoside is poorly active. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 701-704.

248. McGuigan, C.; Tollerfield, S. M.; Riley, P. A., Synthesis and biological evaluation of some phosphate triester derivatives of the anti-viral drug AraA. *Nucleic Acids Res.* **1989**, *17*, 6065-6075.

249. Chang, W.; Bao, D. H.; Chun, B. K.; Naduthambi, D.; Nagarathnam, D.; Rachakonda, S.; Reddy, P. G.; Ross, B. S.; Zhang, H. R.; Bansal, S.; Espiritu, C. L.; Keilman, M.; Lam, A. M.; Niu, C.; Steuer, H. M.; Furman, P. A.; Otto, M. J.; Sofia, M. J., Discovery of PSI-353661, a novel purine nucleotide prodrug for the treatment of HCV infection. *ACS Med. Chem. Lett.* **2011**, *2*, 130-135.

250. McGuigan, C.; Hassan-Abdallah, A.; Srinivasan, S.; Wang, Y. K.; Siddiqui, A.; Daluge, S. M.; Gudmundsson, K. S.; Zhou, H. Q.; McLean, E. W.; Peckham, J. P.; Burnette, T. C.; Marr, H.; Hazen, R.; Condreay, L. D.; Johnson, L.; Balzarini, J., Application of phosphoramidate ProTide technology significantly improves antiviral potency of carbocyclic adenosine derivatives. *J. Med. Chem.* **2006**, *49*, 7215-7226.

251. McGuigan, C.; Murziani, P.; Slusarczyk, M.; Gonczy, B.; Vande Voorde, J.; Liekens, S.; Balzarini, J., Phosphoramidate ProTides of the anticancer agent FUDR successfully deliver the preformed bioactive monophosphate in cells and confer advantage over the parent nucleoside. *J. Med. Chem.* **2011**, *54*, 7247-7258.

252. Derudas, M.; Brancale, A.; Naesens, L.; Neyts, J.; Balzarini, J.; McGuigan, C., Application of the phosphoramidate ProTide approach to the antiviral drug ribavirin. *Bioorg. Med. Chem.* **2010**, *18*, 2748-2755.

253. Saboulard, D.; Naesens, L.; Cahard, D.; Salgado, A.; Pathirana, R.; Velazquez, S.; McGuigan, C.; De Clercq, E.; Balzarini, J., Characterization of the activation pathway of phosphoramidate triester prodrugs of stavudine and zidovudine. *Mol. Pharmacol.* **1999**, *56*, 693-704.

254. McGuigan, C.; Thiery, J. C.; Daverioa, F.; Jiang, W. G.; Davies, G.; Mason, M., Anti-cancer ProTides: tuning the activity of BVDU phosphoramidates related to thymectacin. *Bioorg. Med. Chem.* **2005**, *13*, 3219-3227.

255. Dai, C.; Wang, L.; Sheng, J.; Peng, H.; Draganov, A. B.; Huang, Z.; Wang, B., The first chemical synthesis of boronic acid-modified DNA through a copper-free click reaction. *Chem. Commun.* **2011**, *47*, 3598-3600.

256. Trybulski, E. J.; Zhang, J.; Kramss, R. H.; Mangano, R. M., The synthesis and biochemical pharmacology of enantiomerically pure methylated oxotremorine derivatives. *J. Med. Chem.* **1993**, *36*, 3533-3541.

257. Hong, I. S.; Ding, H.; Greenberg, M. M., Oxygen independent DNA interstrand cross-link formation by a nucleotide radical. *J. Am. Chem. Soc.* **2006**, *128*, 485-491.

258. Kvach, M. V.; Ustinov, A. V.; Stepanova, I. A.; Malakhov, A. D.; Skorobogatyi, M. V.; Shmanai, V. V.; Korshun, V. A., A convenient synthesis of cyanine dyes: reagents for the labeling of biomolecules. *Eur. J. Org. Chem.* **2008**, *2008*, 2107-2117.

259. Haralambidis, J.; Chai, M.; Tregear, G. W., Preparation of base-modified nucleosides suitable for non-radioactive label attachment and their incorporation into synthetic oligodeoxyribonucleotides. *Nucleic Acids Res.* **1987**, *15*, 4857-4876.

8 Appendix

8.1 Authority permission for picture reuse from the American Society for Biochemistry and Molecular Biology

11200 Rockville Pike
Suite 302
Rockville, Maryland 20852

August 19, 2011

American Society for Biochemistry and Molecular Biology

To whom it may concern,

It is the policy of the American Society for Biochemistry and Molecular Biology to allow reuse of any material published in its journals (the Journal of Biological Chemistry, Molecular & Cellular Proteomics and the Journal of Lipid Research) in a thesis or dissertation at no cost and with no explicit permission needed. Please see our copyright permissions page on the journal site for more information.

Best wishes,

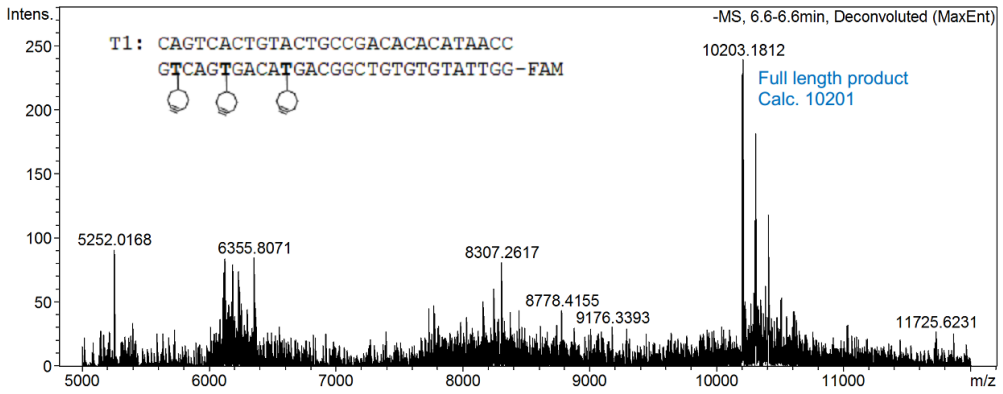
Sarah Crespi

American Society for Biochemistry and Molecular Biology
11200 Rockville Pike, Rockville, MD
Suite 302
240-283-6616
JBC | MCP | JLR

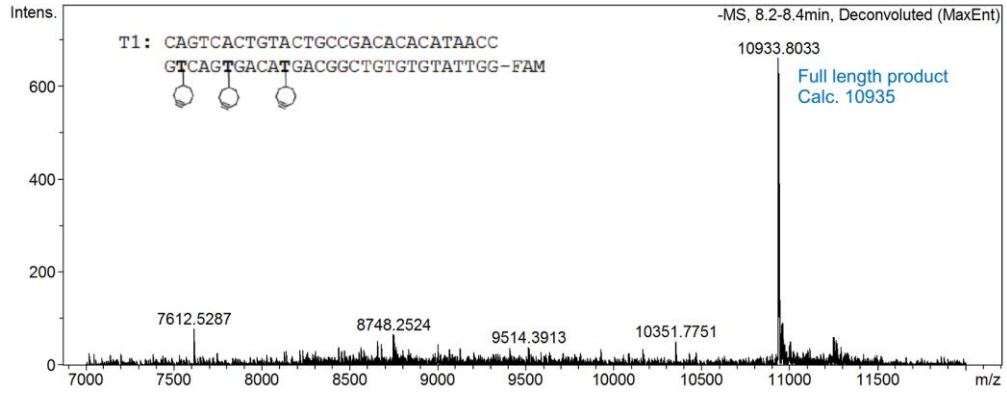
Tel: 240-283-6600 • Fax: 240-881-2080 • E-mail: asbmb@asbmb.org

8.2 Representative mass spectrometry analyses for cyclooctyne-modified primer extension products

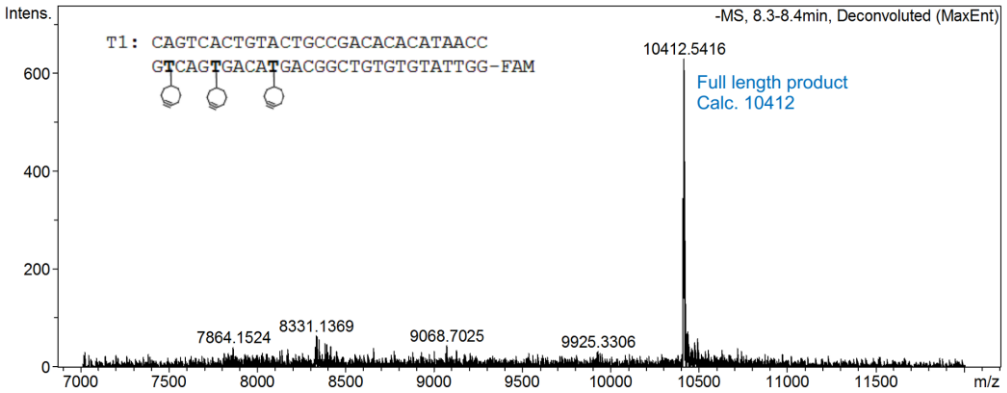
BCN-I dU template T1 product from Klenow polymerase



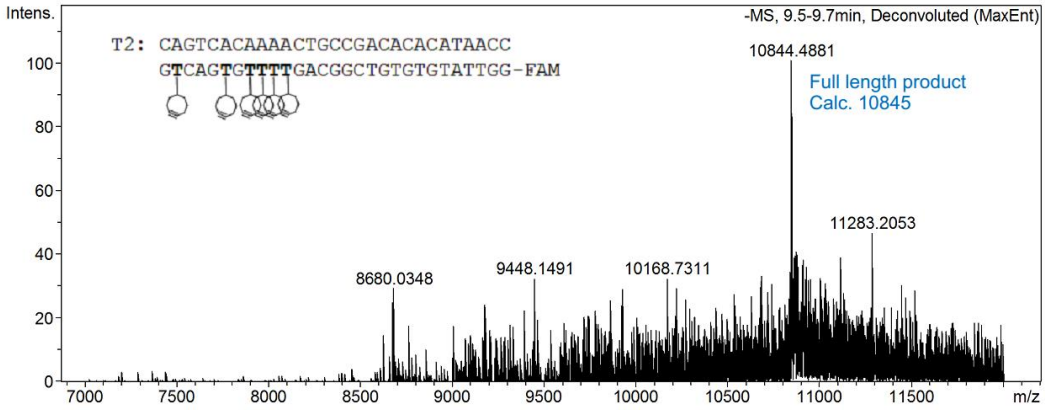
BCN-II dU template T1 product from Gotaq polymerase



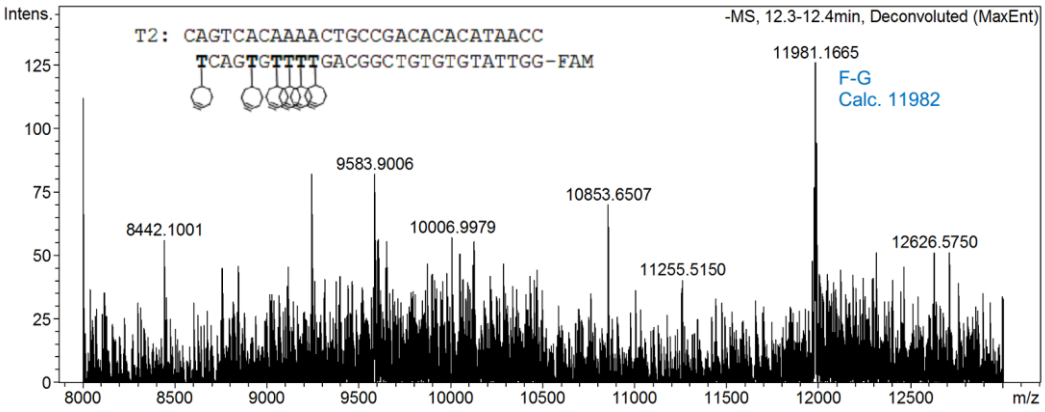
DIBO dU template T1 product from Gotaq polymerase



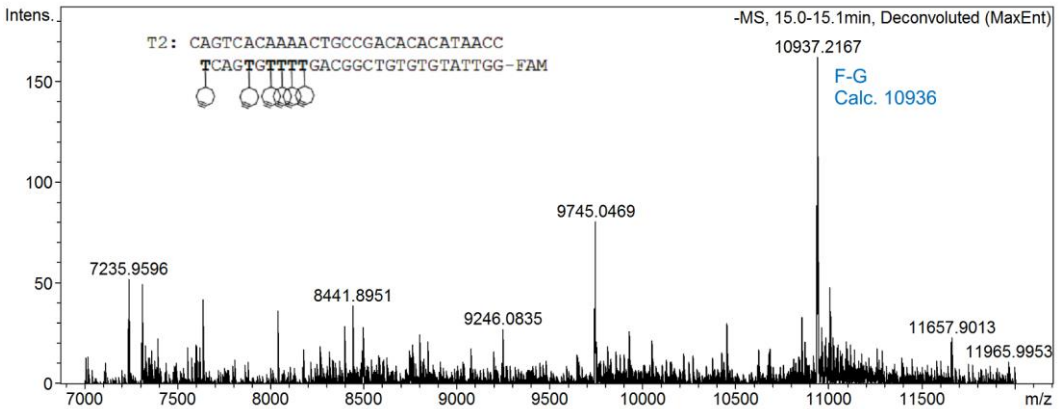
BCN-I dU template T2 product from KOD polymerase



BCN-II dU template T2 product from KOD polymerase

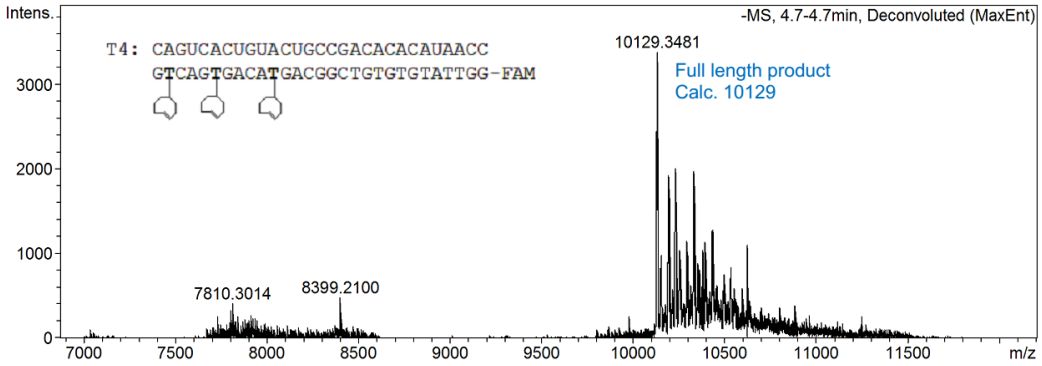


DIBO dU template T2 product from Klenow polymerase

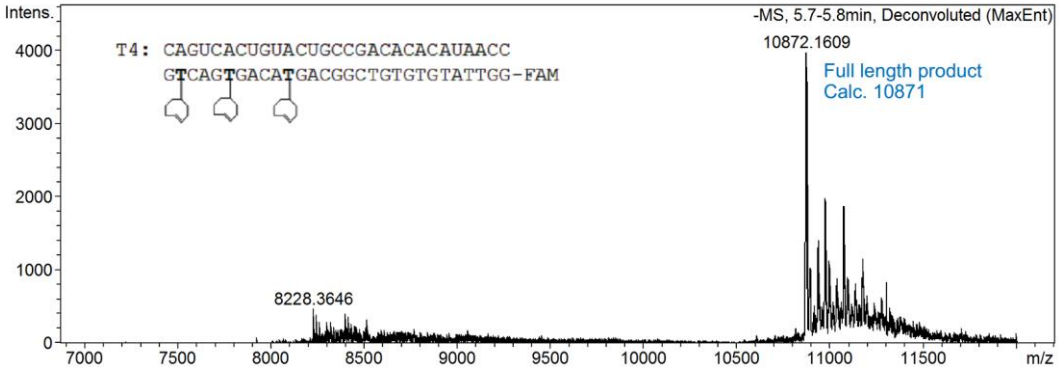


8.3 Representative mass spectrometry analyses for *trans*-cyclooctene-modified primer extension products

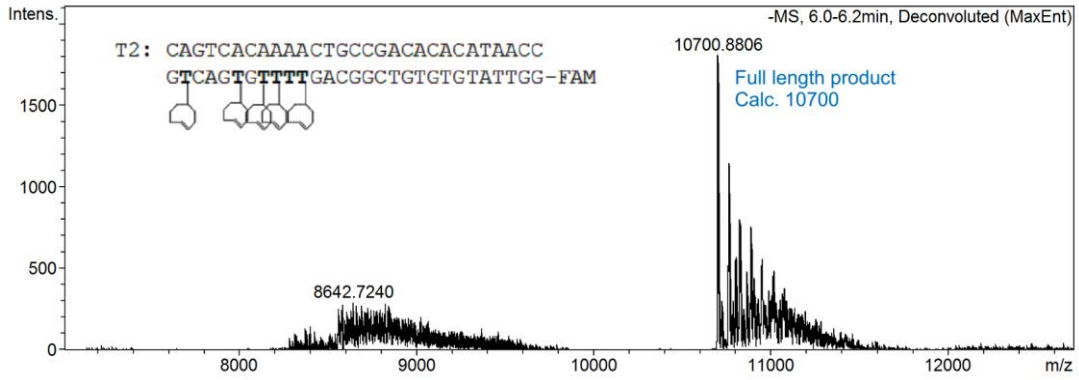
TCO-S dU template T4 product from M-MuLV (RNase H$^-$) reverse transcriptase



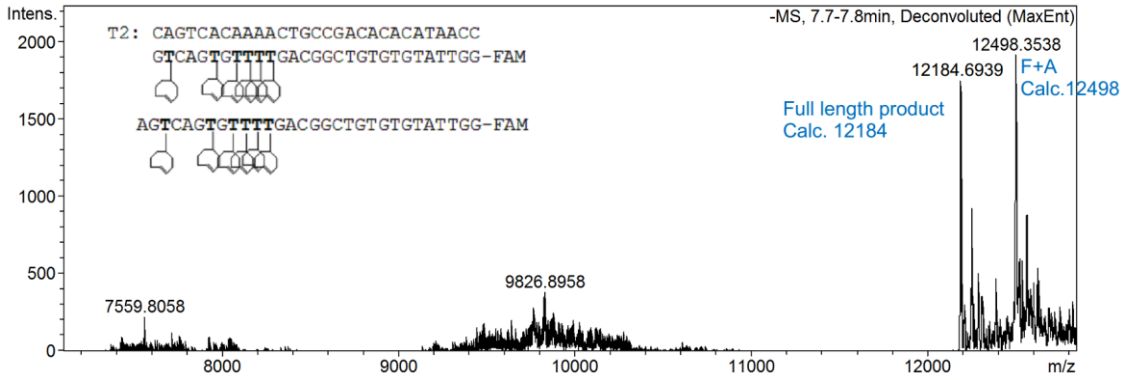
TCO-L dU template T4 product from M-MuLV (RNase H$^-$) reverse transcriptase



