Synthesis and enzymatic incorporation of modified deoxyuridine triphosphates†

Vinciane Borsenberger,‡ Mikiembo Kukwikila‡ and Stefan Howorka*

Received 6th April 2009, Accepted 24th June 2009
First published as an Advance Article on the web 21st July 2009
DOI: 10.1039/b906956a

We describe the synthesis of 2′-deoxyuridine-5′-triphosphate derivatives bearing linkers of varying length, bulk and flexibility, at position 5 of the pyrimidine base. Nucleotide analogues with terminal functional groups are of interest due to their application potential for the functional labelling of DNA strands. In the course of the synthesis of the nucleotide analogues, the methodology for the Yoshikawa phosphorylation procedure was optimised, resulting in an approach which reduces the amount of side-products and is compatible with labile functional groups attached to the base. The effect of linker composition on the enzymatic incorporation into DNA was systematically investigated using two different DNA polymerases. Deep $Vent_R$ exo^- from the B-polymerase family accepted most nucleotide analogues as substrates, while Taq from the A-family was slightly less proficient. Both polymerases had difficulties incorporating 5-(3-amino-prop-1-ynyl)-2′-deoxyuridine triphosphate. A molecular model of the active site of the polymerase was used to rationalise why this nucleotide was not accepted as a substrate.

Introduction

The generation of chemically modified deoxyribonucleotide triphosphates is of importance in light of their application in chemical biology, structural biology, DNA sensing, and nanobiotechnology.[1] In many cases, deoxyribonucleotide triphosphates carrying functional tags such as fluorophores are enzymatically incorporated into DNA strands. This pre-modification approach is complemented by the two-step post-modification strategy in which nucleotide precursors carrying a linker are first enzymatically introduced into DNA and then covalently derivatised with functional tags. The coupling to the tags can exploit *e.g.* click chemistry,[2–4] Diels–Alder cycloaddition,[5–7] or maleimide, disulfide,[8] or amide couplings. The post-modification approach is particularly useful for large tags which are not accepted by polymerases such as oligosaccharides[3] or metal nanoparticles.

In deoxyuridine and deoxycytidine, position 5 of the pyrimidine ring is widely used to introduce new functionalities.[9] Within the pre-modification approach, nucleotides carrying *e.g.* a fluorophore attached *via* a long linker at position 5 are usually well accepted by DNA polymerases,[3,9–12] as this site is not involved in Watson–Crick base-pairing thereby minimising any distortion in the shape of the base pairs. With regard to the post-modification approach, substituents at position 5 protrude from the major groove of the DNA duplex thereby ensuring good steric accessibility to facilitate coupling to reagents.[13]

Nucleotide analogues carrying a non-biogenic substitution at position 5 can be chemically synthesised using palladium-catalysed cross-coupling reactions. For the Sonogashira and Suzuki–Miyaura reactions,[14,15] halogenated nucleosides are used in combination with linkers carrying terminal alkyne or aryl/vinyl boronic acids groups, respectively. The cross-coupling reactions are attractive due to their mild conditions which enable the tagging of thermally sensitive substrates. In addition, the reactions can be performed in aqueous conditions to simplify work-up. The attractive cross-coupling reactions have been exploited[12,16–24] in the pre-modification approach to introduce tags such as fluorophores,[9,25] peptides,[26] and transition metal complexes.[27,28]

The synthesis and use of nucleotides for the post-modification approach has, however, attracted less attention. Nucleotides have been prepared with terminal alkyne groups for click-couplings,[3] a diene for Diels–Alder cycloadditions,[7,29] and an amino group[12,22] for amide coupling. In general, linkers for this strategy must meet criteria with respect to linker length and terminal functional group. For example, the linker must be long enough to ensure that the functional group remains accessible for covalent coupling to the functional tags. At the same time, the linkers have to be short enough to avoid back-folding onto the DNA strand which would reduce the rate of reaction. Furthermore, the terminal functional group of the linker must be tolerated by DNA polymerase enzymes.

Herein we present a systematic study on the effect of substituent length, charge, and bulkiness on the enzyme-catalysed incorporation of deoxyuridine triphosphate analogues. Firstly, we synthesised nine nucleotides which bear functional substitutions at position 5 of deoxyuridine (Fig. 1). The modifications include amine and carboxyl functionalities, alkyne and diene groups, as well as bulky Diels–Alder-coupled peptide tags to probe the influence of steric factors. The compounds were synthesised *via* Sonogashira cross-coupling reactions from 5-iodo-2′-deoxyuridine using the corresponding acetylene-based linkers. The resulting nucleosides were converted into deoxyribonucleotide triphosphates using the Yoshikawa procedure.

University College London, Department of Chemistry, 20 Gordon Street, London, UK WC1H OAJ. E-mail: s.howorka@ucl.ac.uk; Fax: +44 (0) 20 7679 7463; Tel: +44 (0) 20 7679 4702

† Electronic supplementary information (ESI) available: Details of the optimisation of the Yoshikawa phosphorylation procedure. See DOI: 10.1039/b906956a

‡ V. B. and M. K. contributed equally to this work.

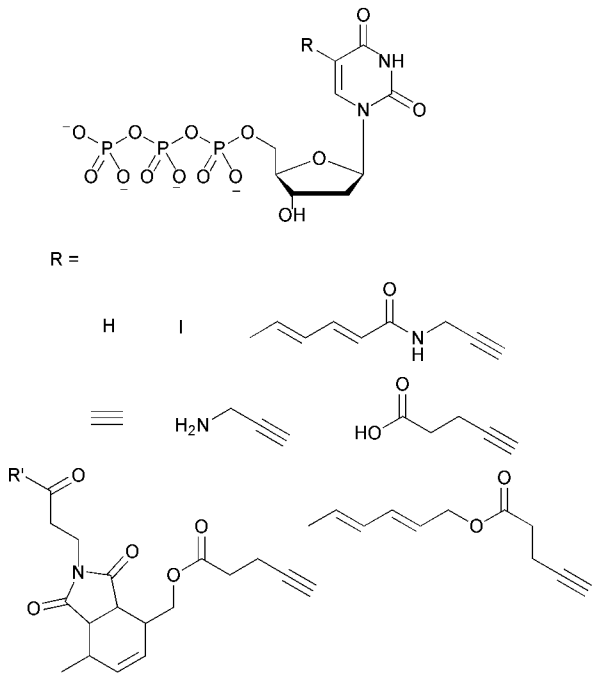
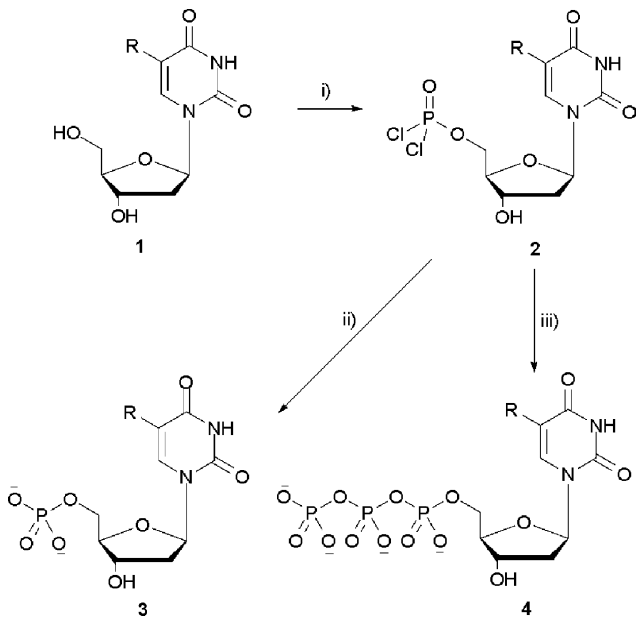


Fig. 1 2′-Deoxyuridine-5′-triphosphate bearing different substituents at position 5 of the pyrimidine. R′ are peptide tags H_6, GVKRKKKP, or GKDDDDYD.