TCO-S dU template T2 product from KOD polymerase

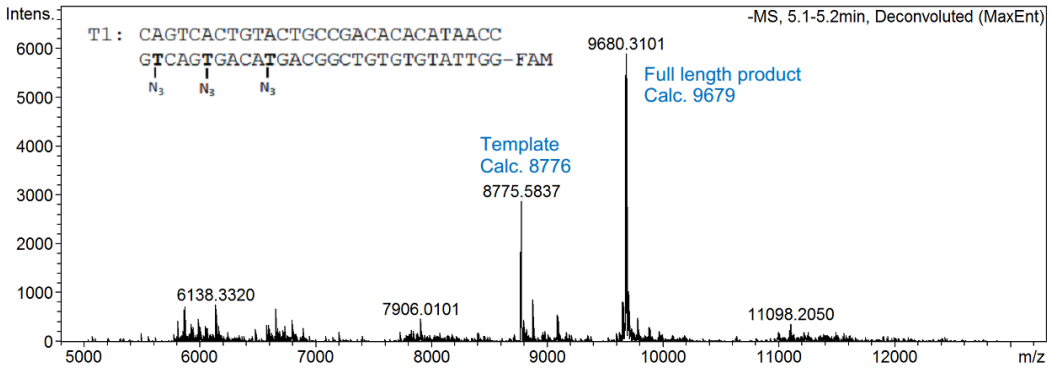


TCO-L dU template T2 product from Gotaq polymerase

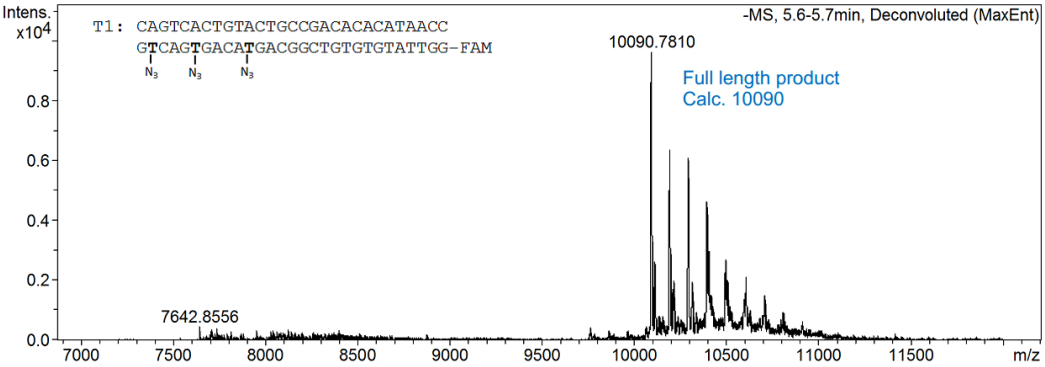


8.4 Representative mass spectrometry results for azide-modified primer extension products

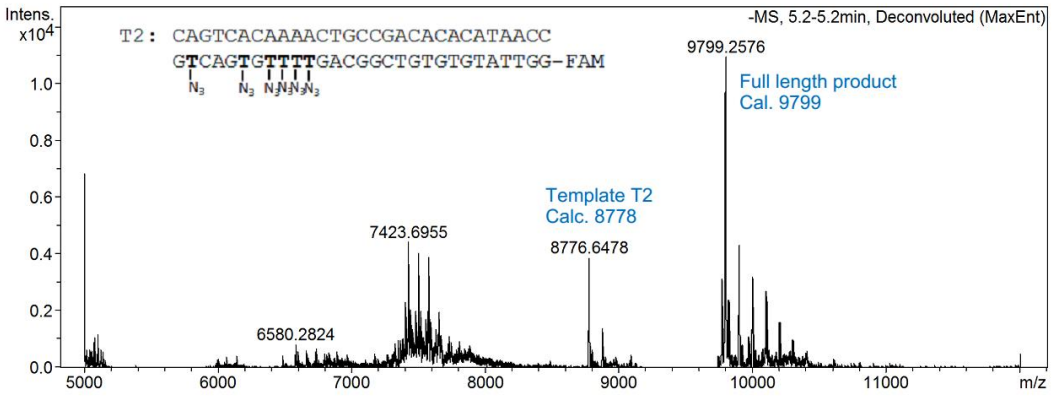
AM dU template T1 product from Klenow polymerase



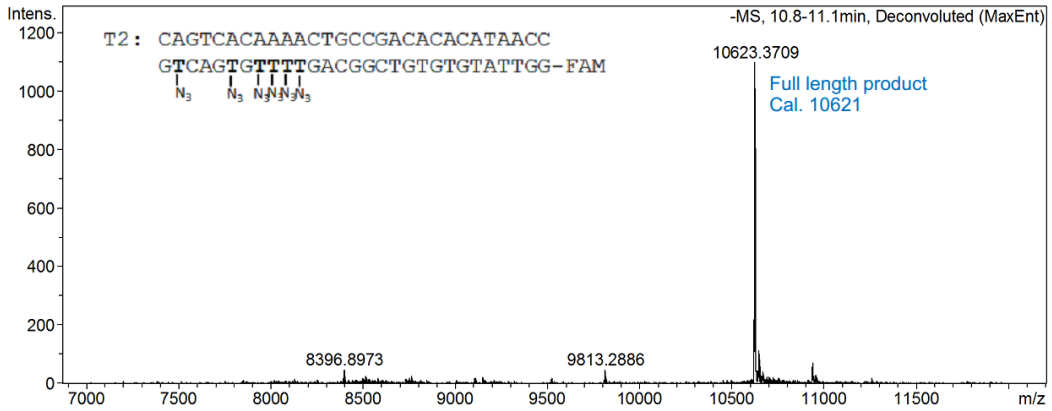
AHP dU template T1 product from Klenow polymerase



AM dU template T2 product from Klenow polymerase



AHP dU template T2 product from Klenow polymerase



8.5 Melting temperature measurements

8.5.1 Melting temperatures of TCO-modified primer extension products

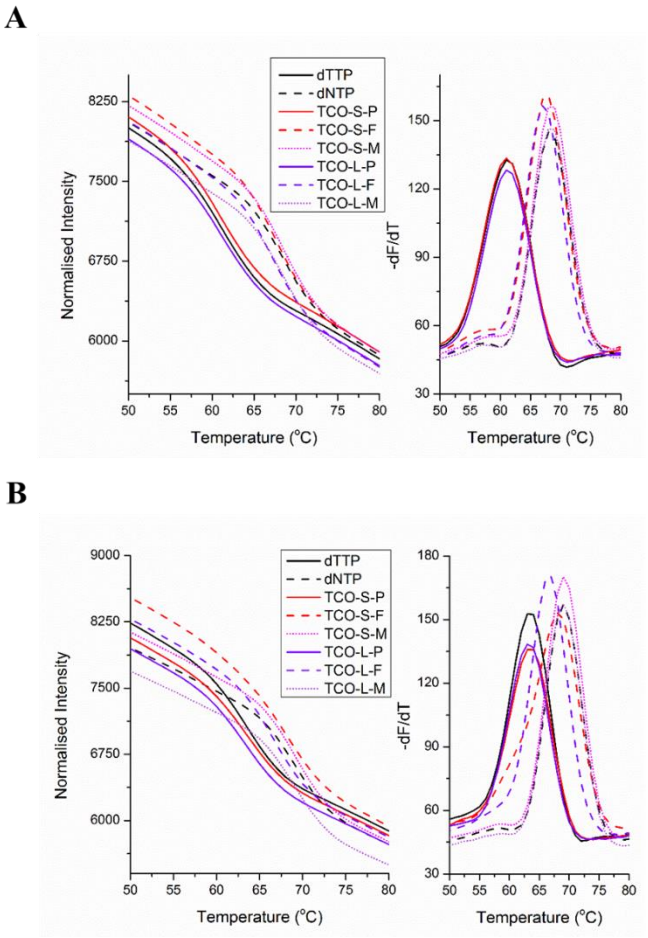


Figure 8.1. Melting curves and derivatives of TCO-modified primer extension products from Gotaq polymerase, template T1 (**A**) and T2 (**B**). dTTP, product from dTTP; dNTP, product from dNTPs; P, product from TCO dUTP; F, product from TCO dUTP + dATP + dCTP + dGTP; M, product from dTTP and TCO dUTP + dATP + dCTP + dGTP.

8.5.2 Melting temperatures of azide-modified primer extension products

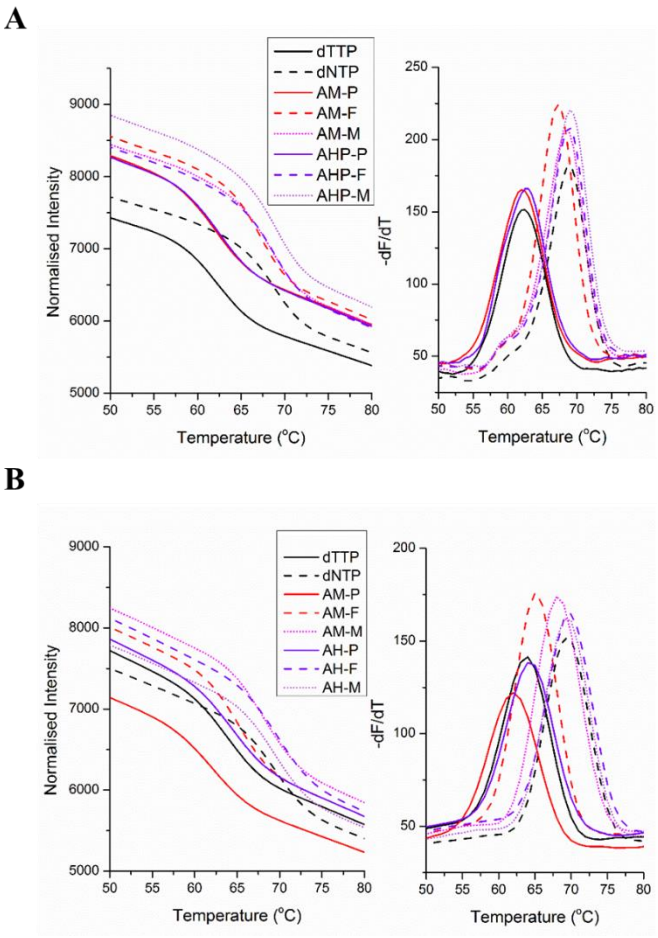


Figure 8.2. Melting curves and derivatives of azide-modified primer extension products from Gotaq polymerase, template T1 (**A**) and T2 (**B**). dTTP, product from dTTP; dNTP, product from dNTPs; P, product from azide dUTP; F, product from azide dUTP + dATP + dCTP + dGTP; M, product from dTTP and azide dUTP + dATP + dCTP + dGTP.

8.6 Michaelis-Menten kinetics

8.6.1 Initial reaction rate (v_0) fitting results of DIBO dUTP using Gotaq polymerase

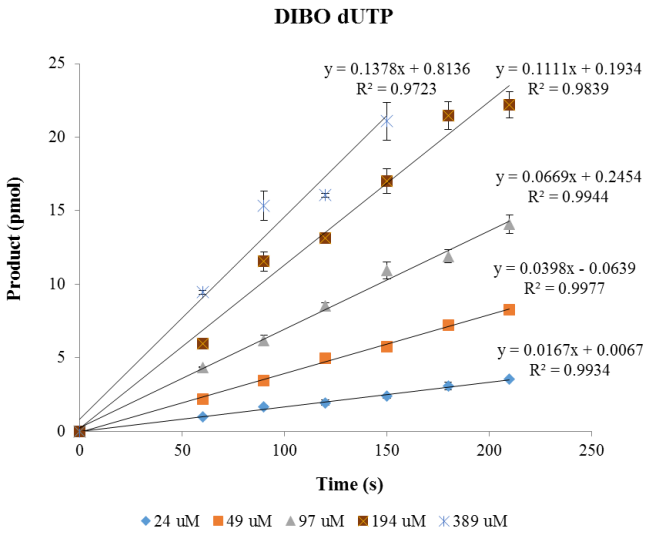


Figure 8.3. Single nucleotide incorporation initial reaction rate (v_0) for different concentrations of DIBO dUTP using Gotaq polymerase.

8.6.2 Michaelis-Menten kinetics using KOD polymerase

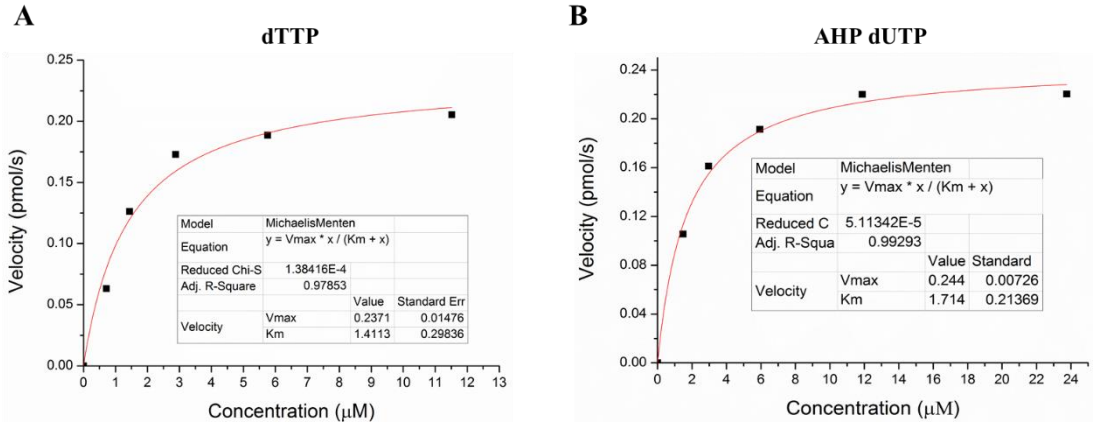


Figure 8.4. Michaelis-Menten kinetics for dTTP (**A**) and AHP dUTP (**B**) single nucleotide incorporation using KOD polymerase.

8.7 Sequencing results of the second-round unmodified PCR products

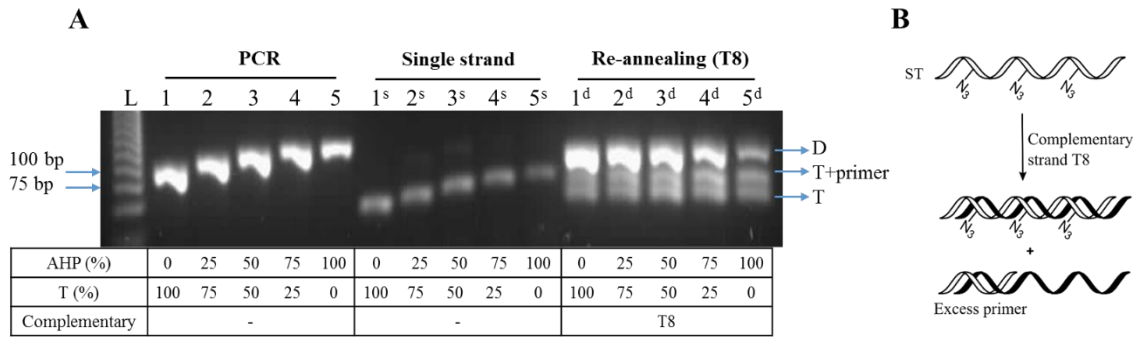
The yellow and grey regions are the two primer regions. P0, amplicon from unmodified template; PH1, amplicon from 25% AHP-modified template; PH1, amplicon from 50% AHP-modified template; PH1, amplicon from 75% AHP-modified template; PH1, amplicon from 100% AHP-modified template; template, the original template T11.

Template starts: ACTGCAATACGAATAATGGCTACAATTTTTTTCAAAAATCTCGTCCGATATTGCATTT

```
                          10         20         30         40         50
P0_primer_10        GAGAAGTTTT CAGGATCAAT GTCATCTGCA GGACGTATAC GAGGAACGCT
PH1_primer_10       GAGAAGTTTT CAGGATCAAT GTCATCTGCA GGACGTATAC GAGGAACGCT
PH2_primer_10       GAGAAGTTTT CAGGATCAAT GTCATCTGCA GGACGTATAC GAGGAACGCT
PH3_primer_10       GAGAAGTTTT CAGGATCAAT GTCATCTGCA GGACGTATAC GAGGAACGCT
PH4_primer_10       GAGAAGTTTT CAGGATCAAT GTCATCTGCA GGACGTATAC GAGGAACGCT
Template            GAGAAGTTTT CAGGATCAAT GTCATCTGCA GGACGTATAC GAGGAACGCT
                          60         70         80         90         100
P0_primer_10        TATAGTATGA GGGCCTACAC CCATAGCAGC TTCCAAGTGT TCAGCATGCA
PH1_primer_10       TATAGTATGA GGGCCTACAC CCATAGCAGC TTCCAAGTGT TCAGCATGCA
PH2_primer_10       TATAGTATGA GGGCCTACAC CCATAGCAGC TTCCAAGTGT TCAGCATGCA
PH3_primer_10       TATAGTATGA GGGCCTACAC CCATAGCAGC TTCCAAGTGT TCAGCATGCA
PH4_primer_10       TATAGTATGA GGGCCTACAC CCATAGCAGC TTCCAAGTGT TCAGCATGCA
Template            TATAGTATGA GGGCCTACAC CCATAGCAGC TTCCAAGTGT TCAGCATGCA
                         110        120        130        140        150
P0_primer_10        TTAGAAGTCC AACAAAGTCA TATTTGTACA TATTTAGTCC AAACAAAACC
PH1_primer_10       TTAGAAGTCC AACAAAGTCA TATTTGTACA TATTTAGTCC AAACAAAACC
PH2_primer_10       TTAGAAGTCC AACAAAGTCA TATTTGTACA TATTTAGTCC AAACAAAACC
PH3_primer_10       TTAGAAGTCC AACAAAGTCA TATTTGTACA TATTTAGTCC AAACAAAACC
PH4_primer_10       TTAGAAGTCC AACAAAGTCA TATTTGTACA TATTTAGTCC AAACAAAACC
Template            TTAGAAGTCC AACAAAGTCA TATTTGTACA TATTTAGTCC AAACAAAACC
                         160        170        180        190        200
P0_primer_10        CCAATACCTA CATCATCAAT ACCACCTTCC ATAGCACGAT CCATTGCTTC
PH1_primer_10       CCAATACCTA CATCATCAAT ACCACCTTCC ATAGCACGAT CCATTGCTTC
PH2_primer_10       CCAATACCTA CATCATCAAT ACCACCTTCC ATAGCACGAT CCATTGCTTC
PH3_primer_10       CCAATACCTA CATCATCAAT ACCACCTTCC ATAGCACGAT CCATTGCTTC
PH4_primer_10       CCAATACCTA CATCATCAAT ACCACCTTCC ATAGCACGAT CCATTGCTTC
Template            CCAATACCTA CATCATCAAT ACCACCTTCC ATAGCACGAT CCATTGCTTC
                         210        220        230        240        250
P0_primer_10        TGTATGATAG GCATAATCAT GTTTTGGACC TGTAGGATGA AGTTCCTCGT
PH1_primer_10       TGTATGATAG GCATAATCAT GTTTTGGACC TGTAGGATGA AGTTCCTCGT
PH2_primer_10       TGTATGATAG GCATAATCAT GTTTTGGACC TGTAGGATGA AGTTCCTCGT
PH3_primer_10       TGTATGATAG GCATAATCAT GTTTTGGACC TGTAGGATGA AGTTCCTCGT
PH4_primer_10       TGTATGATAG GCATAATCAT GTTTTGGACC TGTAGGATGA AGTTCCTCGT
Template            TGTATGATAG GCATAATCAT GTTTTGGACC TGTAGGATGA AGTTCCTCGT
                         260        270        280        290        300
P0_primer_10        AACTTTTTTT GTTATAGGTT TCTTGGAAAA GTATATATGT TCCAATACCA
PH1_primer_10       AACTTTTTTT GTTATAGGTT TCTTGGAAAA GTATATATGT TCCAATACCA
PH2_primer_10       AACTTTTTTT GTTATAGGTT TCTTGGAAAA GTATATATGT TCCAATACCA
PH3_primer_10       AACTTTTTTT GTTATAGGTT TCTTGGAAAA GTATATATGT TCCAATACCA
PH4_primer_10       AACTTTTTTT GTTATAGGTT TCTTGGAAAA GTATATATGT TCCAATACCA
Template            AACTTTTTTT GTTATAGGTT TCTTGGAAAA GTATATATGT TCCAATACCA
                         310        320        330        340        350
P0_primer_10        GCATCCTTTA ATTTCTTGTA GTTTTCTACA GTAGTAGCTG CAATATTTAC
PH1_primer_10       GCATCCTTTA ATTTCTTGTA GTTTTCTACA GTAGTAGCTG CAATATTTAC
PH2_primer_10       GCATCCTTTA ATTTCTTGTA GTTTTCTACA GTAGTAGCTG CAATATTTAC
PH3_primer_10       GCATCCTTTA ATTTCTTGTA GTTTTCTACA GTAGTAGCTG CAATATTTAC
PH4_primer_10       GCATCCTTTA ATTTCTTGTA GTTTTCTACA GTAGTAGCTG CAATATTTAC
Template            GCATCCTTTA ATTTCTTGTA GTTTTCTACA GTAGTAGCTG CAATATTTAC
                         360        370        380        390        400
```

```
P0_primer_10        ATTTACACGT CTAATTGCTC CATTTTTATG TTTTATGCTG TATATGGTTT
PH1_primer_10       ATTTACACGT CTAATTGCTC CATTTTTATG TTTTATGCTG TATATGGTTT
PH2_primer_10       ATTTACACGT CTAATTGCTC CATTTTTATG TTTTATGCTG TATATGGTTT
PH3_primer_10       ATTTACACGT CTAATTGCTC CATTTTTATG TTTTATGCTG TATATGGTTT
PH4_primer_10       ATTTACACGT CTAATTGCTC CATTTTTATG TTTTATGCTG TATATGGTTT
Template            ATTTACACGT CTAATTGCTC CATTTTTATG TTTTATGCTG TATATGGTTT
                           410        420        430        440        450
P0_primer_10        TGATACAGTC AAGAATATAT TCAATAGGAT TGTTTACAGG GTCTTCTCCA
PH1_primer_10       TGATACAGTC AAGAATATAT TCAATAGGAT TGTTTACAGG GTCTTCTCCA
PH2_primer_10       TGATACAGTC AAGAATATAT TCAATAGGAT TGTTTACAGG GTCTTCTCCA
PH3_primer_10       TGATACAGTC AAGAATATAT TCAATAGGAT TGTTTACAGG GTCTTCTCCA
PH4_primer_10       TGATACAGTC AAGAATATAT TCAATAGGAT TGTTTACAGG GTCTTCTCCA
Template            TGATACAGTC AAGAATATAT TCAATAGGAT TGTTTACAGG GTCTTCTCCA
                       456
P0_primer_10        GCCTCA
PH1_primer_10       GCCTCA
PH2_primer_10       GCCTCA
PH3_primer_10       GCCTCA
PH4_primer_10       GCCTCA
Template            GCCTCTAAAGCCAAAC
```