The Yoshikawa procedure[30,31] is a simple and popular phosphorylation approach[1,32] that involves the reaction of a deoxyribonucleoside with phosphorus oxychloride ($POCl_3$) (Scheme 1, step i) to yield a highly reactive phosphorodichloridate intermediate.[30,31] The reactive intermediate can subsequently either be hydrolysed by triethylammonium bicarbonate (TEAB) to the monophosphate product (Scheme 1, step ii) or reacted with tributylammonium pyrophosphate affording the triphosphate deoxyribonucleotide (Scheme 1, step iii).[3,33–35] The reaction with $POCl_3$ is often performed with unprotected deoxyribonucleosides to predominantly achieve 5′-selective phosphorylation. However, the regioselectivity is not perfect, and, consequently, 3′-phosphate and 3′,5′-diphosphate by-products are encountered.[36] While 3′-protection strategies can be enlisted to address this issue, it is advantageous to use unprotected starting material. Indeed, several publications have reported that the extent of side reactions can be suppressed by replacing the traditionally used trialkyl phosphate solvent with other solvent systems or by employing a reaction time of 2–5 hours.[35,37] However, this approach cannot be universally applied to all deoxyribonucleosides due to solubility issues. Our study therefore investigated, secondly, the Yoshikawa phosphorylation procedure in an attempt to optimise yield and regioselectivity for non-protected nucleosides.

Finally, the third aim of our study was to test the incorporation of the nucleotide analogues using DNA polymerases. We chose DNA polymerases Taq and Deep Vent (exo$^-$) which represent the A and B family, respectively. Both enzymes are widely used in the thermal amplification reaction and functional tagging of DNA.[3,7,10]



Scheme 1 The Yoshikawa phosphorylation *via* the phosphorodichloridate intermediate yields mono- and triphosphorylated nucleotides. (i) proton sponge, $POCl_3$; (ii) TEAB; (iii) nBu$_3$NH$^+$·H$_3$P$_2$O$_7^-$; TEAB.

Results and discussion

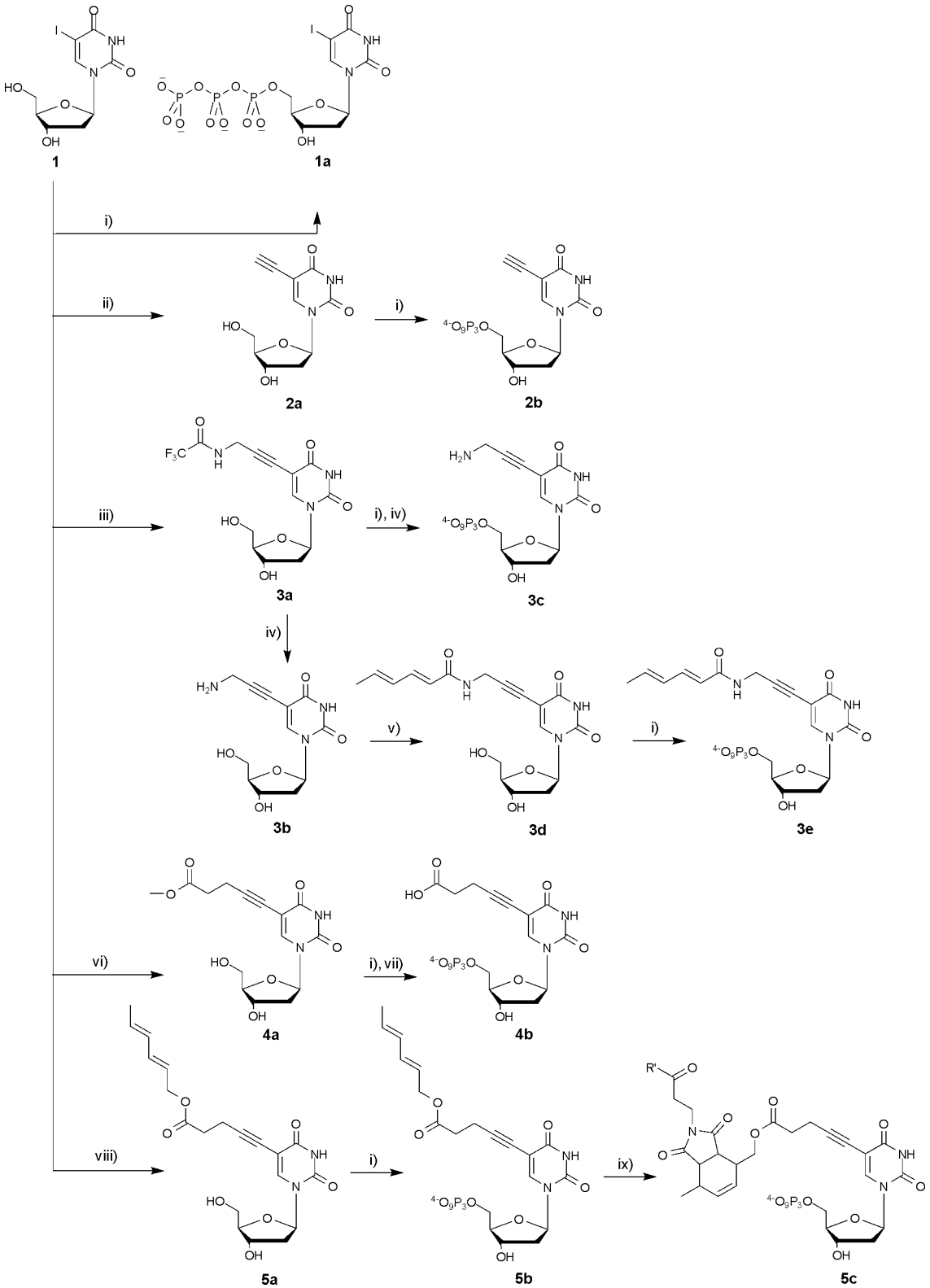
Synthesis of modified nucleosides

The synthesis of the modified nucleosides and their corresponding nucleotides is outlined in Scheme 2. The generation of each nucleoside began with a Sonogashira cross-coupling reaction between 5-iodo-2′-uridine and acetylene-based moieties in the presence of Et_3N, CuI, $Pd(PPh_3)_4$ and DMF. In each reaction, a ratio of approximately 2:1 palladium:copper was employed to avoid oxidation of the catalyst. Furthermore, air was excluded from reaction vessels by repetitive purging and evacuation with argon. Using this procedure, the modified nucleosides were obtained in yields ranging from 33–73% (see Experimental).

The first nucleoside, 5-ethynyl-2′-deoxyuridine, **2a** was obtained *via* a cross-coupling reaction with trimethylsilylacetylene, followed by deprotection to remove the silyl group (Scheme 2, step ii).[3] Deprotection with TBAF proved to be more successful than with sodium carbonate in methanol/water.

The amino-propargyl-substituted derivative, **3c**, was synthesised through a cross-coupling reaction with trifluoroacetate-protected propargylamine yielding **3a** (Scheme 2, step iii). The protected nucleoside was then phosphorylated using an optimised Yoshikawa procedure (see below). The trifluoroacetate protecting group was removed by treatment with concentrated ammonia affording the free amino-terminated nucleotide **3c** (Scheme 2, step iv).

3a was also used to form diene-terminated **3d** (Scheme 2) by, firstly, deprotecting the amine (Scheme 2, step iv) whereby repeated lyophilisation steps were key to eliminate the volatile TFA and ammonium salts. Thereafter, the free amine **3b** was reacted in a one-pot EDC·HCl-mediated amide coupling to potassium sorbate (Scheme 2, step v). HPLC monitoring of the reaction of **3b** showed near quantitative conversion of the amine to **3d** (data not shown). It is noted that this one-pot reaction with EDC did not result in any side-products stemming from the possible reaction of the



Scheme 2 Synthesis of modified nucleotides. *Reagents*: (i) $POCl_3$, proton sponge, TEAB, $n Bu_3NH^+ \cdot H_3P_2O_7^-$; TEAB; (ii) CuI, Et_3N, DMF, $Pd(PPh_3)_4$, trimethylsilylacetylene; TBAF; (iii) CuI, Et_3N, DMF, $Pd(PPh_3)_4$, *N*-propynyltrifluoroacetamide; (iv) TEAB; conc. NH_3; (v) DMF, potassium sorbate, EDC·HCl; (vi) CuI, Et_3N, DMF, $Pd(PPh_3)_4$, pent-4-ynoic methyl ester; (vii) conc. NH_3; (viii) CuI, Et_3N, DMF, $Pd(PPh_3)_4$, pent-4-ynoic acid hexa-2,4-dienyl ester; (ix) maleimide dienophile $R'_1 = H_6$, $R'_2 =$ GVKRKKKP, $R'_3 =$ GKDDDDYD.

activated ester with the hydroxyl groups of the nucleoside. By contrast, coupling routes using *N*-hydroxysuccimide or other more reactive activated esters exhibit less selectivity for the amine and show more side-reactions with hydroxyl groups.[38]

For the synthesis of the carboxyl-terminated derivative **4b**, pent-4-ynoic methyl ester was first coupled to position 5 (Scheme 2, step vi). The protected nucleoside **4a** was then phosphorylated to the triphosphate **4b** (see below). Treatment with concentrated ammonia removed the protecting methyl group and afforded the free carboxylic acid.

A second diene-modified nucleoside, **5a**, was prepared *via* a cross-coupling reaction with sorbyl-4-pentyonate (Scheme 2, step viii). Subsequent phosphorylation (see below) yielded target nucleotide **5b**.[7] The diene-modified nucleotide was reacted with a series of maleimide-terminated peptides to form three Diels–Alder adducts **5c-R$_1$**, **5c-R$_2$** and **5c-R$_3$** (Scheme 2, step ix). The peptides were obtained by solid-phase peptide synthesis, and their sequences were chosen to produce functionally relevant tags. The first peptide in **5c-R$_1$** is the hexahistidine tag[7] which can be useful for metal affinity chromatography purification. The nuclear localisation sequence of the second tag (GVKRKKKP in **5c-R$_2$**) is used in cell-permeating peptide–oligonucleotide conjugates.[39] Finally, the FLAG epitope tag is employed in cell biological fluorescence labelling (GKDDDDYD in **5c-R$_3$**). The Diels–Alder couplings between the peptide–maleimide dienophile and the nucleotide diene were conducted at pH 6.1 to avoid the base-catalysed hydrolysis of the maleimide.[7]

All nucleosides and their corresponding nucleotides were purified by RP-HPLC or flash chromatography and their chemical identity was confirmed by ESI-MS, MALDI-TOF ^{1}H-NMR, ^{13}C-NMR and ^{31}P-NMR. The modified dNTPs were stored as lyophilisates at –20 °C and all are stable for several months except **5c-R$_3$** (see Experimental).

Phosphorylation of nucleosides

The phosphorylation of the nucleosides was conducted using the Yoshikawa procedure. Following a published protocol,[9] we first reacted 5-iodo-2′-deoxyuridine **1** with $POCl_3$ followed by incubation with tributylammonium phosphate for 2 h to obtain the triphosphate nucleotide (Scheme 1, steps i and iii). The methodology yielded the target compound (not shown), but also produced a high content of the diphosphate by-product. These unsatisfactory results prompted a detailed investigation into the Yoshikawa reaction with respect to yield and regioselectivity as a function of reaction duration, temperature and molar excess of $POCl_3$.

To examine the regioselectivity and yield, the model nucleoside 5-iodo-2′-deoxyuridine was reacted with 1.5 equivalents of $POCl_3$ at 0 °C, similar to published conditions.[9,33] At intervals, aliquots of the reaction mixture were withdrawn, treated with TEAB to hydrolyse any phosphorodichloridate intermediate (Scheme 1, step iii), and analysed by RP-HPLC. The chromatogram of the reaction mixture after 3.5 min indicates that the starting material (Fig. 2A, 0 min, symbol S) was partially converted into peaks A, B, and C (Fig. 2A 3.5 min). These peaks constitute the 5′,3′-diphosphate, 5′-monophosphate, and 3′-phosphate derivatives, respectively, as demonstrated by ESI-MS and ^{1}H analysis. Continued incubation of up to 27 min led to a further decrease in starting material, an increase of 5′-monophosphate product peak B and diphosphate by-product peak A, but no major changes for 3′-phosphate peak C (Fig. 2A). A summary of the observed reaction kinetics (Fig. 2B) highlighted that the level of the 5′-monophosphate rose fast within the first 5 min to around 70% and levelled off thereafter (Fig. 2B, filled circles). By comparison, diphosphate increased steadily and constituted 15% at 5 min and close to 30% at 27 min (Fig. 2B, empty squares). This suggests an optimal reaction window between 3 to 7 min that achieves high levels of target and a low level of side-product. The short time needed to complete the phosphorylation reaction is an important finding as most other reports use a reaction time of at least 2 h.[9,33]

After investigating yield and regioselectivity as a function of time, we examined the influence of two other parameters: the reaction temperature and the molar ratio of $POCl_3$ to deoxyribonucleoside. As detailed in the ESI†, the phosphorylation of 5-iodo-2′-deoxyuridine **1** was conducted at either 0 °C or –13 °C, and at molar ratios of 1.0, 1.5, or 3.0 equivalents of $POCl_3$. The kinetic analysis *via* HPLC found, firstly, that a reduction in temperature to –13 °C at a ratio 1.5 had only a minor effect on the generation of 5′-phosphate, but a beneficial consequence by slowing down the formation of the diphosphate by-product. Secondly, increasing the ratio to 3.0 equivalents at –13 °C neither affected product nor by-product formation. Thirdly, a ratio of 1.0 avoided the generation of diphosphate by-product but at the cost of decreased yield of the product.

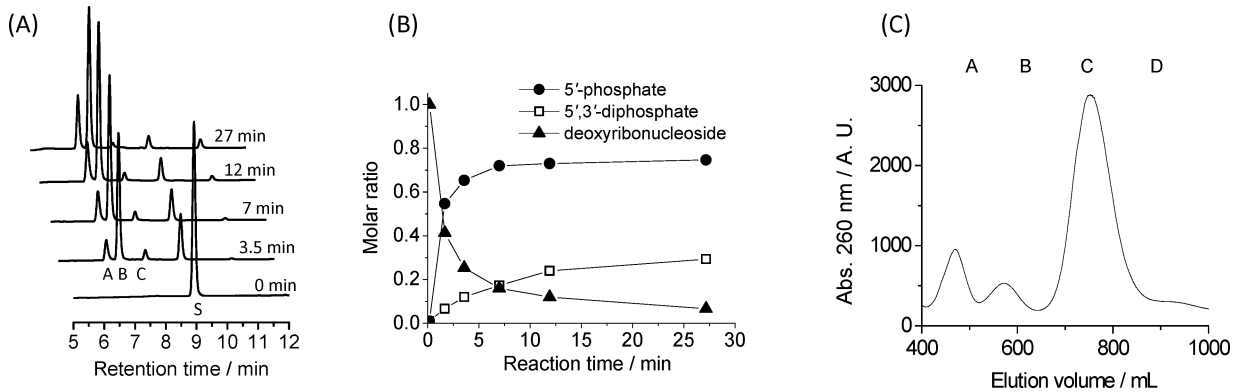


Fig. 2 (A) HPLC traces for the reaction of **1** with 1.5 equivalents of $POCl_3$ at 0 °C. Y-scale: Absorbance at 260 nm. (B) Corresponding plot of peak areas *vs.* time. (C) IEX chromatogram of the product mixture for the phosphorylation of **3e** with $POCl_3$ and pyrophosphate.

The optimal reaction conditions inferred from the monophosphorylation of **1** (1.5 equivalents of $POCl_3$, temperature of at least −13 °C) were applied to generate the triphosphate derivative of our nucleosides. Following Scheme 1 (steps i to iii), the functionalised deoxyuridines were triphosphorylated by treatment with pyrophosphate of the dichloridate species, and subsequently hydrolysed. Fig. 2C illustrates a typical example of an anion exchange chromatogram (IEX) of the crude reaction mixture, with one major component (Fig. 2C, peak C) corresponding to the expected triphosphate, and minor species being mono-, di- and tetra-phosphate nucleotides (Fig. 2C, peaks A, B, and D, respectively). The improved Yoshikawa methodology was extended to the other modified nucleosides giving rise to the corresponding triphosphate derivatives **2b**, **3c**, **3e**, **4b**, **5b**, **5c-R$_1$**, **5c-R$_2$**, **5c-R$_3$** (Scheme 2) in quantitative yields. The successful outcome of the phosphorylation reactions illustrates that the improved Yoshikawa methodology is applicable to a wider range of nucleoside analogues. Furthermore, it is noted that the phosphorylation of **5b** carrying the labile allylic ester was well tolerated and did not lead to hydrolysis.