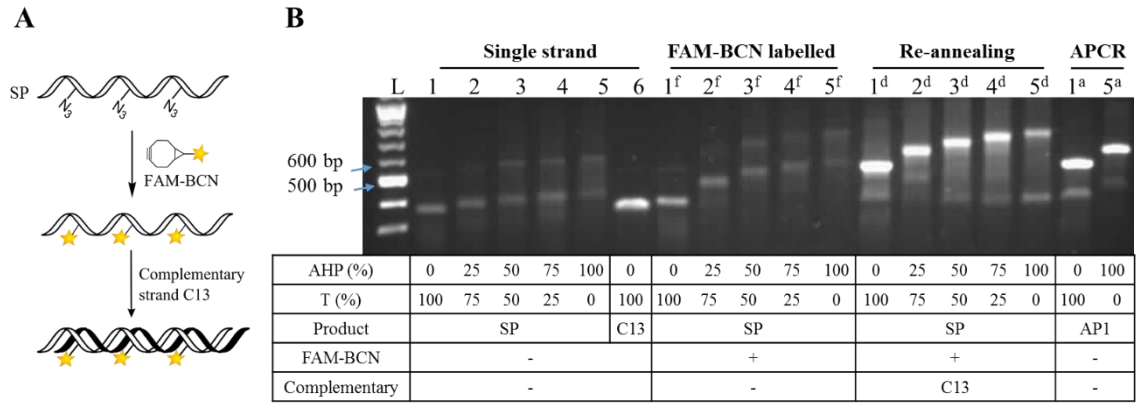
8.8 Agarose gel analysis of λ-exonuclease digestion and re-annealing of AHP-modified ssDNA



AHP (%)	0	25	50	75	100	0	25	50	75	100	0	25	50	75	100
T (%)	100	75	50	25	0	100	75	50	25	0	100	75	50	25	0
Complementary	-					-					T8				

Figure 8.5. **A.** λ-exonuclease digestion of PCR amplicons (81 bp, primer P6 and $P7_p$) with different percentages of AHP dU modifications and re-annealing of the AHP dU single strands (ST) with template T8. Lane L, 25 bp ladder; lane $1/1^s/1^d$, unmodified product; lanes $2/2^s/2^d$ to $5/5^s/5^d$, 25% to 100% AHP-modified products. D = duplex products; T + primer = excess primer annealed to template T8; T = template T8. (2% agarose gel containing SYBR Gold) **B.** Scheme for re-annealing between the AHP-modified single strand (ST, with excess primer) and template T8.

8.9 Agarose gel analysis of FAM-BCN labelled AHP-modified ssDNA (T11)

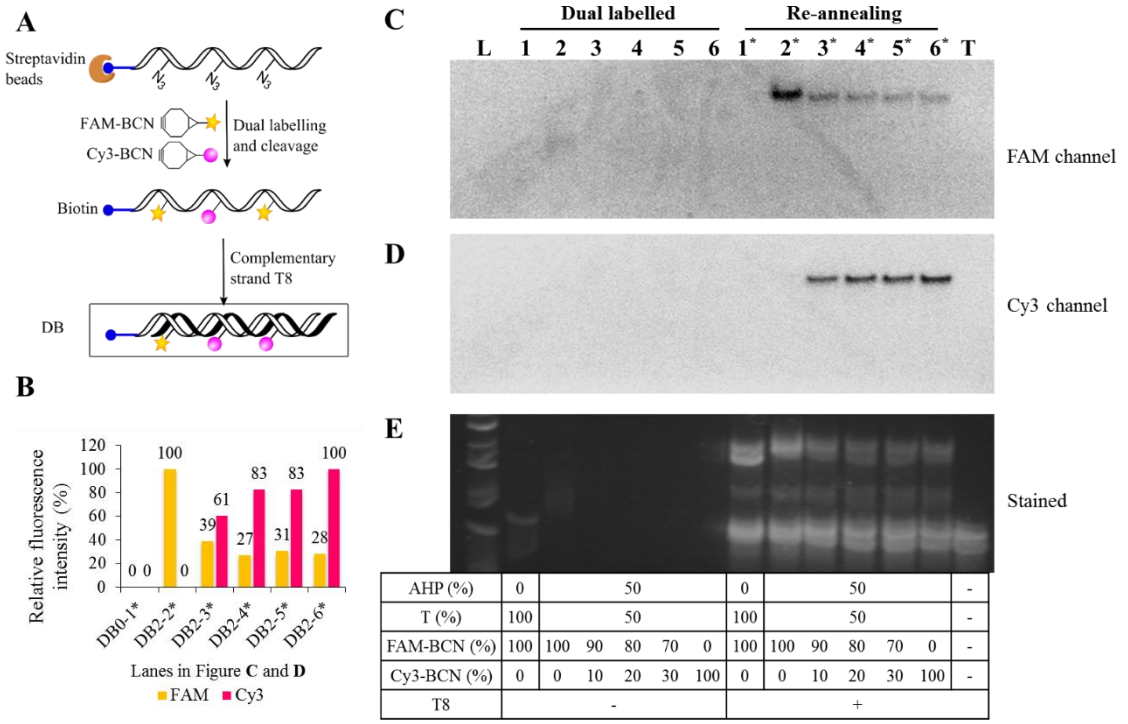


AHP (%)	0	25	50	75	100	0	0	25	50	75	100	0	25	50	75	100	0	100
T (%)	100	75	50	25	0	100	100	75	50	25	0	100	75	50	25	0	100	0
Product	SP					C13	SP					SP					AP1	
FAM-BCN	-						+					+					-	
Complementary	-						-					C13					-	

Figure 8.6. FAM-BCN fluorescent labelling of AHP-modified single strands (SP) and re-annealing with complementary strand C13. **A.** Scheme for labelling and re-annealing. **B.** Agarose gel analysis. Lane L, 100 bp ladder; lanes $1/1^f/1^d/1^a$ and 6, unmodified products; lanes $2/2^f/2^d$ to $5/5^f/5^d/5^a$, 25% to 100% AHP-modified products. (1.5% agarose gel containing ethidium bromide)

8.10 Native PAGE analysis of dual labelled AHP-modified ssDNA (T8) on the streptavidin magnetic beads

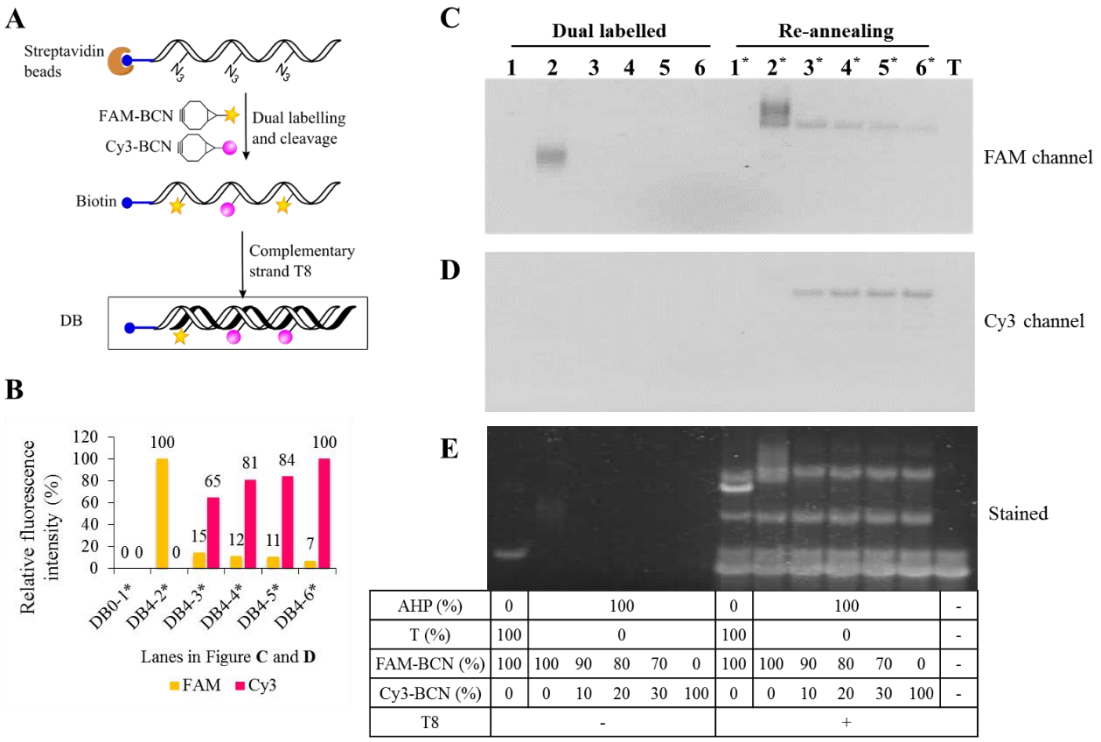
8.10.1 Dual labelling of 50% AHP-modified ssDNA using FAM-BCN and Cy3-BCN



AHP (%)	0	50					0	50					-
T (%)	100	50					100	50					-
FAM-BCN (%)	100	100	90	80	70	0	100	100	90	80	70	0	-
Cy3-BCN (%)	0	0	10	20	30	100	0	0	10	20	30	100	-
T8	-						+						

Figure 8.7. Dual labelling of 50% AHP-modified single strands from streptavidin magnetic separation with different combinations of FAM-BCN (100% to 0%) and Cy3-BCN (0% to 100%). **A.** Scheme for dual labelling and re-annealing. **B.** Relative fluorescence intensity for re-annealed duplex products in FAM and Cy3 channels (lanes 1^* to 6^* in Figure **C** and **D**). Native PAGE analysis visualised before staining by FAM channel (**C**), Cy3 channel (**D**) and after staining with SYBR Gold (**E**). Lane L, 25 bp ladder; lane $1/1^*$, unmodified product, negative control; lanes $2/2^*$ to $6/6^*$, 50% AHP-modified products; lane T, complementary T8. (12% native PAGE)

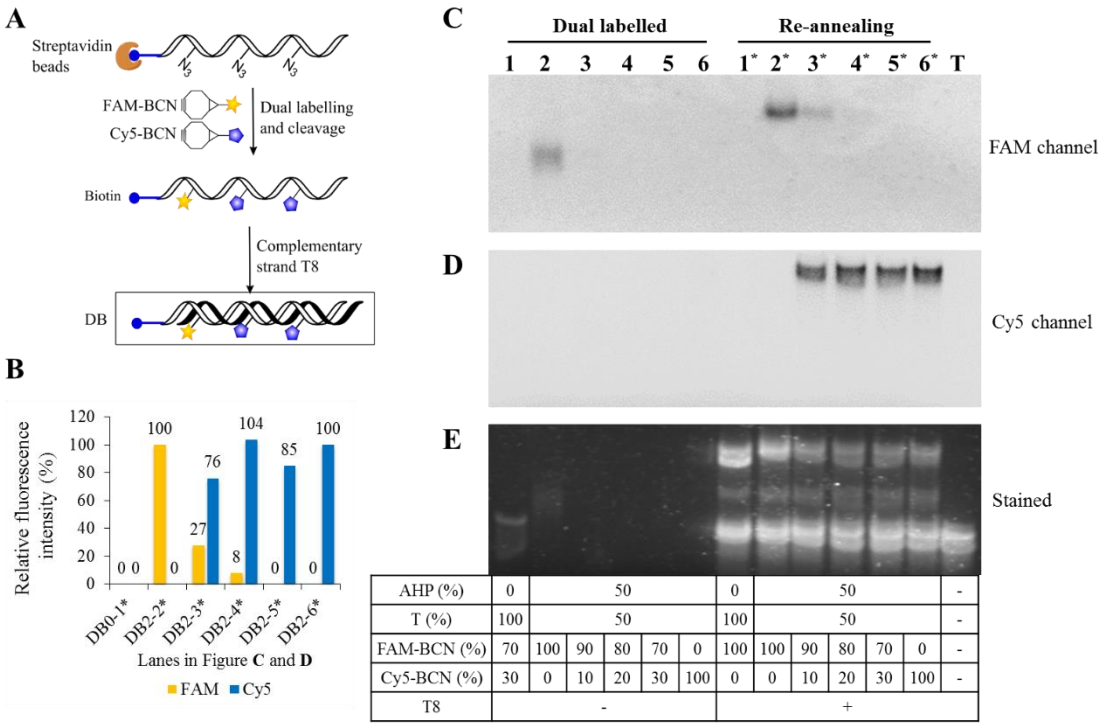
8.10.2 Dual labelling of 100% AHP-modified ssDNA using FAM-BCN and Cy3-BCN



AHP (%)	0	100					0	100					-
T (%)	100	0					100	0					-
FAM-BCN (%)	100	100	90	80	70	0	100	100	90	80	70	0	-
Cy3-BCN (%)	0	0	10	20	30	100	0	0	10	20	30	100	-
T8	-						+						

Figure 8.8. Dual labelling of 100% AHP-modified single strands from streptavidin magnetic separation with different combinations of FAM-BCN (100% to 0%) and Cy3-BCN (0% to 100%). **A.** Scheme for dual labelling and re-annealing. **B.** Relative fluorescence intensity for re-annealed duplex products in FAM and Cy3 channels (lanes 1^* to 6^* in Figure **C** and **D**). Native PAGE analysis visualised before staining by FAM channel (**C**), Cy3 channel (**D**) and after staining with SYBR Gold (**E**). Lane $1/1^*$, unmodified product, negative control; lanes $2/2^*$ to $6/6^*$, 50% AHP-modified products; lane T, complementary T8. (12% native PAGE)

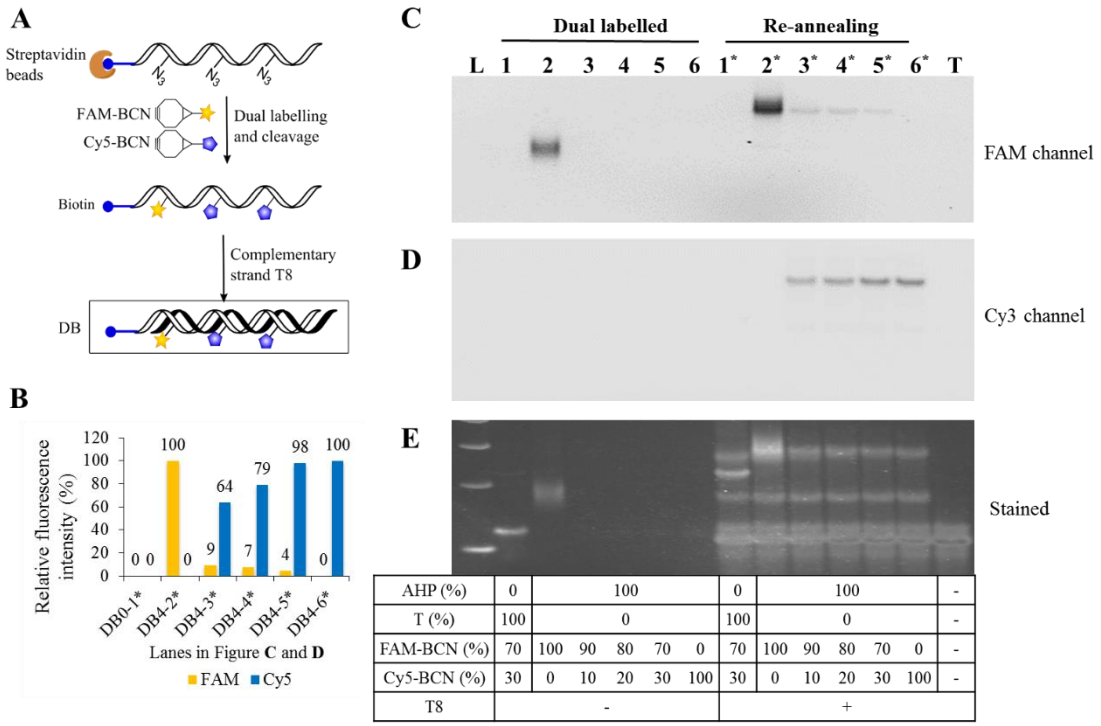
8.10.3 Dual labelling of 50% AHP-modified ssDNA using FAM-BCN and Cy5-BCN



AHP (%)	0	50					0	50					-
T (%)	100	50					100	50					-
FAM-BCN (%)	70	100	90	80	70	0	100	100	90	80	70	0	-
Cy5-BCN (%)	30	0	10	20	30	100	0	0	10	20	30	100	-
T8	-						+						

Figure 8.9. Dual labelling of 50% AHP-modified single strands from streptavidin magnetic separation with different combinations of FAM-BCN (100% to 0%) and Cy5-BCN (0% to 100%). **A.** Scheme for dual labelling and re-annealing. **B.** Relative fluorescence intensity for re-annealed duplex products in FAM and Cy5 channels (lanes 1^* to 6^* in Figure **C** and **D**). Native PAGE analysis visualised before staining by FAM channel (**C**), Cy5 channel (**D**) and after staining with SYBR Gold (**E**). Lane $1/1^*$, unmodified product, negative control; lanes $2/2^*$ to $6/6^*$, 50% AHP-modified products; lane T, complementary T8. (12% native PAGE)

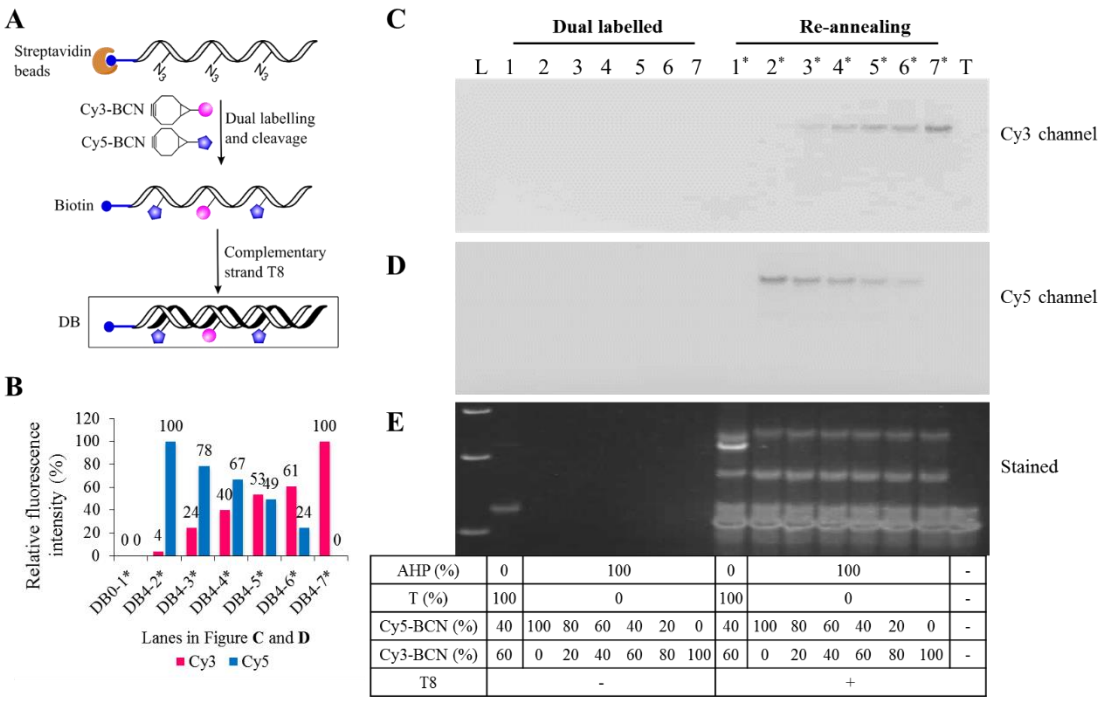
8.10.4 Dual labelling of 100% AHP-modified ssDNA using FAM-BCN and Cy5-BCN



AHP (%)	0	100					0	100					-
T (%)	100	0					100	0					-
FAM-BCN (%)	70	100	90	80	70	0	70	100	90	80	70	0	-
Cy5-BCN (%)	30	0	10	20	30	100	30	0	10	20	30	100	-
T8	-						+						

Figure 8.10. Dual labelling of 100% AHP-modified single strands from streptavidin magnetic separation with different combinations of FAM-BCN (100% to 0%) and Cy5-BCN (0% to 100%). **A.** Scheme for dual labelling and re-annealing. **B.** Relative fluorescence intensity for re-annealed duplex products in FAM and Cy5 channels (lanes 1^* to 6^* in Figure **C** and **D**). Native PAGE analysis visualised before staining by FAM channel (**C**), Cy5 channel (**D**) and after staining with SYBR Gold (**E**). Lane L, 25 bp ladder; lane $1/1^*$, unmodified product, negative control; lanes $2/2^*$ to $6/6^*$, 50% AHP-modified products; lane T, complementary T8. (12% native PAGE)

8.10.5 Dual labelling of 100% AHP-modified ssDNA using Cy3-BCN and Cy5-BCN



AHP (%)	0	100						0	100						-
T (%)	100	0						100	0						-
Cy5-BCN (%)	40	100	80	60	40	20	0	40	100	80	60	40	20	0	-
Cy3-BCN (%)	60	0	20	40	60	80	100	60	0	20	40	60	80	100	-
T8	-							+							

Figure 8.11. Dual labelling of 100% AHP-modified single strands from streptavidin magnetic separation with different combinations of Cy5-BCN (100% to 0%) and Cy3-BCN (0% to 100%). **A.** Scheme for dual labelling and re-annealing. **B.** Relative fluorescence intensity for re-annealed duplex products in Cy3 and Cy5 channels (lanes 1* to 7* in Figure **C** and **D**). Native PAGE analysis visualised before staining by Cy3 channel (**C**), Cy5 channel (**D**) and after staining with SYBR Gold (**E**). Lane L, 25 bp ladder; lane 1/1*, unmodified product, negative control; lanes 2/2* to 7/7*, 50% AHP-modified products; lane T, complementary T8. (12% native PAGE)

8.11 List of publications

1. Ren, X.; Gerowska, M.; El-Sagheer, A. H.; Brown, T., Enzymatic incorporation and fluorescent labelling of cyclooctyne-modified deoxyuridine triphosphates in DNA. *Bioorg. Med. Chem.* **2014**, *22*, 4384-4390.
2. Ren, X.; El-Sagheer, A. H.; Brown, T., Azide and trans-cyclooctene dUTPs: incorporation into DNA probes and fluorescent click-labelling. *Analyst* **2015**, *140*, 2671-2678.

Bioorganic & Medicinal Chemistry 22 (2014) 4384–4390

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Enzymatic incorporation and fluorescent labelling of cyclooctyne-modified deoxyuridine triphosphates in DNA

Xiaomei Ren [a], Marta Gerowska [b], Afaf H. El-Sagheer [a,c], Tom Brown [a,*]

[a] Department of Chemistry, University of Oxford, Chemistry Research Laboratory, Oxford OX1 3TA, UK
[b] School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, UK
[c] Chemistry Branch, Department of Science and Mathematics, Faculty of Petroleum and Mining Engineering, Suez University, Suez 43721, Egypt

ARTICLE INFO

Article history:
Received 8 March 2014
Revised 16 May 2014
Accepted 22 May 2014
Available online 5 June 2014

Keywords:
Copper free click
DNA fluorescent labelling
Cyclooctyne deoxyuridine triphosphates
BCN
DIBO

ABSTRACT

The amino group of 5-aminopropargyl-2′-deoxyuridine-5′-triphosphate was labelled with dibenzocyclooctyne (DIBO) and two derivatives of bicyclo [6.1.0] non-4-yne (BCN) with short and long linkers to produce three different cycloalkyne-modified deoxyuridine triphosphates. BCN was successfully incorporated into DNA at multiple sites by enzyme-mediated primer extension and the polymerase chain reaction (PCR). Efficient fluorescent labelling of the BCN-DNA and DIBO-DNA with Cy3-azide was demonstrated.

1. Introduction

Chemically modified DNA is used extensively in biology, diagnostics, and nanotechnology. Recently the field of nucleic acid chemistry underwent a revolution when click chemistry[1] and in particular the Cu^{I} catalyzed [3+2] azide-alkyne cycloaddition reaction (CuAAC reaction)[2,3] was discovered. This reaction has the considerable advantage in the context of nucleic acids that the reacting alkyne and azide groups are stable in aqueous buffer and do not cross-react with other functional groups. The CuAAC reaction has been utilized extensively[4] as a method of joining together single strands of DNA,[5] cross-linking complementary strands,[6] cyclizing single and double strands,[5,7] labelling oligonucleotides with reporter groups,[8–10] immobilizing DNA on surfaces,[11] constructing nanowires from DNA templates,[12] building DNA nanostructures,[13] producing analogues of DNA with modified nucleobases[14,15] and backbones,[16–21] synthesizing chemically modified RNA constructs,[22,23] monitoring DNA synthesis in vivo[24] and creating chemically modified PCR templates.[8,25,26]

However, despite the spectacular success of the CuAAC reaction in the nucleic acid field, Cu^{I} has one particular limitation. It is cytotoxic, and is therefore not compatible with in vivo applications of nucleic acids, or for use in experiments that involve living cells.[27–30] Unfortunately the uncatalyzed copper-free DNA-templated AAC Huisgen reaction is very slow[31] unless highly activated terminal alkynes are used, but these are unstable in aqueous media so they cannot be used in biology. For carbohydrates and other biomolecules the issue of biocompatibility has been elegantly solved by the development of the ring strain-promoted azide-alkyne [3+2] cycloaddition reaction (SPAAC) (Fig. 1A).[32–35] This involves the uncatalyzed reaction between an azide and a strained alkyne, normally a cyclooctyne derivative, and is promoted by distortion of the sp-orbitals of the alkyne. Reaction to form a triazole (sp^2 hybridization) allows release of energy accompanied by the formation of a stable product. Analogues of cyclooctyne have been synthesized which are modified to increase reactivity,[32,33] and the SPAAC reaction using these has recently been adopted for nucleic acids, both in oligonucleotide labelling[36,37] and copper-free DNA strand ligation.[38] Phosphoramidite monomers have been synthesized for internal[38] and terminal labelling[39] of DNA with cyclooctynes such as DIBO (dibenzocyclooctyne, Fig. 1B).[39,40] For applications in which diastereoisomers of the product of the SPAAC reaction might be undesirable, achiral bicycle [6.1.0] nonyne (BCN, Fig. 1C)[41,42] has been employed.[43,44] Due to its symmetry, BCN produces only enantiomeric products.

Enzymatic incorporation of terminal alkyne and azide modifications into DNA by polymerase enzymes is an efficient and widely employed method of DNA labelling.[10,25,45–48] Clearly it would be desirable if this methodology could be used to incorporate reactive cyclooctyne derivatives for applications such as the synthesis of

* Corresponding author. Tel.: +44 (0)0865 275413.
E-mail address: tom.brown@chem.ox.ac.uk (T. Brown).

http://dx.doi.org/10.1016/j.bmc.2014.05.050
0968-0896/

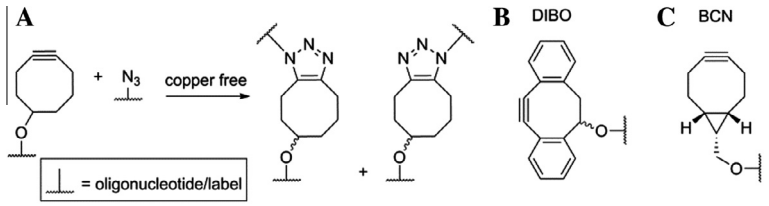


Figure 1. (A) Oligonucleotide ligation and labelling by the ring-strain promoted alkyne-azide cycloaddition reaction (SPAAC). (B) DIBO, (C) BCN.

fluorescent probes. Here we report the synthesis of three new DIBO and BCN-modified deoxyuridine triphosphates. We describe their incorporation into DNA by enzymatic primer extension and polymerase chain reactions, and demonstrate efficient fluorescent labelling of the BCN and DIBO-modified DNA with Cy3 hexylazide via the SPAAC reaction (Fig. 2).

2. Results and discussion

2.1. Preparation of cycloalkyne modified deoxyuridine triphosphates

Deoxyuridine triphosphate is the simplest of the natural dNTPs to modify chemically. The C5-position of uracil is remote from the H-bonding residues that are involved in Watson–Crick base-pairing, and functional groups attached at this site protrude into space in the major groove of the DNA duplex, without disruption of the normal B-conformation. Hence 5-modified dUTPs (dMTPs) are likely to be compatible with enzymatic incorporation by DNA polymerases, provided that the modification does not inhibit polymerase binding.[49–51] The aminopropargyl linkage is a particularly useful functionality because the ethynyl moiety is known to be compatible with DNA polymerases, and the primary amine is a suitable attachment point for carboxylic acid derivatives of reporter groups, thereby enabling the synthesis of various functionalized dMTPs from a common intermediate.[49,52,53] In fact aminopropargyl dUTP is an interesting dMTP in its own right, as it has a short linker in the position normally occupied by the methyl group of dTTP, so it is likely to be a good substrate for various DNA polymerase enzymes.[54] We chose to synthesize three different cyclooctyne dUTP derivatives; a DIBO derivative which has a short linker between the cyclooctyne and the uracil base, and two different BCN dUTP derivatives with short and long linkers (BCN-I and BCN-II, Scheme 1). We expected that the presence of the bulky DIBO ring might lead to less efficient incorporation by DNA polymerases than BCN, and that placing the BCN ring further from the uracil base and away from the enzyme active site would lead to more efficient enzymatic incorporation. In order to synthesize these cyclooctyne deoxyuridine triphosphates we first required the common intermediate, aminopropargyl dUTP. The Sonogashira reaction was used to attach aminoprop-2-yne to 5-iododeoxyuridine[53,55] which was then converted to its 5′-triphosphate by the Yoshikawa procedure.[56] Phosphorus oxychloride ($POCl_3$) in trimethylphosphate was employed to generate the 5′-dichlorophosphate, and the reaction mixture was treated with a tetrabutylammonium salt of pyrophosphate to produce the required triphosphate. In order to minimize the production of 3′,5′-diphosphate by-products, the reaction was conducted at −15 °C. 5-Aminopropargyl-dUTP was then labelled with the active esters of DIBO, BCN-I and BCN-II to yield three activated cyclooctyne-modified triphosphates (dMTPs) (Scheme 1).

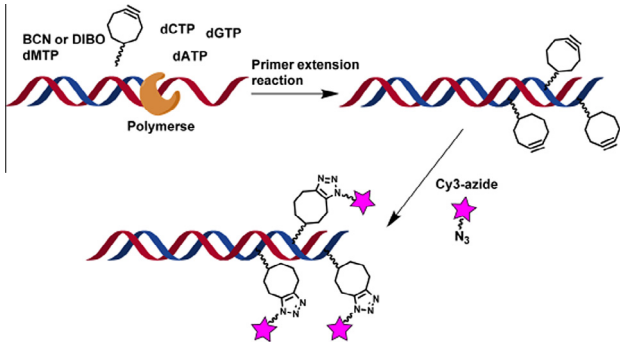
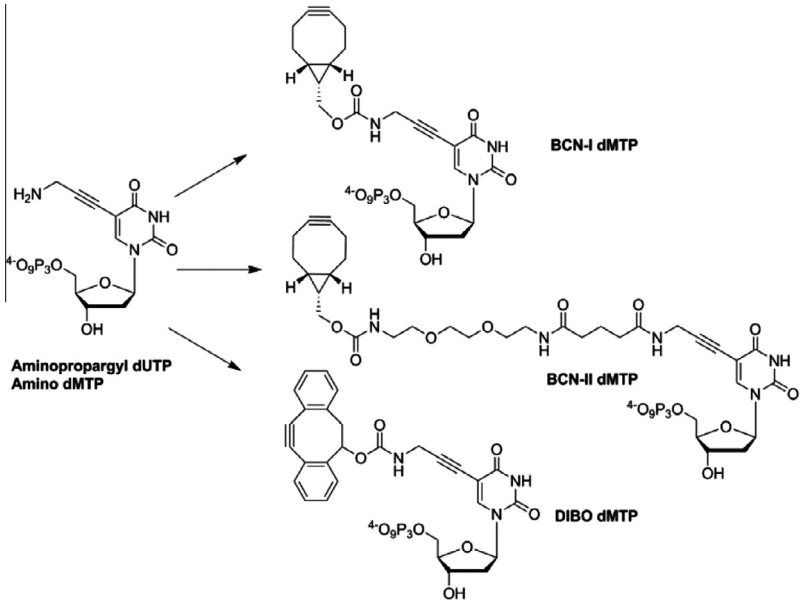


Figure 2. Enzyme-mediated primer extension reaction using DIBO and BCN-modified triphosphates followed by labelling with Cy3-azide.

2.2. Primer extension reactions and PCR

Primer extension is a biochemical approach to the synthesis of chemically modified DNA of defined sequence and substitution pattern. Primer extension reactions utilizing amino, DIBO, BCN-I and BCN-II dMTPs were carried out with family A (Klenow large fragment) and family B polymerases (Therminator II, Gotaq and KOD). The sequences of primers and templates are shown in Table 1. The 5′-end of the primer was labelled with 6-carboxyfluorescein (FAM) for ease of visualization on analytical polyacrylamide gels. Two templates were used with different numbers and arrangements of adenine bases after the starting point of linear copying. Template (T1) begins with one A, and contains just three adenines for replication while the 'highly demanding' template (T2) starts with four consecutive adenines (A) and contains six adenine bases to be replicated. In each case the primer must be extended by 11 bases for complete replication to give a 29-mer product. Alternatively the primer can be extended using only the modified triphosphate to give shorter products, a 19-mer with template T1 and a 22-mer with template T2. The four natural dNTPs (dATP, dCTP, dGTP, dTTP) were used in control experiments.

Polyacrylamide gel electrophoresis (PAGE) and mass spectrometry confirmed that KOD, Therminator II, Gotaq and Klenow large fragment polymerases successfully incorporated all four dMTPs into the template T1 (Fig. 3 and SI). All the copies of the templates containing dMTPs (Fig. 3A, lanes 1 and 2) exhibited lower electrophoretic mobility compared to the natural counterparts (Fig. 3A, lanes 4 and 5). This is due to the extra bulk of the modified uracil bases, particularly in the case of BCN-II dMTP with a long linker which showed a large difference in mobility compared to the natural counterparts. The polymerases incorporated one or two additional nucleotides at the 3′-end of the linear copy of the template when using only dTTP or dMTPs (Fig. 3, lane 4 and SI). When modified dNTPs are incorporated into DNA enzymatically the modified dNTP is usually doped with its natural counterpart to avoid premature termination of replication. We therefore explored the effect of mixing the individual dMTPs with dTTP (Fig. 3A, lane 3) and found that the modified triphosphates were all able to compete effectively, producing fully extended primers with a mixture of dTTP and dMTP opposite to the dA sites in the template. This is evident from the multiple or broader bands of the fully extended mixed



Scheme 1. Labelling of aminopropargyl-dUTP with DIBO, BCN-I and BCN-II active esters to produce 5-cyclooctyne-modified dUTPs. Reagents: DMF and triethylammonium bicarbonate buffer, 4 h, 55 °C.