Enzymatic incorporation of modified nucleotides

All deoxyuridine triphosphate analogues were utilised in primer extension assays to evaluate their ability to act as substrates for commercially available DNA polymerases Deep Vent exo$^-$ and Taq. The principle of the assays is schematically summarised in Fig. 3A. A short primer is hybridised to a complementary section of a longer DNA oligonucleotide. Its non-hybridised section serves as a template to direct the extension of the primer *via* the enzyme-catalysed addition of nucleotides. As the template contains three adenine bases (underlined in Fig. 3A), the primer can be extended by incorporating up to three modified deoxyuridine residues dU*TP. The sequence of the template is designed so that the extension can be controlled by the inclusion or omission of particular bases (Fig. 3B). For instance, dNTP mix 0 lacks a uridine or thymine which is the first nucleotide to be incorporated into the primer. Consequently, no extension is expected to occur for dNTP mix 0. Conversely, dNTP mix F contains a full set of natural bases, and the primer should be extended by 12 nt to the same length as the template (Fig. 3B). By comparison, dNTP mix H lacks dGTP, and the primer is to be extended halfway by 6 nt (Fig. 3B). Similarly, the assay to test for the enzymatic incorporation of the nucleotide analogues was conducted *via* primer extension by six or 12 bases. The corresponding nucleotide mixes dNTP 6 and dNTP 12 contain the chemically tagged uridine derivates in place of dTTP (Fig. 3B). Based on the sequence of the template, the primer extension should lead to the incorporation of two and three dU*TP bases, respectively (Fig. 3B).

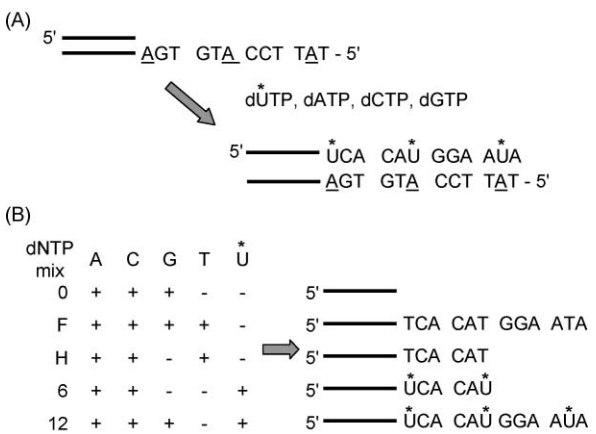


Fig. 3 Template-directed primer extension. The primer and template form a duplex, and the non-hybridised section of the template directs the polymerase-catalysed addition of nucleotides to the 3′ of the primer. (B) dNTP mixes of varying composition determine the number of incorporated nucleotides comprising the nucleotide analogue dU*TP.

The primer extension assays were analysed by determining the length of the extension products *via* gel electrophoresis, followed by staining of the DNA bands. The gel analysis is summarised in Fig. 4. Panel A shows the control reactions which confirm the viability of the primer extension assay. In Fig. 4A, lane 1 and 2 correspond to the 36 nt-long template and the 24 nt-long primer, respectively. By comparison, the use of complete dNTP mix F led to the disappearance of the non-extended primer due to the full extension of the primer which co-migrated with the template in a single band (Fig. 4A, lane 4). By contrast, dNTP mix H which is designed to direct the extension by 6 nt resulted in an up-shift of the primer band (Fig. 4A, lane 5) but not to the co-migration as seen for dNTP mix F (Fig. 4A, lane 4). Furthermore, no extension occurred for dNTP 0 (Fig. 4A, lane 3), in line with expectations.

The results of the 6-nt extension reaction using modified deoxyuridine derivatives and Deep Vent is shown in Fig. 4B. The enzymatic reaction with **1a**, **2b**, **4b**, **3e**, **5b**, and peptide-tagged analogues **5c-R$_1$**, **5c-R$_2$**, **5c-R$_3$** proceeded successfully, as indicated by a major up-shifted band stemming from the extended primer (Fig. 4B, lanes 1, 2, 4 to 9). Enzymatic incorporation of 5-ethynyl-2′-deoxyuridine, **2b** and diene-modified nucleotide **5b** was also observed in other studies.[3,7] These two nucleotides hence serve as positive control and reference for our extension assay. The extended primers carrying the modified nucleotides migrated at about the same height as the strand from natural nucleotide mix H (compare Fig. 4B, lanes 1, 2, 4 to 6 with Fig. 4A, lane 5). The absence of any tag-induced gel shift is most likely due to the small running distance in the gel and the low mass of the chemical tags. By contrast, the bands obtained from peptide-modified nucleotides **5c-R$_1$**, **5c-R$_2$**, and **5c-R$_3$** showed a considerable gel shift (Fig. 4B, lanes 7 to 9) reflecting the approximately ten-fold higher masses of the peptide tags compared to the other chemical tags. These results suggest that in each case the enzyme is able to extend the primer by 6 nucleotides and thereby incorporate two modified nucleotides (Fig. 3B, dNTP mix 6). Interestingly, almost no incorporation occurred for the amino-propargyl derivative **3c** (Fig. 4B, lane 3). The inability of **3c** to serve as substrate is most likely related to the terminal amino group of the substituent at position 5 (see further below).

Similar results were found for the extension by 12 nt using Deep Vent (Fig. 4C). Fully extended primers were obtained for **1a**, **2b**, **4b**, **3e** and **5b** (Fig. 4C, lanes 1, 2, 4 to 6), while amino-propargyl **3c** was not accepted as substrate (Fig. 4C, lane 3). A lower incorporation extent was found for peptide-modified nucleotides **5c-R$_1$**, **5c-R$_2$**, and **5c-R$_3$** (Fig. 4C, lanes 7 to 9). In particular, the addition of a third nucleotide into the partly extended primer was likely blocked due to steric reasons. Complete incorporation

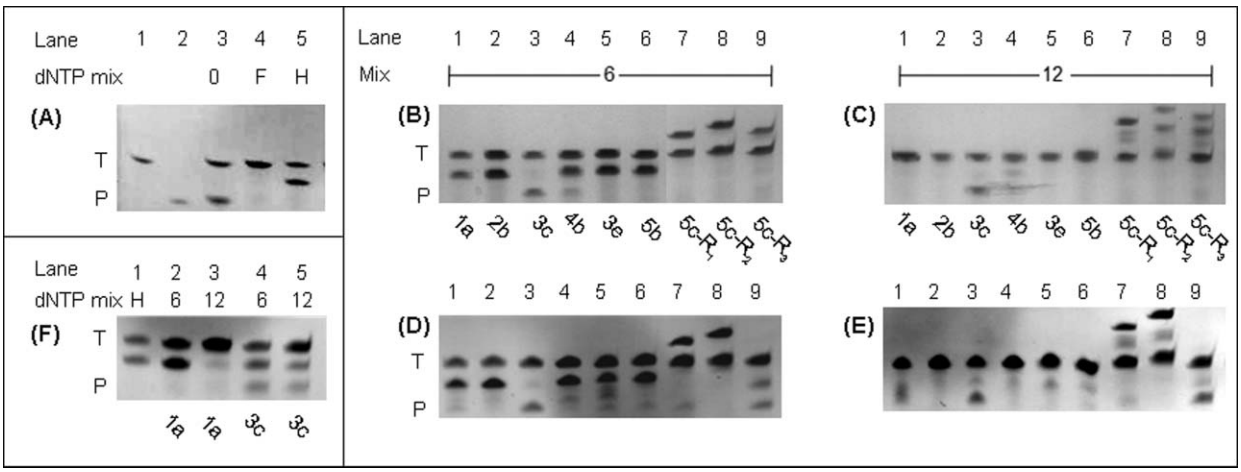


Fig. 4 Denaturing PAGE analysis of extension reactions with modified nucleotides using Deep Vent exo⁻ (B,C) and Taq (D, E) for dNTP mixes 6 and 12 as defined in Fig. 3B, conducted at 57 °C. (A) are control reactions with dNTP mixes 0, F, H, as defined in Fig. 3B. Lane 1 and 2 contain template, T, and primer, P, respectively. (F) shows the extension conducted at 70 °C using mix H, and mixes 6 or 12 containing **1a** or **3c**.

was probably hindered because the third and first tagged base were separated by an almost complete helical turn (10 bases) potentially leading to steric crowding between the two peptides positioned at the same side of the duplex.