Table 1
Oligonucleotide sequences used in primer extension, reverse transcription and PCR

Code	Sequences (**5′**–**3′**)	Mass calcd	Mass found
T1	C**A**GTC**A**CTGT**A**CTGCCGACACACATAACC (DNA template)	8776	8775
T2	C**A**GTC**A**C**AAAA**CTGCCGACACACATAACC (DNA template)	8778	8777
P3	FAM-GGTTATGTGTGTCGGCAG (primer)	6138	6138
T4	C**A**GUC**A**CUGU**A**CUGCCGACACACAUAACC (RNA template)	9169	9169
P5	GCATTCGAGCAACGTAAG (PCR primer)	6548	6548
P6	GGTTATGTGTGTCGGCAG (PCR primer)	7040	7040
T7	FAM-GG**TTAT**GT**G**T**G**TCGGC**A**G**TATT**GT**C**A**GT**G**T**G**A**ATT**CC**A**G**A**G**T**G**T**G**A**G**A**TT**G**T**G**T**G**C**T**GGCG**ATCTTA**CG**TT**G**C**T**CG**AAT**GC (PCR template)		

FAM is 6-carboxyfluorescein. The RNA template has the equivalent sequence to the DNA template T1. The bold A (A/T for PCR) shows the site to incorporate the modified triphosphates.

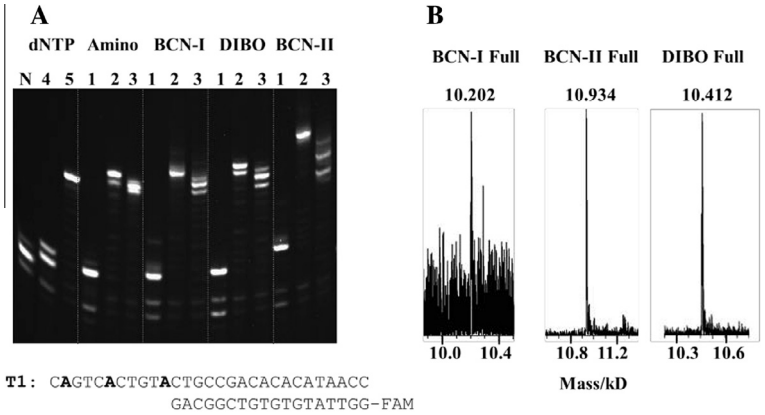


Figure 3. Primer extension reactions using Klenow polymerase (1 unit) with P3 (66 pmol) and T1 (132 pmol). 3.2 nmol of dMTPs or dNTPs for each triphosphate were used unless otherwise stated. (A) Lane N, P3 + T1 without triphosphates; lane 1, dMTP; lane 2, dMTP + dATP + dCTP + dGTP; lane 3, dMTP (2.1 nmol) + dTTP (1.1 nmol) + dATP + dCTP + dGTP; lane 4, dTTP; lane 5, dTTP + dATP + dCTP + dGTP (20% PAGE gel). (B) Mass spectra of fully extended primers: BCN-I Full is produced from BCN-I dMTP + dATP + dCTP + dGTP (calcd 10202, found 10202); BCN-II Full is produced from BCN-II dMTP + dATP + dCTP + dGTP (calcd 10934, found 10934); DIBO Full is produced from DIBO dMTP + dATP + dCTP + dGTP (calcd mass 10412, found 10412).

products of dMTP with dTTP, dATP, dCTP, dGTP (Fig. 3A, lane 3) compared to the control dNTP (lane 5).

Next template T2 with four consecutive adenines was used to evaluate the incorporation of the various dMTPs in a highly demanding case. Encouragingly dMTP containing short chain BCN-I was successfully incorporated by KOD, Therminator II and Klenow polymerases (Fig. 4). The dMTP containing long chain BCN-II gave fully and partially extended products, whereas the more bulky DIBO did not produce a fully extended product with any of the polymerases used (see SI). It is noteworthy that the most

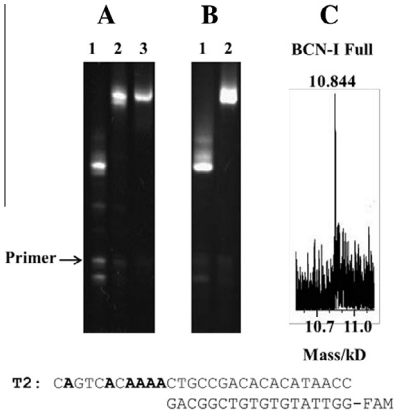


Figure 4. Primer extension reactions with P3 (66 pmol), T2 (132 pmol), BCN-I dMTP (3.2 nmol for each triphosphate). (A) Lane 1, BCN-I dMTP using Klenow polymerase (2 units); lane 2, dMTP + dATP + dCTP + dGTP using Klenow polymerase (2 units); lane 3, dMTP + dATP + dCTP + dGTP using Therminator II polymerase (1 unit). (20% PAGE gel) (B) Reactions using KOD polymerase (2 units, 0.4 mM $MgCl_2$). Lane 1, BCN-I dMTP; lane 2, BCN-I dMTP + dATP + dCTP + dGTP. (20% PAGE gel) (C) Mass spectrum of fully extended primer (BCN-I Full) from B-lane 2 (calcd mass 10845, found 10844).

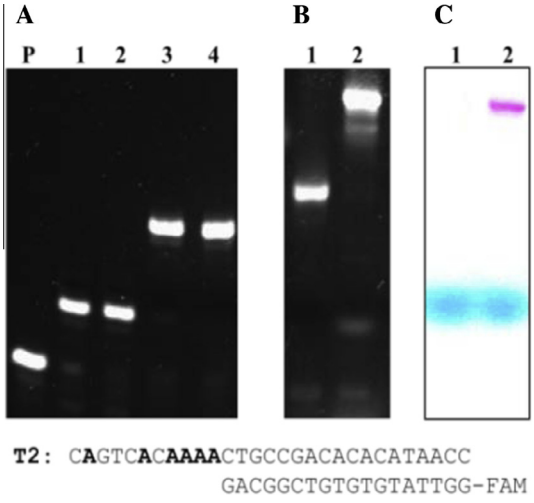


Figure 6. Cy3-azide labelling of primer extension products from P3, T2 and BCN-I dMTP + dATP + dCTP + dGTP using KOD polymerase (1 unit, 1.0 mM $MgCl_2$). (A) Lane P, P3 + T2; lane 1, dTTP; lane 2, lane 1 with Cy3 azide; lane 3, dATP + dCTP + dGTP + dTTP; lane 4, lane 3 with Cy3-azide. (B) Lane 1, BCN-I dMTP + dATP + dCTP + dGTP(non-labelled); lane 2, lane 1 labelled with Cy3-azide. Pictures were taken using transilluminator light and UV filter. (C) Picture of gel B was taken using a digital camera. (20% PAGE gel).

suitable cyclooctyne dUTP in this study is the one with the smallest mass attached to the 5-position of the uracil base (BCN-I), even though this has a much shorter linker than BCN-II.

2.3. Fluorescent labelling

Cy3 is commonly used in the synthesis of fluorescent DNA probes.[57] In this context it has a number of favourable properties which make it suitable for the multiple labelling of DNA strands. It is very bright, sterically undemanding and not excessively hydrophobic. After primer extension with template T1 and ethanol precipitation to remove unincorporated triphosphates, the BCN-I and DIBO-functionalized linear extension products were labelled efficiently with Cy3 hexylazide via the SPAAC reaction (Figs. 5 and 6) (see SI for the structure and synthesis of the Cy3 hexylazide). Natural dTTP and dNTP were used as negative controls. The labelled products exhibited high fluorescence and much lower electrophoretic mobility than the unlabelled controls. This confirms that the BCN and DIBO moieties are stable to the conditions of their incorporation and labelling.

Efficient incorporation of cyclooctynes into DNA during PCR amplification and subsequent fluorescence labelling by the SPAAC

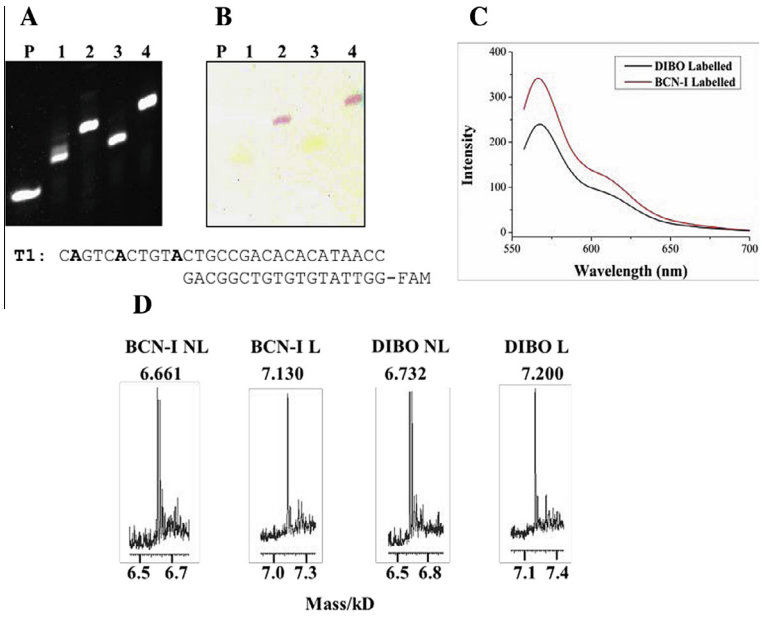


Figure 5. Cy3-azide labelling of primer extension products from P3, T1 and dMTPs using Gotaq polymerase (1 unit). Lane P, P3; lane 1, BCN-I dMTP (non-labelled); lane 2, BCN-I dMTP labelled with Cy3-azide; lane 3, DIBO dMTP (non-labelled); lane 4, DIBO dMTP labelled with Cy3-azide. (20% PAGE gel) (A) Picture was taken using transilluminator light and UV filter. (B) Picture was taken using normal digital camera. (C) Fluorescent spectra of lane 2 (BCN-I labelled) and lane 4 (DIBO labelled) products (Excited at 540 nm on PE LS50B fluorimeter). (D) Mass spectra of one base extended primers, non-labelled product (NL) and Cy3-azide labelled product (L). BCN-I NL (lane 1 in PAGE): calcd 6657 [M], found 6661 [M] and 6683 [M+Na]; BCN-I L (lane 2 in PAGE): calc. 7125, found 7130; DIBO NL (lane 3 in PAGE), calc. 6727 [M], found 6732 [M] and 6754 [M+Na]; DIBO L (lane 4 in PAGE), calc.7195, found 7200.

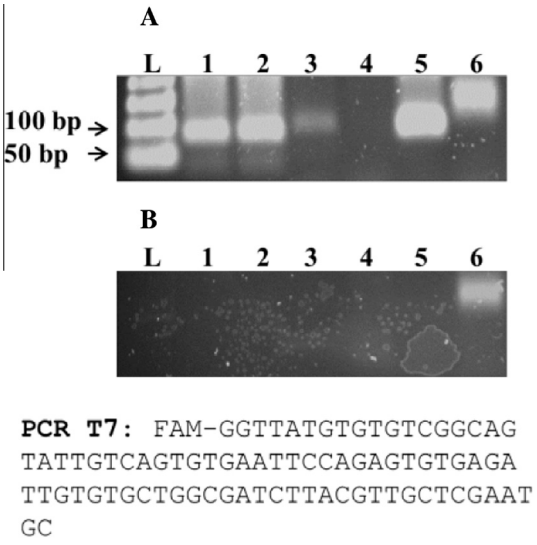


Figure 7. Cy3-azide labelling of polymerase chain reaction product from P5, P6, T7 and BCN-I dMTP with dTTP + dATP + dGTP + dCTP using KOD polymerase (0.5 unit, 1.0 mM $MgCl_2$). Unless otherwise stated 4 nmol of each triphosphate was used. Lane L, 50 bp ladder; lane 1, dTTP + dATP + dGTP + dCTP (non-labelled); lane 2, lane 1 with Cy3 azide; lane 3, BCN-I dMTP (2 nmol) + dTTP (2 nmol) + -dATP + dGTP + dCTP (non-labelled); lane 4, lane 3 labelled with Cy3-azide; lane 5, BCN-I dMTP (1 nmol) + dTTP (3 nmol) + dATP + dGTP + dCTP (non-labelled); lane 6, lane 5 labelled with Cy3-azide. (A) Gel after staining with ethidium bromide. (B) Gel before staining. (2% agarose gel).

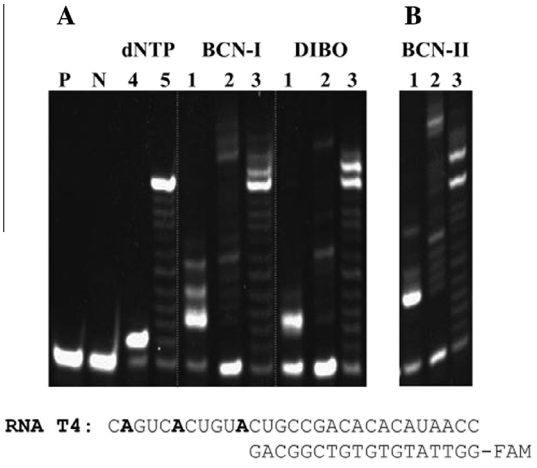


Figure 8. Reverse transcription reactions with P3 (66 pmol), RNA template T4 (132 pmol), dMTPs or dNTPs (3.2 nmol for each triphosphate unless otherwise stated) using M-MuLV (RNase H^-) reverse transcriptase (100 units). (A) The reaction time was 4 h. Lane P, P3 + T4 without triphosphate; lane N, dATP + dCTP + dGTP; lane 1, dMTP; lane 2, dMTP + dATP + dCTP + dGTP; lane 3, dMTP (1.6 nmol) + dTTP (1.6 nmol) + dATP + dCTP + dGTP; lane 4, dTTP; lane 5, dATP + dCTP + dGTP + dTTP. (B) The reaction time was 1 h. Lane 1, dMTP; lane 2, dMTP + dATP + dCTP + dGTP; lane 3, dMTP (2.1 nmol) + dTTP (1.1 nmol) + -dATP + dCTP + dGTP. (20% PAGE gel).

reaction would enable the synthesis of long fluorescent probes in practical quantities. To this end PCR was carried out with BCN-I dMTP using KOD polymerase and 81-mer DNA template T7. The modified triphosphate was successfully incorporated when it was mixed with the four natural triphosphates, using a 1:3 ratio of dMTP to dTTP. Labelling of the PCR product created from this mixture with Cy3 hexylazide gave a fluorescent band (Fig. 7.B, lane 6). As expected, subsequent staining with ethidium bromide revealed slower migration of this band compared with the unlabelled amplicon (Fig. 7A, lanes 5 and 6).

2.4. Reverse transcription

Reverse transcriptase (RT) enzymes can be used to copy DNA and RNA templates in order to prepare chemically modified complementary DNA (cDNA). This methodology can efficiently produce functionalized complementary DNA (cDNA) from RNA that has been isolated from natural sources. In the current work we chose to evaluate the cyclooctyne-modified dMTPs with Moloney Murine Leukemia Virus reverse transcriptase (M-MuLV RT, RNase H^-) which conveniently lacks the catalytic activity necessary to cleave RNA strands. RNA template T4 of the equivalent sequence to the previously used DNA template with dispersed adenines was employed. BCN-II dMTP was incorporated efficiently within 1 h while BCN-I dMTP needed 4 h. Even with this extended period the DIBO reaction was incomplete (Fig. 8, lane 1). The above reverse transcription reactions would clearly benefit from further optimization of the reaction conditions.

3. Conclusions

The strain-promoted alkyne-azide cycloaddition reaction is a valuable tool for DNA labelling, and in this context the efficient incorporation of reactive cyclooctyne derivatives into DNA by enzymatic methods is an important objective. To this end we have synthesized one DIBO and two BCN dUTP analogues from a common aminopropargyl dUTP intermediate and incorporated them into DNA by linear extension using two different templates and a variety of DNA polymerases. All three cyclooctynes (DIBO, BCN-I and BCN-II) were incorporated successfully against template with dispersed adenine bases, whereas BCN was the most successful dMTP when using a template with four consecutive adenines. Labelling of the modified extension products with Cy3 hexylazide via the copper-free SPAAC click reaction demonstrated the efficient incorporation and stability of these modified triphosphates to the enzymatic reaction conditions. The acceptance of cyclooctyne-modified triphosphates as substrates for various DNA polymerases in linear extension and PCR is encouraging in the context of future in vitro and in vivo applications.

4. Experimental section

4.1. General method for oligonucleotide synthesis and purification

Standard DNA phosphoramidites, solid supports, and additional reagents were purchased from Link Technologies and Applied Biosystems Ltd. All oligonucleotides were synthesized on an Applied Biosystems 394 automated DNA/RNA synthesizer using a standard 1.0 μmol scale phosphoramidite cycle of acid-catalyzed detritylation, coupling, capping, and iodine oxidation. Stepwise coupling efficiencies and overall yields were determined by the automated trityl cation conductivity monitoring facility and in all cases were >98.0%. All β-cyanoethyl phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use. The coupling time for normal A, G, C, and T monomers was 40 s. Cleavage of oligonucleotides from the solid support and deprotection was achieved by exposure to concentrated aqueous ammonia solution for 60 min at room temperature followed by heating in a sealed tube for 5 h at 55 °C.

The building blocks for the RNA analogues were prepared using 2′-TBS protected RNA phosphoramidite monomers with *t*-butylphenoxyacetyl protection of the A, G and C nucleobases and unprotected U (Sigma–Aldrich). A solution of 0.3 M benzylthiotetrazole in acetonitrile (Link Technologies) was used as the coupling agent, *t*-butylphenoxyacetic anhydride was employed as the capping

agent and 0.1 M iodine as the oxidizing agent (Sigma–Aldrich). All RNA phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use, and the coupling time for all monomers was 10 min. Stepwise coupling efficiencies were determined by automated trityl cation conductivity monitoring and in all cases were >96%. Cleavage of oligonucleotides from the solid support and deprotection were achieved by exposure to concentrated aqueous ammonia/ethanol (3/1 v/v) for 2 h at room temperature followed by heating in a sealed tube for 45 min at 55 °C. After cleavage from the solid support and deprotection of the nucleobases and phosphotriesters, RNA oligonucleotides were concentrated to a small volume in vacuo, transferred to 15 mL plastic tubes and freeze dried. The residues were dissolved in DMSO (300 μL) and triethylamine trihydrofluoride (300 μL) was added after which the reaction mixtures were kept at 65 °C for 2.5 h. Sodium acetate (3 M, 50 μL) and butanol (3 mL) were added with vortexing and the samples were kept at −80 °C for 30 min then centrifuged at 4 °C at 13,000 rpm for 10 min. The supernatant was decanted and the precipitate was washed twice with ethanol (0.75 mL) then dried under vacuum.

The fully deprotected oligonucleotides (DNA or RNA) were purified by reversed-phase HPLC on a Gilson system using a Luna 10 μ C8 100 Å pore Phenomenex 10 × 250 mm column with a gradient of acetonitrile in ammonium acetate, in case of DNA, (0% to 50% buffer B over 20 min, flow rate 4 mL/min), (buffer A: 0.1 M ammonium acetate, pH 7.0, buffer B: 0.1 M ammonium acetate, pH 7.0, with 50% acetonitrile). Elution was monitored by UV absorption at 295 nm. After HPLC purification, oligonucleotides were desalted using NAP-10 columns (GE Healthcare). For HPLC purification of RNA, triethylammonium bicarbonate buffer was used (buffer A: 0.1 M triethylammonium bicarbonate, pH 7.5, buffer B: 0.1 M triethylammonium bicarbonate, pH 7.5, with 50% acetonitrile). The fractions from HPLC were evaporated without need for additional desalting.

All oligonucleotides were characterized by negative-mode electrospray HPLC-mass spectrometry, using a Brueker Daltronics micro-TOF mass spectrometer and an Acquity UPLC system with a BEH C18 1.7 μm column (Waters). A gradient of acetonitrile in triethylammonium acetate (TEAA) and hexafluoroisopropanol (HFIP) was employed, increasing from 5% to 40% buffer B over 14 min, with a flow rate of 0.1 mL/min (buffer A: 10 mM TEAA, 100 mM HFIP in H_2O; buffer B: 20 mM TEAA in CH_3CN). Raw data were processed and deconvoluted using the Data Analysis function of the Bruker Daltronics Compass™ 1.3 software package.

4.2. Primer extension experiments

In a 20 μL reaction, 66 pmol of FAM labelled DNA primer (P3), 132 pmol of template (T1 or T2) and 3.2 nmol of each dMTP or dNTP were mixed with the polymerase enzyme (1 or 2 units, see figure legends) and 1× buffer (supplied with the enzymes). In the case of KOD polymerase, 0.4 or 1 mM $MgCl_2$ was added separately to the reaction buffer. The reaction mixtures were heated for 1 or 1.5 h (KOD, Therminator II and GoTaq were at 72 °C while Klenow was at 37 °C) and 20 μL formamide was added before analysis by 20% denaturing polyacrylamide gel electrophoresis under a constant 20 W. For the mass spectrometry analysis, 40 μL reactions were carried out, desalted on a NAP-10 column and concentrated by freeze-drying. KOD was purchased from Merck Millipore, Therminator II and Klenow large fragment from New England Bio and Gotaq from Promega.

4.3. Polymerase chain reaction

Samples containing 10 pmol of two primers (P5, P6), 50 pg of template (T7) and 4 nmol dMTP or dNTP in total were added followed by polymerase, 10X KOD polymerase buffer (pH = 8) and 25 mM $MgCl_2$. The final reaction volume was 20 μL with 1× reaction buffer and 1 mM $MgCl_2$. Amplification was performed using the following procedure: an initial denaturing at 95 °C for 1 min, followed by 25 cycles of denaturing at 95 °C for 15 s, primer annealing at 54 °C for 20 s, and extension at 72 °C for 30 s, then further extension at 72 °C for 5 min. The samples were analyzed by 2% agarose gel electrophoresis in 1× TBE buffer (126 V).

4.4. Reverse transcription

Samples containing 66 pmol of FAM labelled DNA primer (P3), 132 pmol of RNA template (T4) and 3.2 nmol dMTP or dNTP were mixed with M-MuLV (RNase H^-) reverse transcriptase (100 units), 5× buffer and 1× dithiothretitol (DTT, supplied with the enzyme). The final reaction volume was 20 μL. The reaction mixtures were heated at 42 °C for 1 h or 4 h, 20 μL of formamide was added and samples were analyzed by 20% denaturing polyacrylamide gel electrophoresis under constant 20 W.

4.5. Fluorescent labelling

Primer extension reactions or PCR (2 × 20 μL + 10 μL water) were prepared as explained above in Section 4.2 or 4.3 using BCN-1 dMTP, DIBO dMTP or dNTP negative control. DNA was precipitated by mixing with 5 μL 3 M sodium acetate (pH 5.3) followed by 150 μL ethanol and was then left on dry ice for 10 min and at −20 °C for 20 min. This was followed by centrifugation at 4 °C for 30 min (13,000 rpm). The precipitate was re-dissolved in 40 μL 1× Gotaq Green buffer and 20 μL was reacted with 1.6 μL Cy3 hexylazide (10 mM DMSO solution) for 1 h. for primer extensions, 20 μL of formamide was added to the labelled DNA and the unlabelled primer extension reactions (20 μL) followed by analysis on a 20% PAGE gel at 20 W. for PCR, after the labelling reactions, a second ethanol precipitation was carried out using 2 μL of 3 M sodium acetate (pH 5.3) and 60 μL of ethanol. The samples were analyzed by 2% agarose gel electrophoresis in 1× TBE buffer (126 V) For mass spectrometry and fluorescent spectra analysis, 40 μL of non-labelled and labelled reactions were carried out as above followed by desalting on NAP-10 columns.

Acknowledgments

T.B. and A.H.E.-S. are grateful to the UK BBSRC for funding via the sLOLA Grant BB/J001694/1: 'Extending the Boundaries of Nucleic Acid Chemistry'.

Supplementary data

Supplementary data (chemical synthesis, primer extension experiments and fluorescent labelling reactions) associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmc.2014.05.050.

References and notes

1. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Edit.* **2001**, *40*, 2004.
2. Tornoe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057.
3. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Edit.* **2002**, *41*, 2596.
4. El-Sagheer, A. H.; Brown, T. *Chem. Soc. Rev.* **2010**, *39*, 1388.
5. Kumar, R.; El-Sagheer, A. H.; Tumpane, J.; Lincoln, P.; Wilhelmsson, L. M.; Brown, T. *J. Am. Chem. Soc.* **2007**, *129*, 6859.
6. Kocalka, P.; El-Sagheer, A. H.; Brown, T. *ChemBioChem* **2008**, *9*, 1280.
7. El-Sagheer, A. H.; Kumar, R.; Findlow, S.; Werner, J. M.; Lane, A. N.; Brown, T. *ChemBioChem* **2008**, *9*, 50.
8. Gierlich, J.; Burley, G. A.; Gramlich, P. M. E.; Hammond, D. M.; Carell, T. *Org. Lett.* **2006**, *8*, 3639.

9. Motea, E. A.; Lee, I.; Berdis, A. J. *Nucleic Acids Res.* **2012**, *40*, 2357.
10. Wenge, U.; Ehrenschwender, T.; Wagenknecht, H.-A. *Bioconjugate Chem.* **2013**, *24*, 301.
11. Rozkiewicz, D. I.; Gierlich, J.; Burley, G. A.; Gutsmiedl, K.; Carell, T.; Ravoo, B. J.; Reinhoudt, D. N. *ChemBioChem* **2007**, *8*, 1997.
12. Fischler, M.; Simon, U.; Nir, H.; Eichen, Y.; Burley, G. A.; Gierlich, J.; Gramlich, P. M. E.; Carell, T. *Small* **2007**, *3*, 1049.
13. Lundberg, E. P.; El-Sagheer, A. H.; Kocalka, P.; Wilhelmsson, L. M.; Brown, T.; Norden, B. *Chem. Commun.* **2010**, 3714.
14. Seela, F.; Ming, X. *Helv. Chim. Acta* **2008**, *91*, 1181.
15. Kocalka, P.; Andersen, N. K.; Jensen, F.; Nielsen, P. *ChemBioChem* **2007**, *8*, 2106.
16. Fujino, T.; Yamazaki, N.; Isobe, H. *Tetrahedron Lett.* **2009**, *50*, 4101.
17. Isobe, H.; Fujino, T.; Yamazaki, N.; Guillot-Nieckowski, M.; Nakamura, E. *Org. Lett.* **2008**, *10*, 3729.
18. Nuzzi, A.; Massi, A.; Dondoni, A. *Qsar Combinatorial Sci.* **2007**, *26*, 1191.
19. Lucas, R.; Elchinger, P. H.; Faugeras, P. A.; Zerrouki, R. *Nucleoside, Nucleotide Nucleic Acid* **2010**, *29*, 168.
20. Lucas, R.; Zerrouki, R.; Granet, R.; Krausz, P.; Champavier, Y. *Tetrahedron* **2008**, *64*, 5467.
21. Vergnaud, J.; Faugeras, P.; Chaleix, V.; Champavier, Y.; Zerrouki, R. *Tetrahedron Lett.* **2011**, *52*, 6185.
22. El-Sagheer, A. H.; Brown, T. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 15329.
23. Paredes, E.; Das, S. R. *ChemBioChem* **2011**, *12*, 125.
24. Salic, A.; Mitchison, T. J. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 2415.
25. Gierlich, J.; Gutsmiedl, K.; Gramlich, P. M. E.; Schmidt, A.; Burley, G. A.; Carell, T. *Chem. Eur. J.* **2007**, *13*, 9486.
26. El-Sagheer, A. H.; Sanzone, A. P.; Gao, R.; Tavassoli, A.; Brown, T. *Proc. Nat. Acad. Sci. U.S.A.* **2011**, *108*, 11338.
27. Chiou, S. H. *J. Biochem.* **1983**, *94*, 1259.
28. Stoewe, R.; Prutz, W. A. *Free Radical Biol. Med.* **1987**, *3*, 97.
29. Kobayashi, S.; Ueda, K.; Morita, J.; Sakai, H.; Komano, T. *BBA* **1988**, *949*, 143.
30. Won, D. E.; Lee, C. S.; Shin, Y. K.; Lee, K. S. *J. Med.* **1989**, *14*, 59–74.
31. El-Sagheer, A. H.; Brown, T. *Pure Appl. Chem.* **2010**, *82*, 1599.
32. Jewett, J. C.; Sletten, E. M.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2010**, *132*, 3688.
33. Chang, P. V.; Prescher, J. A.; Sletten, E. M.; Baskin, J. M.; Miller, I. A.; Agard, N. J.; Lo, A.; Bertozzi, C. R. *Proc. Nat. Acad. Sci. U.S.A.* **2010**, *107*, 1821.
34. Jewett, J. C.; Bertozzi, C. R. *Chem. Soc. Rev.* **2010**, *39*, 1272.
35. Sletten, E. M.; Bertozzi, C. R. *Org. Lett.* **2008**, *10*, 3097.
36. van Delft, P.; Meeuwenoord, N. J.; Hoogendoorn, S.; Dinkelaar, J.; Overkleeft, H. S.; van der Marel, G. A.; Filippov, D. V. *Org. Lett.* **2010**, *12*, 5486.
37. Jayaprakash, K. N.; Peng, C. G.; Butler, D.; Varghese, J. P.; Maier, M. A.; Rajeev, K. G.; Manoharan, M. *Org. Lett.* **2010**, *12*, 5410.
38. Shelbourne, M.; Chen, X.; Brown, T.; El-Sagheer, A. H. *Chem. Commun.* **2011**, 6257.
39. Shelbourne, M.; Brown, T.; El-Sagheer, A. H.; Brown, T. *Chem. Commun.* **2012**, 11184.
40. Marks, I. S.; Kang, J. S.; Jones, B. T.; Landmark, K. J.; Cleland, A. J.; Taton, T. A. *Bioconjugate Chem.* **2011**, *22*, 1259.
41. Debets, M. F.; Van Berkel, S. S.; Dommerholt, J.; Dirks, A. J.; Rutjes, F.; Van Delft, F. L. *Acc. Chem. Res.* **2011**, *44*, 805.
42. Borrmann, A.; Milles, S.; Plass, T.; Dommerholt, J.; Verkade, J. M. M.; Wiessler, M.; Schultz, C.; van Hest, J. C. M.; van Delft, F. L.; Lemke, E. A. *ChemBioChem* **2012**, *13*, 2094.
43. Dommerholt, J.; Schmidt, S.; Temming, R.; Hendriks, L. J. A.; Rutjes, F.; van Hest, J. C. M.; Lefeber, D. J.; Friedl, P.; van Delft, F. L. *Angew. Chem., Int. Ed.* **2010**, *49*, 9422.
44. Jawalekar, A. M.; Malik, S.; Verkade, J. M. M.; Gibson, B.; Barta, N. S.; Hodges, J. C.; Rowan, A.; Delft, F. L. V. *Molecules* **2013**, *18*, 7346.
45. Gutsmiedl, K.; Fazio, D.; Carell, T. *Chem. Eur. J.* **2010**, *16*, 6877.
46. Weisbrod, S. H.; Marx, A. *Chem. Commun.* **2007**, 1828.
47. Rao, H.; Sawant, A. A.; Tanpure, A. A.; Srivatsan, S. G. *Chem. Commun.* **2012**, 498.
48. Rao, H.; Tanpure, A. A.; Sawant, A. A.; Srivatsan, S. G. *Nat. Protoc.* **2012**, *7*, 1097.
49. Jager, S.; Rasched, G.; Kornreich-Leshem, H.; Engeser, M.; Thum, O.; Famulok, M. *J. Am. Chem. Soc.* **2005**, *127*, 15071.
50. Sagi, J.; Szemzo, A.; Ebinger, K.; Szabolcs, A.; Sagi, G.; Ruff, E.; Otvos, L. *Tetrahedron Lett.* **1993**, *34*, 2191.
51. Znosko, B. M.; Barnes, T. W.; Krugh, T. R.; Turner, D. H. *J. Am. Chem. Soc.* **2003**, *125*, 6090.
52. Giller, G.; Tasara, T.; Angerer, B.; Muhlegger, K.; Amacker, M.; Winter, H. *Nucleic Acids Res.* **2003**, *31*, 2630.
53. Borsenberger, V.; Kukwikila, M.; Howorka, S. *Org. Biomol. Chem.* **2009**, *7*, 3826.
54. Lee, S. E.; Sidorov, A.; Gourlain, T.; Mignet, N.; Thorpe, S. J.; Brazier, J. A.; Dickman, M. J.; Hornby, D. P.; Grasby, J. A.; Williams, D. M. *Nucleic Acids Res.* **2001**, *29*, 1565.
55. Hobbs, F. W. *J. Org. Chem.* **1989**, *54*, 3420.
56. Yoshikawa, M.; Kato, T.; Takenishi, T. *Tetrahedron Lett.* **1967**, *50*, 5065.
57. Hall, L. M.; Gerowska, M.; Brown, T. *Nucleic Acids Res.* **2012**, *40*.

Analyst

PAPER

View Article Online
View Journal | View Issue

Cite this: *Analyst*, 2015, **140**, 2671

Azide and *trans*-cyclooctene dUTPs: incorporation into DNA probes and fluorescent click-labelling†

Xiaomei Ren,[a] Afaf H. El-Sagheer[a,b] and Tom Brown*[a]

Received 23rd January 2015,
Accepted 25th February 2015

DOI: 10.1039/c5an00158g

www.rsc.org/analyst

5-Azidomethyl dUTP and two 5-*trans*-cyclooctene dUTPs with different linkers between the TCO and the uracil base have been incorporated into DNA by primer extension, reverse-transcription and PCR amplification. For azidomethyl dUTP the PCR reaction was successful even when the modified dUTP was not supplemented with dTTP. In one case 335 azidomethyl dU residues were incorporated into the 523 base pair amplicon using this methodology. 5-Azidomethyl dUTP was found to be a better substrate for DNA polymerases than the *trans*-cyclooctene dUTPs. However, the inverse electron demand Diels–Alder reaction between cyclooctene DNA and a tetrazine Cy3-dye was more efficient than the strain-promoted reaction between azide DNA and a bicyclo [6.1.0] non-4-yne Cy3 dye.

Introduction

Nascent DNA synthesis can be monitored in proliferating cells by fluorescence. Cells are incubated with the thymidine analogue 5-ethynyl-2′-deoxyuridine which is converted to its triphosphate by cellular enzymes and incorporated into genomic DNA by endogenous polymerases. Subsequent cell-fixation and derivatisation of the alkyne-modified DNA with a fluorescent azide (*e.g.* TAMRA azide) *via* the CuAAC reaction[1,2] produces labelled DNA (Fig. 1A) which can be visualised by fluorescent cell imaging.[3] The same approach has been adapted to detect RNA synthesis in cells by incorporating the ribonucleoside 5-ethynyluridine into newly transcribed RNA.[4] The methodology is compatible with fixed cells, but the toxicity of copper precludes its use on live cells. Recently it has been shown that the need for toxic metal ion catalysis is negated if an azide-modified nucleoside is incorporated into genomic DNA and a cyclooctyne derivative of the dye[5] (such as bicyclo [6.1.0] non-4-yne (BCN)),[6] is used in the fluorescent labelling step (Fig. 1B).[7] Relief of ring strain by a change of hybridisation from sp to sp^2 drives a fast chemical reaction between the cyclooctyne attached to the dye and the azide-modified DNA. This reaction, known as the strain-promoted alkyne–azide cycloaddition (SPAAC) reaction,[7,8] has also been used to label

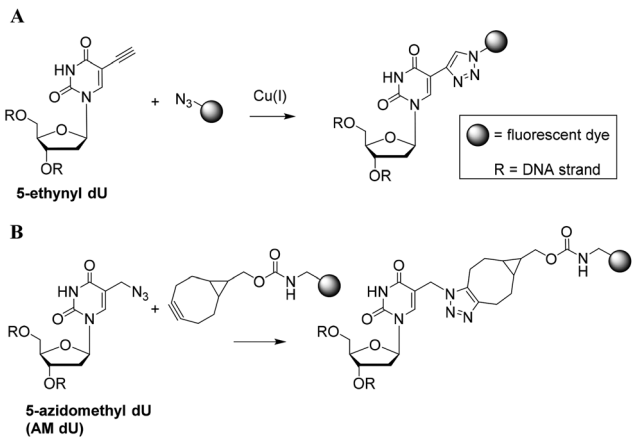


Fig. 1 Fluorescent DNA labelling in cells: (A) 5-ethynyl dU and a fluorescent azide (CuAAC); (B) 5-azidomethyl dU and a BCN-labelled fluorophore (SPAAC).

synthetic oligonucleotides.[9] Unlike bulky dye-labelled nucleosides, ethynyl dU and azidomethyl dU are sufficiently similar to their natural congener (thymidine) to be converted to their triphosphates by endogenous enzymes within cells. This intracellular conversion of nucleoside to dNTP is essential; unlike free nucleosides, dNTPs cannot be used in cell-labelling experiments as they do not pass through the cell membrane.

In addition to cell imaging, there are many applications that require labelled DNA, for which click chemistry is particularly useful.[10] A popular approach is to incorporate functional groups such as alkynes into the DNA during PCR amplification, primer extension or nick translation *via* a modified dNTP, and to subsequently label the modified DNA with an azide derivative of a reporter group such as a fluorescent

[a]*Department of Chemistry, University of Oxford, Chemistry Research Laboratory, 12 Mansfield Road, Oxford, OX1 3TA, UK. E-mail: tom.brown@chem.ox.ac.uk*
[b]*Chemistry Branch, Department of Science and Mathematics, Faculty of Petroleum and Mining Engineering, Suez University, Suez, 43721, Egypt*

†Electronic supplementary information (ESI) available: Synthesis of modified triphosphates and Cy3-BCN. MS and NMR spectra characterization. PAGE and mass spectra results of primer extension and reverse transcription. Melting experiments and labelling experiments of PCR products. See DOI: 10.1039/c5an00158g

dye.[11–14] Similarly, azides can be incorporated into DNA or RNA for fluorophore attachment *via* CuAAC, SPAAC or Staudinger reactions.[15–18] Such two-step labelling processes are necessary because direct enzyme-mediated incorporation of fluorophore-labelled dNTPs during PCR is inefficient due to the bulky nature of fluorescent dyes. In contrast, dNTPs that are modified by addition of sterically undemanding groups such as terminal alkynes are incorporated much more efficiently.

Crystal structures of the thermostable DNA polymerase KlenTaq with alkyne-modified substrate dNTPs in its active site illustrate the plasticity of the enzyme and the presence of flexible regions of the dNTP binding pocket. Due to these factors alkyne modifications at the C5 atom of dU cause only minor disruption to the ternary complex of the enzyme.[13,19]

We are interested in synthesising fluorescent DNA for a variety of biological applications including those described above, and are investigating the incorporation of modified dNTPs into DNA by primer extension, reverse transcription and PCR, and determining the efficiency of subsequent labelling using various click reactions. Here we present studies on azidomethyl dUTP (AM dUTP) and two *trans*-cyclooctene (TCO) modified dUTP analogues with different linkers between the TCO moiety and the nucleobase.