The assay with Taq polymerase yielded results as shown in Fig. 4D and Fig. 4E. Incorporation by 6 nt was mostly complete for **1a**, **2b**, **4b**, **3e**, **5b**, **5c-R$_1$**, and **5c-R$_2$** (Fig. 4D, lanes 1, 2, 4 to 8). In comparison to Deep Vent, the extent of incorporation by Taq was slightly lower, as exemplary illustrated by a range of minor bands stemming from prematurely terminated extension product of **3e** (Fig. 4D, lane 5). As in the case of Deep Vent, amino-propargyl nucleotide **3c** was a poor substrate for enzymatic polymerisation by Taq (Fig. 4D, lane 3), even though another study reported good incorporation, albeit with the different enzyme Tth.[12] Our low extent of incorporation for **5c-R$_3$** carrying a FLAG peptide epitope (Fig. 4D, lane 9) reflected the partial hydrolysis of the triphosphate (IEX analysis not shown) which was most likely caused by the low pH in the solution of the acidic FLAG peptide tag. Re-purification of **5c-R$_3$** resulted in full incorporation (not shown). The assays for the extension by 12 nt (Fig. 4E) supported the previous findings on the incorporation of most nucleotide analogues as well as for the poor incorporation of **3c** (Fig. 4E, lane 3) and the sterically blocked extension of peptide-tags **5c-R$_1$** and **5c-R$_2$** (Fig. 4E, lanes 7 and 8). Hydrolysis accounted for the poor incorporation of **5c-R$_3$** (Fig. 4E, lane 9).

The likely molecular reason for the poor incorporation of amino-propargyl nucleotide **3c** was investigated with a molecular model of a polymerase. A model was constructed using the X-ray structure of the T7 DNA polymerase,[40] which—like Taq—belongs to the type A family and shares a conserved active site. The X-ray structure of T7 DNA polymerase is a ternary complex with the double stranded DNA template containing a single stranded overhang, and an incoming dTTP nucleotide complementary to base A in the overhang.[40] Our model of the active site (Fig. 5) shows the protein (grey), the DNA template (magenta), and the nucleotide. The nucleotide 5-(3-amino-propynyl)-2′-deoxyuridine 5′-triphosphate **3c** was placed in the same position as dTTP in the X-ray file. Visual inspection of the active site (Fig. 5) revealed that the primary amino group of the base substituent is close to

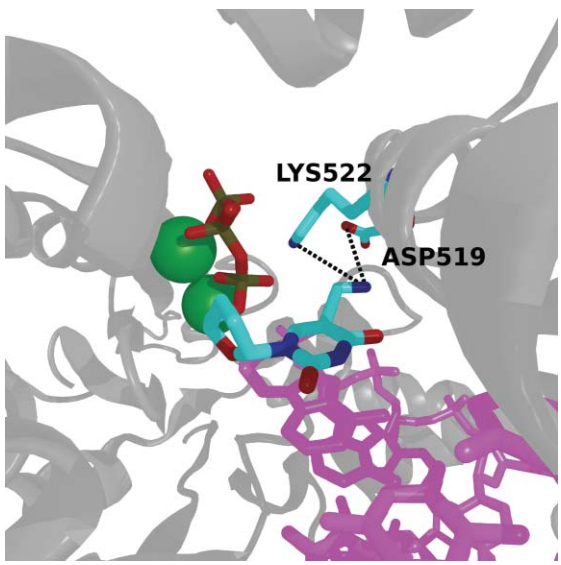


Fig. 5 Molecular model of the active site of the T7 DNA polymerase from the A-type family. The model shows the electrostatic interactions between the amino group of 5-(3-amino-propynyl)-2′-deoxyuridine 5′-triphosphate **3c** and the side chains of aspartic acid and lysine. The protein is in grey, the DNA strand of the template is in magenta, and the two Mg^{2+} ions are represented as green spheres. The metal ions electrostatically interact with the triphosphate moiety of the nucleotide. The nucleotide **3c** was placed into the position of the incoming dTTP nucleotide in the crystal structure of the ternary protein·DNA·dTTP complex (PDB ID 1T7P).

two charged amino acid moieties. As measured from its nitrogen atom, the amino-propargyl group is 4.5 Å apart from the nitrogen of lysine at position 522, and 4.7 Å away from the closest oxygen of aspartic acid at position 519 (Fig. 5). Based on the importance of substrate recognition by key residues in DNA polymerases,[41,42] we propose that the electrostatic interaction between the amino group and the charged residues is the key to the inhibition of the polymerase activity. In particular, the electrostatic attraction of amino-propargyl to Asp, and the electrostatic repulsion to Lys, is most likely altering the position of the pyrimidine base within the active site (not shown) thereby affecting the ability of the

polymerase to accept the nucleotide as a substrate. Several observations are consistent with the notion that electrostatic interactions cause the poor acceptance of amino-propargyl nucleotide **3c**. Firstly, abolishing the ionic interaction by neutralising the amino group of **3c** *via* acylation with hexa-2,4-dienoic acid (thereby yielding nucleotide **3e**, see Scheme 2) resulted in a high polymerisation extent for Taq (lanes 5 in Fig. 4B and C). In addition, weakening the interaction by increasing the temperature of polymerisation from 57 to 70 °C improved the acceptance of **3c** by the enzyme, as seen for dNTP mix 6 (Fig. 4F, lane 4) and dNTP mix 12 (Fig. 4F, lane 5). The extent of incorporation of **3c** was, however, not complete. While the ionic interaction between nucleotide and enzyme is a plausible explanation, we cannot, however, rule out other reasons. For example, a modified nucleotide added to a DNA strand could affect the interaction between primer and template strands and thereby lead to DNA duplex stability problems.[12] Further work using point-mutated enzymes would be required to unequivocally clarify whether the charged residues in the active site are the key to the poor acceptance of amino-propargyl derivative **3c**.

Conclusions

In this study, we have described the synthesis of deoxyuridine triphosphate derivates using the Sonogashira cross-coupling. As the triphosphate analogues were being used for the post-modification labelling of DNA strands, we explored pyrimidine derivatives that carried at position 5 linkers of varying length, flexibility and terminal functional groups. In the course of the synthesis of the nucleotide derivatives, the Yoshikawa procedure was optimised. Compared to the long incubation time in the literature, we found an optimal shorter reaction window between 3 to 7 min that achieved high levels of target and a low level of side-product. The procedure is also compatible with labile functional groups attached to the base. As shown by primer extension assays, almost all nucleotides were well accepted by two widely used DNA polymerases. Blockade of nucleotide incorporation was observed for sterically demanding peptide-tagged nucleotides as well as for the amino-propargyl derivative of deoxyuridine. The poor extent of incorporation was most likely related to electrostatic interactions with amino acid side chains, as demonstrated using a molecular model of the active site of an enzyme. In conclusion, by covering the chemistry and biology of functional deoxyuridine triphosphates our study provides important insights into their synthesis and biochemical use, thereby complementing existing reports on purine-containing nucleotide triphosphates.[3,9–12,43]

Experimental

General

Trimethylphosphate was dried under a 3Å-molecular sieve prior to use. All nucleosides and proton sponge were stored in a desiccator over P_2O_5 several days before use. Anhydrous tributyl ammonium pyrophosphate solution was prepared following a published procedure.[44] All other chemicals and solvents were of commercial quality and used as received. Deep Vent$_R$® (exo$^-$) polymerase was purchased from New England Biolabs (NEB) and DNA oligonucleotides from Integrated DNA Technologies. Preparative ion exchange chromatography was performed on a Sephadex A-25 DEAE (120 mL) column, a Resource Q (1 mL) or a Resource S (1 mL) column on an Äkta Purifier system. Analytical HPLC was performed on a Varian Pro Star system equipped with a dual wavelength 325 detector and a Gemini C18 column (250 × 4.6 mm). Preparative HPLC was performed on the same system using a Polaris C18 column (100 × 21.2 mm, 5 µm beads). Mass spectrometric analysis was performed on a Waters Acquity Ultra Performance LC/MS system equipped with an Acquity UPLC BEH C18 column (50 × 2.1 mm, 1.7 µm beads). NMR spectra were recorded at 293K on a Bruker AVANCE500 spectrometer (for ^{1}H- and ^{13}C-NMR) and on a Bruker AMX300 spectrometer (for ^{31}P-NMR). Coupling constants are reported in Hz. Polyacrylamide gels were prepared with 18% acrylamide in *tris*-borate buffer containing 7 M urea and 18% formamide, and stained using ethidium bromide.

Synthesis of nucleosides

5-Ethynyl-2′-deoxyuridine (2a). 5-iodo-2′-deoxyuridine, **1**, (370 mg, 1.04 mmol), palladium tetrakis(triphenylphosphine) (113 mg, 98 µmol), and copper iodide (47 mg, 250 µmol) were placed in a dried flask flushed with argon. Anhydrous DMF (12.0 mL) was added *via* syringe, followed by triethylamine (0.4 mL, 2.9 mmol) and trimethylsilylacetylene (0.71 mL, 5.0 mmol). The mixture was stirred at room temperature for 3.5 h, then solvents were removed *in vacuo*. The resulting residue was submitted to flash column chromatography (10% MeOH in DCM) affording 381 mg of yellow foam. The foam was dissolved in THF (4 mL) and TBAF in THF (1.2 mL, 1M, 1.2 mmol). After stirring the reaction mixture at room temperature for 3 h, the resulting white precipitate was removed by filtration and rinsed first with THF and then with 10% MeOH in DCM (2 mL). The combined filtrates were concentrated and purified by flash column chromatography (10% MeOH in DCM) to afford 191 mg (73%) of white powder. δ_H(300 MHz; DMSO-d_6) 2.12 (2H, dd, J = 5.1, 5.1), 3.60 (2H, q, J = 11.7), 3.78 (1H, q, J = 3.3), 4.09 (1H, s), 4.20–4.25 (1H, m), 5.13 (1H, t, J = 4.6), 5.24 (1H, d, J = 4.1), 6.09 (1H, t, J = 6.6), 8.29 (1H, s), 11.62 (1H, s, H-3). δ_C(300 MHz; DMSO-d_6) 40.1, 60.7, 69.9, 76.4, 84.7, 87.5, 97.5, 144.5, 149.4, 161.6. ESI-MS(pos): 275.3 m/z [M + Na]$^+$, calculated mass for $C_{11}H_{12}N_2O_5$: 252.2.

5-[3-(Trifluoroacetamido)propynyl]-2′-deoxyuridine (3a). 5-iodo-2′-deoxyuridine, **1**, (740 mg, 2.1 mmol), palladium tetrakis(triphenylphosphine) (61 mg, 53 µmol), and copper iodide (29 mg, 152 µmol) were placed in a dried flask flushed with argon. Anhydrous DMF (7.0 mL) was added *via* syringe, followed by triethylamine (0.6 mL, 4.0 mmol) and *N*-propynyltrifluoroacetamide (762 mg, 5.0 mmol) in anhydrous DMF (7 mL). After stirring the mixture at room temperature for 4.5 h, the solvents were removed *in vacuo*, and the residue was purified by flash column chromatography (8% MeOH in DCM) affording 790 mg of yellow foam. δ_H(300 MHz; DMSO-d_6) 2.11 (3H, t, J = 5.4), 3.52–3.62 (2H, m), 3.78 (1H, q, J = 2.7, 3.2), 4.21 (3H, d, J = 5.35), 5.08 (1H, t, J = 4.3), 5.23 (1H, d, J = 3.7), 6.09 (1H, t, J = 6.4), 8.19 (1H, s), 10.05 (1H, t, J = 5.3), 11.62 (1H, s). δ_C(300 MHz; DMSO-d_6) 29.4, 40.0, 60.9, 70.1, 75.3, 84.7, 87.4, 87.6, 144.1, 149.3, 161.5. δ_F(282 MHz; DMSO-d_6) -74.71.

5-(Hexa-2,4-dienoic acid prop-2-ynylamide)-2′-deoxyuridine (3d). **3a** was taken up in a 1:1 solution of conc. ammonia and water (10 mL). After 2 h, solvents were evaporated *in vacuo*. The residue was taken in water and lyophilised. The resulting foam (**3b**) was dissolved in anhydrous DMF, and mixed with potassium sorbate (153 mg, 1.0 mmol) and EDC·HCl (304 mg, 2.0 mmol). After stirring under argon for 48 h, the reaction mixture was filtered, and DMF was removed using co-evaporation with water. The brown residue was taken up in 5% MeOH in DCM and triturated, which resulted in the precipitation of a yellow solid. The solid was filtered off and dried in air, then washed with water, and a small amount of methanol, affording 125 mg of white crystals (33%). δ_H(300 MHz; DMSO-d_6) 1.79 (3H, d, $J = 5.9$, H-16), 2.10 (2H, t, $J = 5.4$, H-2′), 3.50–3.64 (2H, m, H-5′), 3.78 (1H, q, $J = 3.2$, H-4′), 4.14 (2H, d, J = 5.4, H-9), 4.18–4.24 (1H, m, H-3′), 5.09 (1H, t, $J = 4.8$, 5′-OH), 5.22 (1H, d, $J = 4.3$, 3′-OH), 5.87 (1H, d, $J = 14.9$, H-12), 6.07–6.25 (3H, m, H-1′, H-13, H-14), 7.03 (1H, dd, $J = 10.2$, 10.2, H-15), 8.16 (1H, s, H-6), 8.47 (1H, t, $J = 5.4$), 11.62 (1H, s).

5-(Pent-4-ynoic acid methyl ester)-2′-deoxyuridine (4a). 5-iodo-2′-deoxyuridine, **1**, (354 mg, 1.0 mmol), palladium tetrakis(triphenylphosphine) (110 mg, 95 µmol), and copper iodide (46 mg, 158 µmol) were placed in a dried flask filled with argon. Anhydrous DMF (6.0 mL) was added *via* syringe, followed by triethylamine (0.4 mL, 2.9 mmol) and a solution of pent-4-ynoic acid methyl ester (5.7 ml, 5.1 mmol) in anhydrous DMF (5 mL). After 3 h stirring at room temperature, the solvent was removed *in vacuo*. The resulting brown oil was purified by flash column chromatography (10% MeOH in DCM), yielding 410 mg of orange oil (125%) (contamination by triethylammonium). δ_H(300 MHz; DMSO-d_6) 2.09 (2H, t, $J = 7.2$), 2.57 (4H, m), 3.61 (3H, s), 3.57 (2H, d, $J = 4.8$), 3.77 (1H, d, $J = 3.2$), 4.21 (1H, m), 5.09 (1H, t, $J = 5.1$), 5.25 (1H, d, $J = 4.3$), 6.09 (1H, t, $J = 6.7$), 8.10 (1H, s). δ_C(126 MHz, DMSO-d_6) 14.7, 32.6, 39.9, 51.5, 60.9, 70.1, 84.6, 87.5, 143.0. ESI-MS(pos): 449.5 m/z $[M + Na + Et_3N]^+$, calculated mass for $C_{15}H_{18}N_20_7$: 338.3.

5-(Pent-4-ynoic acid hexa-2,4-dienyl ester)-2′-deoxyuridine (5a). 5-iodo-2′-deoxyuridine, **1**, (395 mg, 1.05 mmol), palladium tetrakis(triphenylphosphine) (55 mg, 48 µmol), and copper iodide (28 mg, 121 µmol) were placed in a dried flask flushed with argon. Anhydrous DMF (5.0 mL) was added *via* syringe, followed by triethylamine (0.4 mL, 2.9 mmol) and a solution of pent-4-ynoic acid hexa-2,4-dienyl ester (1.0 g, 6.02 mmol) in anhydrous DMF (5 mL). After 22 h stirring at room temperature, the solvent was removed *in vacuo*. The resulting brown oil was purified by flash column chromatography using a gradient from 3–5% MeOH in DCM, yielding 158 mg of white powder (39%). δ_H(500 MHz, DMSO-d_6) 1.71 (3H, d, $J = 6.6$,), 2.09 (2H, dd, $J = 6.6$, 4.7,), 2.53–2.63 (4H, m), 3.54 (1H, ddd, $J = 11.8$, 5.0, 3.8,), 3.59 (1H, ddd, $J = 11.8$, 5.0, 3.6,), 3.78 (1H dd, $J = 6.6$, 3.5,), 4.19–4.23 (1H, m), 4.56 (1H, d, $J = 6.3$,), 5.07 (1H, t, $J = 4.9$,), 5.22 (1H, d, $J = 4.1$), 5.62 (1H, dt, J = 15.1, 6.6,), 5.72 (1H, dq, $J = 15.1$, 6.6,), 6.05 (1H, ddd, J = 15.1, 10.4, 1.6), 6.09 (1H, t, $J = 6.6$), 6.25 (1H, dd, J = 10.4, 15.1), 8.10 (1H, s), 11.55 (1H, s). δ_C(126 MHz, DMSO-d_6) 171.1 161.6, 149.4, 143.0, 134.0, 130.6, 130.5, 124.2, 98.6, 91.5, 87.5, 84.6, 73.1, 70.1, 64.3, 60.9, 39.7, 32.8, 17.8, 14.8. HRMS 403.15, calculated mass for $C_{20}H_{23}N_2O_7$: 403.15.