Experimental

Primer extension experiments

In a 20 μL reaction, 66 pmol of FAM labelled DNA primer (P4), 132 pmol of template (T1, T2 or T3 in Table 1) and 3.2 nmol of each modified dUTP or unmodified dNTP were mixed with the polymerase enzyme (1 unit for GoTaq and Klenow, 0.5 unit for KOD) and 1× buffer (supplied with the enzymes). In the case of KOD polymerase, 1 mM $MgCl_2$ was added separately to the reaction mixture. The reaction mixtures were heated for 1.5 h (KOD and GoTaq at 60 °C, Klenow at 37 °C) and 20 μL formamide was added before analysis by 20% denaturing polyacrylamide gel electrophoresis (PAGE) at a constant 20 W of power in 1× TBE buffer. For the mass spectrometry analysis, two 20 μL reactions were carried out, followed by ethanol precipitation. DNA was precipitated by mixing with 4 μl of 3 M sodium acetate (pH 5.3) followed by 120 μl ethanol. The mixture was then left on dry ice for 10 min and at −20 °C for further 20 min. This was followed by centrifugation at 4 °C for 30 min (13 000 rpm). The precipitate was dissolved in 15 μL of water and analyzed on a Bruker microTOF mass spectrometer (ES^- mode). KOD DNA polymerase was purchased from Merck Millipore, Klenow large fragment from New England Biolabs and GoTaq from Promega.

Reverse transcription

Samples containing 66 pmol of FAM labelled DNA primer (P4), 132 pmol of RNA template (T5 and T6 in Table 1) and 3.2 nmol modified dUTP or unmodified dNTP were mixed with reverse transcriptase (M-MuLV RT (RNase H^-) 100 units, or AMV RT 5 units) and 1× buffer. 1× dithiothreitol (DTT, supplied with the enzyme) was added to the M-MuLV reaction, while 20 units of RNase inhibitor were added to AMV reactions. The final reaction volume was 20 μL. The reaction mixtures were heated at 42 °C for 16 h. 20 μL of formamide was added and samples were analyzed by 20% denaturing PAGE at a constant 20 W of power. For the mass spectrometry analysis, two 20 μL reactions were carried out, followed by ethanol precipitation. The DNA precipitate was dissolved in 15 μL of water and analyzed by mass spectrometry. M-MuLV RT (RNase H^-), AMV RT and RNase inhibitor were purchased from New England Biolabs.

Polymerase chain reaction

For template T9, samples containing 10 pmol of two primers (P7, P8, Table 1), 50 pg of template (T9, Table 1) and 5 nmol modified dUTP or unmodified dNTP in total were added followed by GoTaq DNA polymerase (1 unit), 5× GoTaq buffer and SYBR Green (0.6 μL, 3.75×). The final reaction volume was 20 μL with 1× reaction buffer. For plasmid template T12, Samples containing 10 pmol of two primers (P10, P11, Table 1), 300 pg

Table 1 Oligonucleotide sequences used in primer extension, reverse transcription and PCR reactions[a]

Code	Sequences(5′–3′)
T1	C**A**GTC**A**CTGT**A**CTGCCGACACACATAACC (DNA template)
T2	C**A**GTC**A**C**AAAA**CTGCCGACACACATAACC (DNA template)
T3	CTGTC**A**CTGTGCTGCCGACACACATAACC (DNA template)
P4	FAM-GGTTATGTGTGTCGGCAG (primer)
T5	C**A**GUC**A**CUGU**A**CUGCCGACACACAUAACC (RNA template)
T6	C**A**GUC**A**C**AAAA**CUGCCGACACACAUAACC (RNA template)
P7	GCATTCGAGCAACGTAAG (PCR primer)
P8	GGTTATGTGTGTCGGCAG (PCR primer)
T9	FAM-GG**TTA**TGTGTGTCGGC**A**GT**A**TTGTC**A**GTGTGA**A**TTCCAGAGTGTGAGA**T**TGTGTGCTGGCGA**T**CT**TA**CGTTGCTCGA**A**TGC (PCR template)
P10	GTTTGGCTTTAGAGGCTGGAG (Plasmid primer)
P11	ACTGCAATACGAATAATGGCTAC (Plasmid primer)
T12	Template from plasmid HydGdCTD5 (sequence in ESI)

[a] FAM is 6-carboxamidohexylfluorescein. The RNA template has the equivalent sequence to the DNA template. The bold A (A and T for PCR template) show the sites of incorporation of the modified triphosphates. All oligonucleotides were characterized by mass spectrometry.

of plasmid template (HydGdCTD5, sequence in ESI†) and 10 nmol modified dUTP or unmodified dNTP in total were added followed by KOD DNA polymerase (0.5 unit), 10× KOD polymerase buffer (pH 8.0), 25 mM $MgCl_2$ and SYBR Green (0.6 μL, 3.75×). The final reaction volume was 20 μL with 1× reaction buffer and 2 mM $MgCl_2$.

Amplification was performed on a real-time PCR instrument (Bio-Rad CFX96) using the following procedure: an initial denaturing at 95 °C for 2 min, followed by 25 cycles of denaturing at 95 °C for 30 s, primer annealing at 54 °C for 30 s, and extension at 72 °C for 30 s. Further extension was carried out at 72 °C for 5 min. This was followed by melting experiments involving heating the amplification mixture to 95 °C for 30 s, cooling to 30 °C, then increasing the temperature at 1 or 0.5 °C per second to 95 °C and holding at each temperature for 5 s. The samples were analyzed by 1.5% or 2% agarose gel electrophoresis with 20 μL of ethidium bromide (1 mg ml^{-1}). Gel was run at constant voltage (126 V) in 1× TBE buffer.

Fluorescent labelling

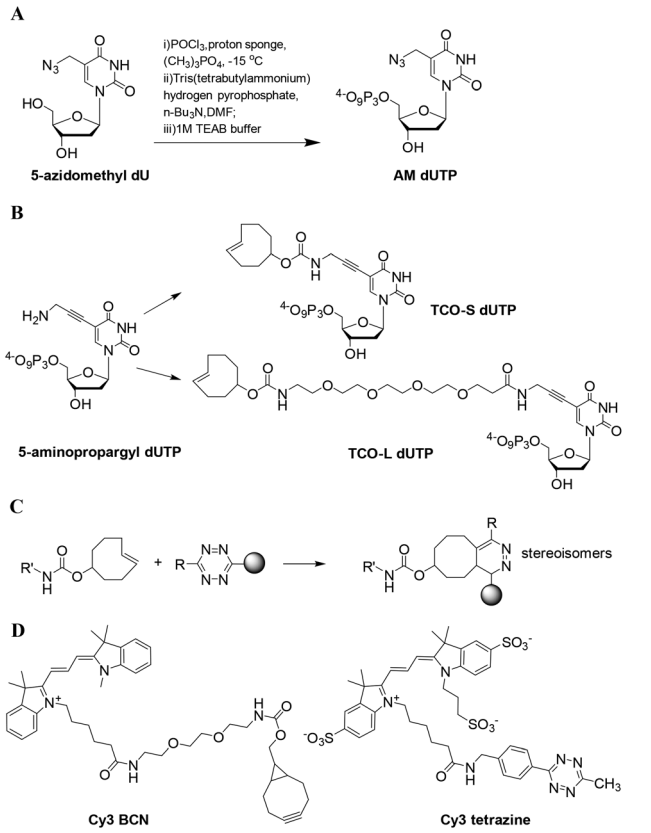
Two 20 μL primer extension reactions were performed as explained in the previous section using AM, TCO dUTPs, or unmodified dNTP negative control. Ethanol precipitation of the product was carried out after primer extension reaction. The precipitate was re-dissolved in 40 μL of 1× GoTaq Green buffer and 20 μL of the AM dUTP extension product was reacted with 6.4 nmol Cy3-bicyclo [6.1.0] non-4-yne (Cy3-BCN). Similarly 20 μL of TCO dUTP products were reacted with 6.4 nmol 6-methyl-tetrazine-sulfo-Cy3. When the reaction was complete, another ethanol precipitation was performed. 20 μl of formamide was added to the labelled and unlabelled primer extension reactions followed by analysis on a 20% PAGE gel at a constant 20 W of power. For the mass spectrometry analysis, two 20 μL labelling reactions were carried out, followed by gel-filtration on a NAP-10 column.

For PCR reactions, products were first extracted from agarose gel using a Qiagen gel extraction kit according to the manufacturer's instructions. The solution was quantified using a Nanodrop 1000 UV spectrometer. 460 ng of T9 products or 736 ng of T12 products were diluted to 40 μL 1× GoTaq Green buffer solution. Next, 20 μL of each solution was labelled with 10 nmol of 6-methyl-tetrazine-sulfo-Cy3 for 30 min. After the labelling reactions, ethanol precipitation was applied to remove extra free dyes. Samples were analyzed by 1.5% or 2% agarose gel electrophoresis at constant voltage (126 V) in 1× TBE buffer. The gel was photographed first, and then stained with ethidium bromide solution (1 μg ml^{-1}) for 20 min to enable the unlabelled DNA to be visualized.

Results and discussion

Preparation of azide and *trans*-cyclooctene modified triphosphates

Modified deoxynucleoside triphosphates were synthesized by the Yoshikawa procedure.[20] Phosphorus oxychloride ($POCl_3$) in trimethylphosphate was employed to generate the 5′-dichlorophosphate derivative of the nucleoside, and the reaction mixtures were treated with tetrabutylammonium pyrophosphate to produce the required triphosphates. The reactions were carried out at −15 °C to minimize the production of by-products. 5-Azidomethyl dUTP (Scheme 1A) and 5-aminopropargyl-dUTP[14] were obtained by this method. 5-Aminopropargyl-dUTP was labelled with the two active esters of *trans*-cyclooctene (structures in ESI Fig. S-1†) to yield two TCO-modified triphosphates with different length of linkage (Scheme 1B).



Scheme 1 Synthesis of (A) AM dUTP and (B) TCO-S dUTP and TCO-L dUTP; (C) general reaction of cyclooctene with a tetrazine- derivatised fluorescent dye; (D) Cy3 derivatives used in this study.

Primer extension reactions

The efficiency of incorporation of the azidomethyl and TCO dUTPs in primer extension experiments was evaluated using three different DNA templates (Table 1). Template T1 has three adenine sites for incorporation of the modified dUTPs, and these are well separated. The more challenging template T2 carries six such sites, four of which are clustered together, presenting a particularly difficult region for the polymerase to incorporate unnatural dNTPs against. 6-Carboxyfluorescein (FAM) was attached to the 5′-end of the primer for ease of visualization in gel electrophoresis analysis. Polyacrylamide gel electrophoresis (PAGE) and mass spectrometry confirmed that KOD and GoTaq thermostable polymerases and Klenow

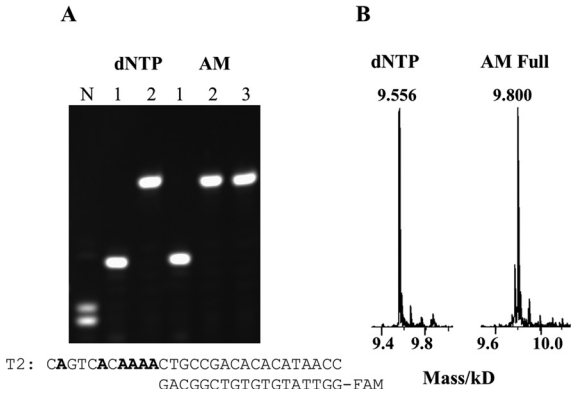


Fig. 2 Successful primer extension using azidomethyl dUTP (AM dUTP) with Klenow polymerase, primer P4 and template T2. 3.2 nmol of each triphosphate was used unless otherwise stated. (A) Lane N, primer and template with dATP, dCTP and dGTP but without dTTP; dNTP lanes, unmodified triphosphate reactions; AM lanes, modified AM dUTP reactions; Lane 1, dTTP or modified dUTP; Lane 2, dTTP or modified dUTP + dATP + dCTP + dGTP; Lane 3, dTTP + modified dUTP (1.6 nmol of each) + dATP + dCTP + dGTP. (20% PAGE gel) (B) Mass spectra (MS) of fully extended primers (lanes 2 from A), unmodified dNTP product (calc. 9553, found 9556); AM dUTP product (calc. 9799, found 9800).

large fragment DNA polymerase all successfully incorporated azidomethyl dUTP (AM dUTP) into the templates T1 and T2 (Fig. 2 and ESI Fig. S-2 and S-3, Table S-1†). In some control reactions in which dTTP or modified dUTP was absent, single nucleotide extension products were observed due to nucleotide misincorporation. The resultant 3′-mismatched complex could not extend further and two closely running bands were observed on PAGE gels (*e.g.* Fig. 2 lane N and Fig. 3 dNTP lane 1).

For the two *trans*-cyclooctene modified triphosphates studied, TCO-S dUTP was consistently incorporated more efficiently than TCO-L dUTP. Polyacrylamide gel electrophoresis and mass spectrometry confirmed that KOD, GoTaq and Klenow large fragment polymerases successfully incorporated TCO-S dUTP into the templates T1 and T2 (Fig. 3 and ESI Fig. S-4, Table S-1†). However, for template T2 the fully extended products of TCO-L dUTP were only obtained using GoTaq polymerase (ESI Fig. S-4, Table S-1†), so the extended ethylene glycol chain did not confer any advantages, in fact it was detrimental. As anticipated, copies of the templates containing TCO dUTPs exhibited lower electrophoretic mobility compared to the natural counterparts (Fig. 3, lanes 1, 2). This is due to the extra bulk of the modified uracil bases, particularly TCO-L dUTP.

In some applications it is advantageous to control (*i.e.* reduce) the density of reactive groups (and ultimately fluorophores) introduced *via* the modified dUTP to avoid fluorescence quenching in the labelled DNA. This can be done using a mixture of the modified dUTP and dTTP in the same extension reaction. This will only be successful if the modified dUTP competes with dTTP for incorporation. To confirm this is the case for the modified dUTPs we carried out linear

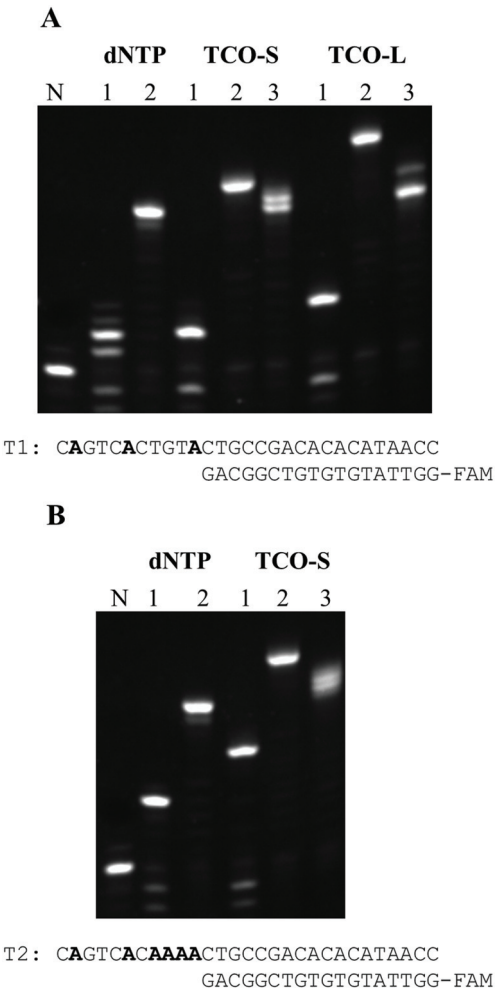


Fig. 3 Successful primer extension using *trans*-cyclooctene dUTPs with Klenow DNA polymerase. 3.2 nmol of each triphosphate was used unless otherwise stated. (A) Reaction with primer P4 and template T1; (B) reaction with primer P4 and template T2. Lane N, primer and template without dNTPs; dNTP lanes, unmodified triphosphate reactions; TCO-S and TCO-L lanes, modified TCO dUTP reactions; Lane 1, dTTP or modified dUTP; Lane 2, dTTP or modified dUTP + dATP + dCTP + dGTP; Lane 3, dTTP + modified dUTP (1.6 nmol of each) + dATP + dCTP + dGTP. (20% PAGE gel).

copying to explore the consequences of mixing them with dTTP in the presence of dATP, dCTP and dGTP (Fig. 2A and 3, lane 3, ESI Fig. S-5†). The modified triphosphates were all able to compete with dTTP, producing fully extended primers with mixtures of dTTP and modified triphosphate opposite to the dA sites in the templates. This is evident from the multiple or broad bands on the gels and from the mass spectra of the fully extended products (ESI Fig. S-5†).

Fluorescent labelling

The products of single nucleotide primer extension were fluorescently labelled to determine the efficiency of single terminal base labelling. The AM dUTP linear extension product was labelled with Cy3-bicyclo [6.1.0] non-4-yne (Cy3-BCN)

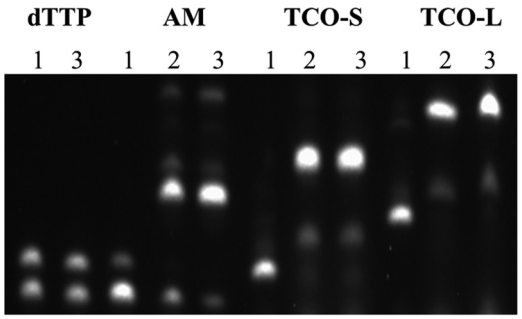


Fig. 4 Efficient fluorescent labelling of single nucleotide primer extension products obtained from modified dUTPs only, GoTaq polymerase, primer P4 and template T1. dTTP and AM dUTP products were labelled with Cy3-BCN; TCO dUTPs products were labelled with 6-methyl-tetrazine-sulfo-Cy3. The bold 3′-T in the extended primer contains either thymidine or the modified nucleotide. Lane 1, unlabelled products; Lane 2, 30 min labelling; Lane 3, 60 min labelling. (20% PAGE gel).

via the strain-promoted cycloaddition (SPAAC) reaction and the *trans*-cyclooctene functionalized product was labelled with 6-methyl-tetrazine-sulfo-Cy3 *via* the inverse electron-demand Diels–Alder (iEDDA) reaction (Scheme 1C and D). The latter reaction went to completion within 30 min, whereas for AM dUTP, some unlabelled starting material remained after 1 h (Fig. 4). Primer extension reaction products obtained from dTTP were employed as negative controls for the labelling reactions and gel mobility studies. The labelled products, which were characterized by mass spectrometry (Fig. 5), exhibited much lower electrophoretic mobility than the unlabeled controls (Fig. 4).

When the labelling reactions were carried out on fully extended products (multiple labelling), the products containing the AM dU nucleotide exhibited much lower labelling efficiency (ESI Fig. S-6†). This is in accordance with the known rate constants for the two reactions. The *trans*-cyclooctene-tetrazine is the fastest known bioorthogonal click reaction with a second-order rate constant up to 22 000 M^{-1} s^{-1}. This is several orders of magnitude faster than the SPAAC reaction (which in turn is much faster than the CuAAC reaction).[21] It is important to bear in mind that these reported kinetic data were obtained from small molecule studies; in the context of DNA labelling various factors such as steric hindrance from the oligonucleotide chain, and clustering of labelled dU residues (sequence context) influence reaction rates. Moreover, when selecting click reactions on the basis of reaction rates it should be stated that the CuAAC reaction can be accelerated by addition of Cu(I)-binding ligands[22] and the SPAAC reaction rate can be greatly increased by the use of electron deficient aryl azides.[23] It should also be noted that some "click" reactions are not truly bioorthogonal, for example cyclooctynes can react with nucleophiles such as thiols which are prevalent in cells, thereby sequestering reactants.[5]

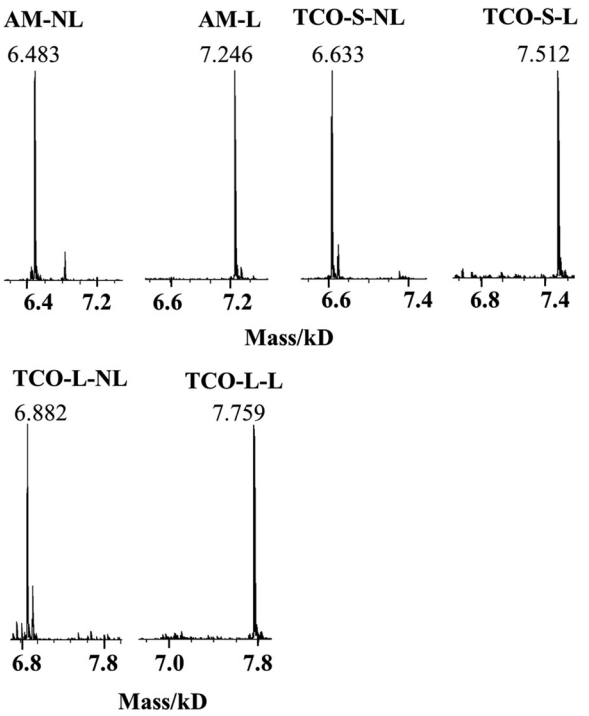


Fig. 5 Mass Spectrometry characterization of single fluorescently labelled nucleotide primer extension products obtained from modified dUTPs only, GoTaq polymerase, primer P4 and template T1. AM dUTP products were labelled with Cy3-BCN; TCO dUTPs products were labelled with 6-methyl-tetrazine-sulfo-Cy3. AM-NL (not labelled): calc. 6483, found 6483; AM-L (labelled with Cy3-BCN): calc. 7246, found 7246; TCO-S-NL: calc. 6633, found 6633; TCO-S-L: calc. 7513, found 7512; TCO-L-NL: calc. 6880, found 6882; TCO-L-L: calc. 7760, found 7759. (NL = not labelled, L = labelled).

Reverse transcription

It is important for biological imaging and diagnostic studies to develop methods to produce fluorescently labelled cDNA probes from RNA molecules of biological origin. To this end our three modified dUTPs were tested for incorporation against a synthetic RNA template employing Moloney Murine Leukemia Virus reverse transcriptase (M-MuLV RT, RNase H^-) and Avian Myeloblastosis Virus reverse transcriptase (AMV RT). The two RNA templates T5 and T6 have the equivalent sequence to the previously used DNA template. AM dUTP and TCO-S dUTP were incorporated efficiently into both templates to give fully extended products with both reverse transcriptase enzymes (Fig. 6 and ESI Fig. S-7 to S-9, Table S-1†). All modified dUTPs were incorporated more efficiently with M-MuLV RT (RNase H^-) than with AMV RT (ESI Fig. S-7, S-9†).

Polymerase chain reaction

In the above linear copying studies we have demonstrated that the AM and TCO dUTPs can be incorporated into DNA by primer extension. The generation of PCR amplification products that contain dNTP analogues is another important method to prepare fluorescent probes. This requires the chemical modifications to be tolerated by DNA polymerases when they are in

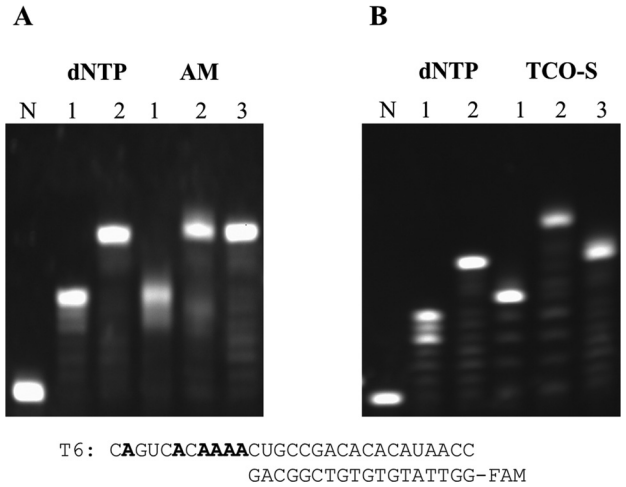


Fig. 6 Efficient reverse transcription reactions using M-MuLV (RNase H⁻) reverse transcriptase with primer P4 and template T6. 3.2 nmol of each triphosphate were used unless otherwise stated. (A) Reactions with AM dUTP; (B) reactions with TCO-S dUTP. Lane N, primer and template without dNTPs; dNTP lanes, unmodified triphosphate reactions; AM lanes, modified AM dUTP reactions; TCO-S lanes, modified TCO-S dUTP reactions; Lane 1, dTTP or modified dUTP; Lane 2, dTTP or modified dUTP + dATP + dCTP + dGTP; Lane 3, dTTP + modified dUTP (1.6 nmol of each) + dATP + dCTP + dGTP. (20% PAGE gel).

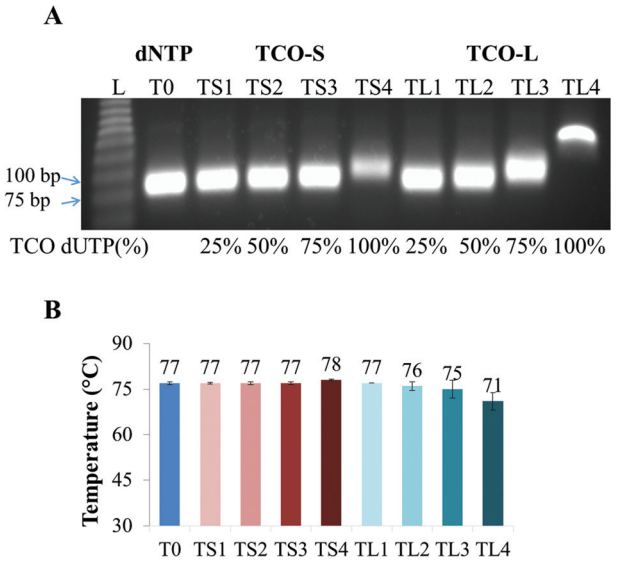


Fig. 8 (A) Efficient PCR amplification with different percentages of TCO dUTPs relative to dTTP (5 nmol in total of TCO dUTP/dTTP + 5 nmol each of dATP + dCTP + dGTP) using GoTaq polymerase and template T9. Lane L, 25 bp ladder; lane T0, unmodified amplicon; lane TS1/TL1 to TS4/TL4, 25% to 100% modified TCO-S/TCO-L dUTP relative to dTTP. (2% agarose gel) (B) Melting temperatures of PCR amplicons.

the template strand as well as when they are attached to dUTP for incorporation into the template. We have investigated this with our AM dUTP and TCO dUTP analogues. Initial PCR reactions were carried out using 81-mer DNA template T9. In all cases amplification was successful, even when using the modified triphosphate in combination with dATP, dCTP and dGTP with no added dTTP (Fig. 7A, lane TA4 and Fig. 8A, lane TS4, TL4). Melting experiments on the amplicons showed that the AM and TCO-L modifications destabilize the DNA duplex in a manner proportional to the number of modifications in the PCR products, whereas TCO-S does not (Fig. 7B and 8B). Such effects of modified nucleotides on duplex stability should be evaluated and taken into account when designing PCR protocols and cycles for incorporation of modified dNTPs.

Encouraged by these results, PCR amplification of a more demanding 523 bp region of a plasmid with a total of 335 thymidine sites was evaluated. Amplification was attempted with AM dUTP along with dATP, dCTP and dGTP in the absence of dTTP (Fig. 9). Although initial PCR with this template using 1 mM $MgCl_2$ was inefficient the problem was easily solved by increasing the Mg^{2+} ion concentration to achieve efficient amplification. This requirement for higher Mg^{2+} concentration might be due to the positive effect of Mg^{2+} on the stability of the duplex formed between the extended primer and the template.

The efficiency of PCR amplification of plasmid template T12 was sensitive to TCO-S dUTP and more so to TCO-L dUTP. We varied the percentage of dTTP and were able to obtain reasonable yields of PCR amplicons with 50% TCO-S dUTP

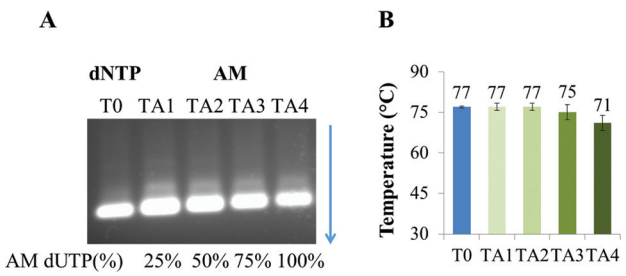


Fig. 7 (A) Efficient PCR amplification with different percentages of AM dUTP relative to dTTP (5 nmol in total of AM dUTP/dTTP + 5 nmol each of dATP + dCTP + dGTP) using GoTaq polymerase and template T9. Lane T0, unmodified amplicon; lane TA1 to TA4, 25% to 100% modified AM dUTP relative to dTTP. (2% agarose gel) (B) Melting temperatures of PCR amplicons.

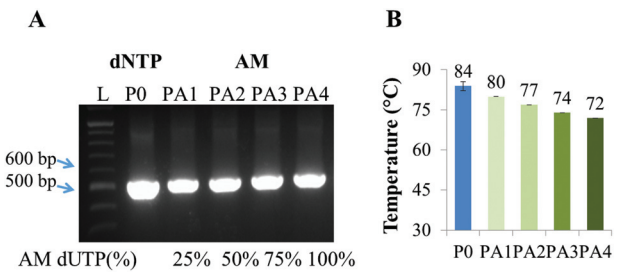


Fig. 9 (A) Efficient PCR amplification with different percentages of AM dUTP relative to dTTP (10 nmol in total of AM dUTP/dTTP + 10 nmol each of dATP + dCTP + dGTP) using KOD polymerase and template T12 (plasmid template). Lane L, 100 bp ladder; lane P0, unmodified amplicon; lane PA1 to PA4, 25% to 100% modified AM dUTP relative to dTTP. (1.5% agarose gel) (B) Melting temperatures of PCR amplicons.

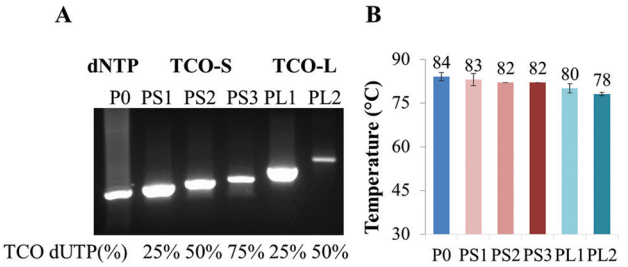


Fig. 10 (A) Efficient PCR amplification with different percentages of TCO dUTP relative to dTTP (10 nmol in total of TCO dUTP/dTTP + 10 nmol each of dATP + dCTP + dGTP) using KOD polymerase and template T12 (plasmid template). Lane P0, unmodified amplicon; lane PS1 to PS3, 25% to 75% modified TCO-S dUTP relative to dTTP; lane PL1 to PL2 25% to 50% modified TCO-L dUTP relative to dTTP. (1.5% agarose gel) (B) Melting temperatures of PCR amplicons.

and 25% TCO-L dUTP (Fig. 10A, lane PS2 and PL1). As observed with the previous shorter amplicon, DNA duplex stability was affected less by TCO-S dUTP than by TCO-L dUTP. Although it might be possible to optimize the PCR reaction for incorporation of higher percentages of TCO-L, the use of TCO-S is a better option for introduction of *trans*-cyclooctene into DNA.

Labelling the PCR products containing *trans*-cyclooctene dUTPs with 6-methyl-tetrazine-sulfo-Cy3 gave the expected pink bands as shown in Fig. 11B for the 523 bp labelled product. The labelled products migrated slower in a manner proportionate to the percentage of modified dUTP used in the PCR amplification (Fig. 11A and ESI Fig. S-14†).

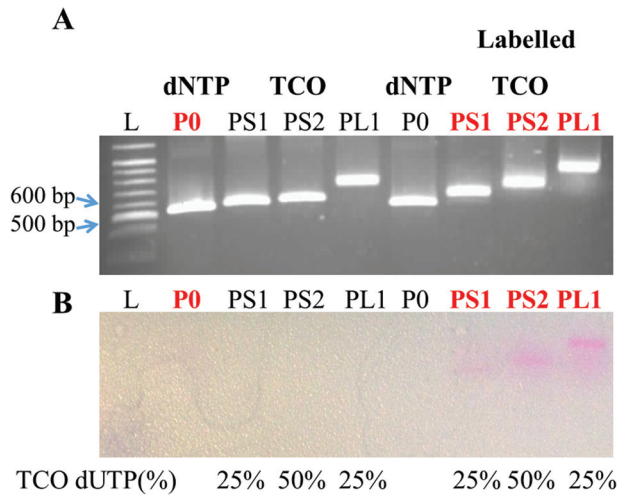


Fig. 11 Fluorescent labelling of PCR products obtained from different percentages of dTTP and TCO dUTPs (10 nmol in total of TCO dUTP/dTTP + 10 nmol each of dATP + dCTP + dGTP) using KOD polymerase and template T12 (plasmid template). Lane L, 100 bp ladder; lane P0, unmodified amplicon; lane PS1 to PS2, 25% to 50% modified TCO-S dUTP relative to dTTP; lane PL1, 25% modified TCO-L dUTP relative to dTTP. Labels above gels in black are for products without 6-methyl-tetrazine-sulfo-Cy3, while labels in bold red are for products that have been reacted with 6-methyl-tetrazine-sulfo-Cy3. The P0 red negative control is positioned distant from the Cy3-labelled products to illustrate that contamination with free dye is not a significant problem. (A) Fluorescent gel was visualized by staining with ethidium bromide; (B) gel picture was taken under normal light. (1.5% agarose gel).

Conclusions

5-Azidomethyl dUTP and two *trans*-cyclooctene dUTPs have been successfully incorporated into DNA by primer extension, reverse-transcription and PCR amplification. PCR was successful with AM dUTP even when the modified dUTP was not supplemented with dTTP. In general AM dUTP is a better substrate than the TCO dUTPs for the DNA polymerases used in this study, and TCO-S dUTP with a short linker is incorporated by DNA polymerases more efficiently than the longer version. The iEDDA click reaction between cyclooctene and tetrazine is more efficient than the SPAAC reaction between azide and BCN, as would be expected from its much faster reported kinetics.[21,24,25]

Very recently we have investigated the use of 5-azidomethyl dU (AM dU) and TCO dU to label nascent genomic DNA in cell culture with fluorescent cyclooctynes, and compared this directly with ethynyl dU incorporation followed by labelling with fluorescent azides. The ethynyl dU experiments yielded excellent fluorescent images whereas the AM dU and TCO dU studies gave disappointing results, with almost no detectable fluorescence. There are several possible reasons for this and further detailed studies are in progress.

Acknowledgements

TB and AHE-S are grateful to the UK BBSRC for funding *via* the sLOLA grant BB/J001694/1: "Extending the Boundaries of Nucleic Acid Chemistry."

Notes and references

1 C. W. Tornoe, C. Christensen and M. Meldal, *J. Org. Chem.*, 2002, **67**, 3057–3064.
2 V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem., Int. Ed.*, 2002, **41**, 2596–2599.
3 A. Salic and T. J. Mitchison, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 2415–2420.
4 C. Y. Jao and A. Salic, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 15779–15784.
5 M. F. Debets, S. S. Van Berkel, J. Dommerholt, A. J. Dirks, F. Rutjes and F. L. Van Delft, *Acc. Chem. Res.*, 2011, **44**, 805–815.
6 J. Dommerholt, S. Schmidt, R. Temming, L. J. A. Hendriks, F. Rutjes, J. C. M. van Hest, D. J. Lefeber, P. Friedl and F. L. van Delft, *Angew. Chem., Int. Ed.*, 2010, **49**, 9422–9425.
7 A. B. Neef and N. W. Luedtke, *ChemBioChem*, 2014, **15**, 789–793.

8 J. C. Jewett and C. R. Bertozzi, *Chem. Soc. Rev.*, 2010, **39**, 1272–1279.
9 M. Shelbourne, T. Brown, Jr., A. H. El-Sagheer and T. Brown, *Chem. Commun.*, 2012, **48**, 11184–11186.
10 A. H. El-Sagheer and T. Brown, *Chem. Soc. Rev.*, 2010, **39**, 1388–1405.
11 J. Gierlich, K. Gutsmiedl, P. M. E. Gramlich, A. Schmidt, G. A. Burley and T. Carell, *Chem. – Eur. J.*, 2007, **13**, 9486–9494.
12 F. Seela, V. R. Sirivolu and P. Chittepu, *Bioconjugate Chem.*, 2008, **19**, 211–224.
13 S. Obeid, H. Busskamp, W. Welte, K. Diederichs and A. Marx, *Chem. Commun.*, 2012, **48**, 8320–8322.
14 X. Ren, M. Gerowska, A. H. El-Sagheer and T. Brown, *Bioorg. Med. Chem.*, 2014, **22**, 4384–4390.
15 H. Rao, A. A. Tanpure, A. A. Sawant and S. G. Srivatsan, *Nat. Protoc.*, 2012, **7**, 1097–1112.
16 H. Rao, A. A. Sawant, A. A. Tanpure and S. G. Srivatsan, *Chem. Commun.*, 2012, **48**, 498–500.
17 M. L. Winz, A. Samanta, D. Benzinger and A. Jaschke, *Nucleic Acids Res.*, 2012, **40**, 13.
18 S. H. Weisbrod and A. Marx, *Chem. Commun.*, 2007, 1828–1830.
19 S. Obeid, A. Baccaro, W. Welte, K. Diederichs and A. Marx, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 21327–21331.
20 M. Yoshikawa, T. Kato and T. Takenishi, *Tetrahedron Lett.*, 1967, **50**, 5065–5068.
21 A. C. Knall and C. Slugovc, *Chem. Soc. Rev.*, 2013, **42**, 5131–5142.
22 T. R. Chan, R. Hilgraf, K. B. Sharpless and V. V. Fokin, *Org. Lett.*, 2004, **6**, 2853–2855.
23 J. Dommerholt, O. van Rooijen, A. Borrmann, C. F. Guerra, F. M. Bickelhaupt and F. L. van Delft, *Nat. Commun.*, 2014, **5**, 5378.
24 D. Wang, W. Chen, Y. Zheng, C. Dai, K. Wang, B. Ke and B. Wang, *Org. Biomol. Chem.*, 2014, **12**, 3950–3955.
25 K. Wang, D. Wang, K. Ji, W. Chen, Y. Zheng, C. Dai and B. Wang, *Org. Biomol. Chem.*, 2015, **3**, 909–915.