Triphosphorylation of nucleotides

Time-course on the monophosphorylation of 5-iodo-2′-deoxyuridine. 5-Iodo-2′-deoxyuridine (78 mg, 0.22 mmol) and 1-8-bis (dimethylamino) naphthalene (proton sponge) (71 mg, 0.33 mmol) were placed in a dried 2-neck flask, which was subsequently flushed with argon. Trimethyl phosphate (2.0 mL) was added *via* syringe, and the mixture was stirred at room temperature for at least 10 min until the compound was completely dissolved. A sample (1.5 µL) was withdrawn while maintaining a positive pressure of argon, and diluted with TEAB (0.1 M, 150 µL). The reaction mixture was then brought to the desired temperature using an ice bath (0 °C experiment) or an ice-salt bath (–13 °C experiment) and $POCl_3$ (20–60 µL, 0.22–0.66 mmol) was added *via* syringe. Samples were taken at regular intervals and TEAB was added within 10 sec as described to hydrolyse the phosphorodichloridate intermediate. The samples were analysed by HPLC. Peak areas of starting material and phosphorylated derivatives were normalised to the peak area of the proton sponge. δ_P(121.4 MHz; D_2O) –2.26. ESI-MS(neg): 433 m/z $[M - H]^-$, calculated mass for $C_9H_{12}IN_2O_8P$: 434.1.

General method for the triphosphorylation of nucleosides. The nucleoside (0.22 mmol) and proton sponge (71 mg, 0.33 mmol) were placed in a dried 2-neck flask, and dissolved in trimethyl phosphate (2.0 mL) The reaction mixture was cooled to between –13 °C and –20 °C, and $POCl_3$ (30 µL, 0.33 mmol) was added *via* syringe. After 20 min stirring, a suspension of tributylammonium pyrophosphate in anhydrous DMF (0.5 M, 2 mL) and tributylamine (0.16 mL) was added. After 45 min, the mixture was combined with TEAB (0.1M, 20 mL), and warmed up to room temperature. Volatile solvents were removed *in vacuo*, and the resultant residue purified by ion exchange chromatography (Sephadex A-25 DEAE) followed by HPLC.

5-Iodo-2′-deoxyuridine 5′-triphosphate triethylammonium salt (1a). Made from commercially available 5-iodo-2′-deoxyuridine. δ_H(400 MHz; D_2O) 1.22 (38H, t, $J = 7.4$), 2.34–2.39 (2H, m), 3.09 (24H, q, $J = 7.4$), 4.15–4.27 (3H, m), 4.62–4.67 (1H, m), 6.26 (1H, t, $J = 6.6$), 8.26 (1H, s) δ_C(126 MHz; D_2O) 39.3, 65.9, 71.0, 85.8, 145.9. δ_P(121.4 MHz; D_2O) –21.30 (1P, t, $J = 21.4$), –10.35 (1P, d, $J = 21.4$), –5.26 (1P, d, $J = 21.4$). ESI-MS(neg): 593 m/z $[M - H]^-$, calculated mass for $C_9H_{14}IN_2O_{14}P_3$: 594.0.

5-Ethynyl-2′-deoxyuridine 5′-triphosphate triethylammonium salt (2b). Made from **2a** (55 mg, 0.22 mmol). δ_H(400 MHz; D_2O) 1.28 (27 H, t, $J = 7.4$), 2.33–2.46 (2H, m), 3.20 (18H, q, $J = 7.4$), 3.60 (1H, s),4.19–4.23 (3H, m), 4.61–4.65 (1H, m), 6.28 (1H, t, $J = 6.7$), 8.22 (1H, s). ESI-MS(neg): 491 m/z $[M - H]^-$, calculated mass for $C_{11}H_{15}N_2O_{14}P_3$: 492.2.

5-(3-Amino-propynyl)-2′-deoxyuridine 5′-triphosphate triethylammonium salt (3c). Made from **3a**. The synthesis was performed as described above. After neutralisation by TEAB, 20 mL of concentrated ammonia was added to the reaction mixture to deprotect the amine. After 18 h stirring at room temperature, the product was purified in order to remove the acetamide protecting group. δ_H(500 MHz; D_2O) 1.26 (18H, t, $J = 7.4$), 2.29–2.43 (2H, m), 3.18 (12H, q, $J = 7.4$), 3.14 (2H, s), 4.14–4.26 (3H, m), 4.58–4.64 (1H, m), 6.28 (1H, t, $J = 6.2$), 8.31 (s, 1H). δ_P(121.4 MHz; D_2O) –21.38 (1P, t, $J = 21.4$), –10.40 (1P, d, $J = 21.4$), –0.05

(1P, d, J = 21.4). ESI-MS(neg): 520 m/z [M – H]$^-$, calculated mass for $C_{12}H_{18}N_3O_{14}P_3$: 521.2.

5-(Hexa-2,4-dienoic acid prop-2-ynylamide)-2′-deoxyuridine 5′-triphosphate triethylammonium salt (3e). Made from **3d**. δ_H(500 MHz; D_2O) 1.27 (27H, t, J = 7.4), 1.82 (3H, d, J = 6.0), 2.36 (1H, dd, J = 14.0, 6.3,), 2.41 (1H, ddd, J = 14.0, 6.3, 4.4), 3.18 (18H, q, J = 7.4,), 4.17–4.26 (5H, m), 4.61–4.66 (1H, m), 5.99 (1H, d, J = 15.1), 6.19–6.33 (3H, m), 7.14 (1H, dd, J = 15.1, 10.1), 8.17 (1H, s). ESI-MS(pos): 516.4 m/z [M + H]$^+$, calculated mass for $C_{18}H_{24}N_3O_{15}P_3$: 615.3.

5-(Pent-4-ynoic acid)-2′-deoxyuridine 5′-triphosphate triethylammonium salt (4b). Made from **4a**. The synthesis was done as described above. After neutralisation by TEAB, 20 mL of concentrated ammonia was added to the reaction mixture to deprotect the carboxylic acid. After 24 h stirring at 4 °C, the product was purified in order to remove the protecting group. δ_H(300 MHz; D_2O) 2.35 (2H, t, J = 6.9), 2.53 (2H, t, J = 6.9), 2.68 (2H, t, J = 6.9,), 4.16 (2H, d, J = 4.2), 4.60 (1H, m), 6.25 (1H, t, J = 6.6), 8.03 (1H, s). δ_C(126 MHz; D_2O) 15.8, 34.2, 39.0, 70.9, 85.9, 144.3. δ_P(121.4 MHz; D_2O) –21.6 (1P, m), –10.41 (1P, d, J = 18.8), –1.75 (1P, d, J = 18.8). ESI-MS(neg): 767.4 m/z [M + 2Et$_3$N + H]$^+$, calculated mass for $C_{14}H_{19}N_2O_{16}P_3$: 564.2.

5-(Pent-4-ynoic acid hexa-2,4-dienyl ester)-2′-deoxyuridine 5′-triphosphate triethylammonium salt (5b). Made from **5a**. δ_H(500 MHz; D_2O) 1.19 (27H, t, J = 7.3), 1.61 (3H, d, J = 6.6,), 2.26 (2H, dd, J = 14.0, 6.6,), 2.31 (1H, ddd, J = 14.0, 6.3, 3.9), 2.58–2.62 (2H, m); 2.63–2.67 (2H, m), 3.10 (18H, q, J = 7.3,), 4.07–4.15 (3H, m), 4.51–4.55 (1H, m), 4.58 (2H, d, J = 6.3,), 4.58 (1H, dt, J = 15.1, 6.6), 5.68 (1H, dq, J = 15.1, 6.6), 5.95 (1H, ddd, J = 15.1, 10.4, 1.6), 6.17 (1H, t, J = 6.6,), 6.20 (1H, dd, J = 15.1, 10.4), 7.92 (1H, s). δ_C(126 MHz; D_2O) 175.0, 164.8, 150.8, 144.2, 135.3, 132.5, 130.5, 123.7, 100.3, 94.1, 85.9, 85.8, 72.2, 70.9, 66.0, 46.9, 38.8, 33.4, 17.7, 15.5, 8.5. δ_P(121.4 MHz; D_2O) –21.69 (1P, t, J = 21.1), –10.17 (1P, d, J = 21.1), –8.75 (1P, d, J = 21.1). HRMS: 643.05, calculated mass for $C_{20}H_{26}N_2O_{16}P_3$: 643.05.

Peptide synthesis. The maleimide–peptide conjugates were synthesised using standard Fmoc solid-phase peptide synthesis (SPPS) on a Syro automated system using chloro-trityl resin pre-loaded with His, or Wang resin pre-loaded with Gly, and HBTU/DIPEA coupling chemistry. Four equivalents of coupling agent and amino acid were used in the coupling step of 40 min duration. Fmoc deprotection was carried out by shaking the resin in 40% piperidine for 3 min, followed by a second step in 20% piperidine for 10 min. Maleimide-β-alanine was added as the last amino acid using the same coupling chemistry. After completion of the synthesis, peptides were deprotected and cleaved from the resin by incubation in TFA/TES/H_2O (95:2.5:2.5 v/v/v) for 3 h. The peptide solution was collected and precipitated in diethyl ether at –20 °C and centrifuged to a pellet. The supernatant was removed and fresh diethyl ether was added. The pellet was broken up by sonication, the fragments supended in ether, and the slurry centrifuged to form a pellet. This purification process was repeated three times. Afterwards, the precipitate was dissolved in deionised water, frozen in a dry ice/acetone bath and freeze-dried overnight. The purity and chemical identity of the peptides was confirmed by HPLC and mass spectrometry, respectively.

Diels–Alder reaction of nucleotide triphosphate (5c). A solution of 100 mM of the maleimide-terminated peptide was prepared by dissolving lyophilised peptide in 500 mM Na_2HPO_4 buffer (final pH 6.1).[7] This solution (15 μL, 1.5 μmol) was added to the nucleotide solution (30 mM, 25 μL), and the reaction mixture was incubated at 25 °C for 2 h. LC-MS and ion exchange chromatography analysis indicated full conversion of the starting material for all reactions. The product was purified by ion exchange chromatography using a Resource Q column for the negatively charged products (His_6 tag or FLAG nucleotides) or a Resource S column for the positively charged product (NLS). The fractions containing products were lyophilised.

Purification gradient for His_6 nucleotide (R_1). Resource Q column at a flow rate of 2.5 mL/min using 0.1 M TEAB pH 7.5 for 5 mL followed by a gradient of 0.1 M to 0.45 M TEAB over 10 mL. HRMS: 1693.48, calculated mass for $C_{65}H_{80}N_{22}O_{27}P_3$: 1693.47.

Purification gradient for NLS nucleotide (R_2). Resource S column at a flow rate of 2.5 mL/min using 5 mM TEAB pH 7.5 for 5 mL followed by a gradient of 5 mM to 0.4 M TEAB over 10 mL. ESI-MS(neg): 1734.1 m/z [M – H]$^-$, calculated mass for $C_{69}H_{114}N_{18}O_{28}P_3$: 1735.7.

Purification gradient for FLAG nucleotide (R_3). Resource Q column at a flow rate of 2.5 mL/min using 0.45 M TEAB pH 7.5 for 5 mL followed by a gradient of 0.45 M to 0.8 M TEAB over 10 mL, followed by a step of 10 mL at 0.8 M.

Primer extension. Samples containing 150 pmol of DNA oligonucleotide primer of sequence 5′-ATG GGA CTA ACT AAT CTT TGC TTA-3′ and 170 pmol of template 5′-TAT TCC ATG TGA TAA GCA AAG ATT AGT TAG TCC CAT-3′ were dissolved in a polymerase extension mix. The mix contained 0.5 U Deep Vent$_R$® exo$^-$ (NEB), 1X Thermopol™ buffer (NEB), and 40 μM of natural dNTP nucleotides or dUTP derivatives in a total volume of 25 μL. The reaction mixtures were heated to 85 °C for 1 min to denature DNA duplexes, and incubated at 57 °C for 10 min to anneal and extend the primers, followed by immediate chilling on ice. The extension products were visualised by denaturing PAGE (18%) using 4–5 μL aliquots of the mixtures denatured by addition of formamide and subsequent heating at 95 °C.

Acknowledgements

This work has been supported by UCL Chemistry, UCL Business PLC, and the EPSRC (EP/D030005). We thank A. B. Tabor for use of a peptide synthesizer, and Hugh Martin for helping prepare illustrations.

References

1 G. M. Blackburn, M. J. Gait, D. Loakes and D. Williams, *Nucleic acids in chemistry and biology*, Royal Society of Chemistry, London, 2006.
2 V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem., Int. Ed.*, 2002, **41**, 2596–2599.
3 G. A. Burley, J. Gierlich, M. R. Mofid, H. Nir, S. Tal, Y. Eichen and T. Carell, *J. Am. Chem. Soc.*, 2006, **128**, 1398–1399.
4 J. Gierlich, G. A. Burley, P. M. E. Gramlich, D. M. Hammond and T. Carell, *Org. Lett.*, 2006, **8**, 3639–3642.
5 O. Diels and K. Alder, *Liebigs Annalen der Chemie*, 1928, **460**, 98–122.

6 D. C. Rideout and R. Breslow, *J. Am. Chem. Soc.*, 2000, **102**, 7816–7817.
7 V. Borsenberger and S. Howorka, *Nucleic Acids Res.*, 2009, **37**, 1477–1485.
8 H. A. Held and S. A. Benner, *Nucleic Acids Res.*, 2002, **30**, 3857–3869.
9 G. Giller, T. Tasara, B. Angerer, K. Muhlegger, M. Amacker and H. Winter, *Nucleic Acids Res.*, 2003, **31**, 2630–2635.
10 T. Tasara, B. Angerer, M. Damond, H. Winter, S. Dorhofer, U. Hubscher and M. Amacker, *Nucleic Acids Res.*, 2003, **31**, 2636–2646.
11 Z. Földes-Papp, B. Angerer, W. Ankenbauer and R. Rigler, *J. Biotechnol.*, 2001, **86**, 237–253.
12 O. Thum, S. Jäger and M. Famulok, *Angew. Chem., Int. Ed.*, 2001, **40**, 3990–3993.
13 T. Kottysch, C. Ahlborn, F. Brotzel and C. Richert, *Chem.–Eur. J.*, 2004, **10**, 4017–4028.
14 F. W. Hobbs, *J. Org. Chem.*, 1989, **54**, 3420–2422.
15 M. Hocek and M. Fojta, *Org. Biomol. Chem.*, 2008, **6**, 2233–2241.
16 P. Baumhof, N. Griesang, M. Bachle and C. Richert, *J. Org. Chem.*, 2006, **71**, 1060–1067.
17 L.-A. Fendt, I. Bouamaied, S. Thoeni, N. Amiot and E. Stulz, *J. Am. Chem. Soc.*, 2007, **129**, 15319–15329.
18 T. Nguyen, A. Brewer and E. Stulz, *Angew. Chem., Int. Ed.*, 2009, **48**, 1974–1977.
19 S. I. Khan, A. E. Beilstein and M. W. Grinstaff, *Inorg. Chem.*, 1999, **38**, 418–419.
20 N. Amann, E. Pandurski, T. Fiebig and H. A. Wagenknecht, *Angew. Chem., Int. Ed.*, 2002, **41**, 2978–2980.
21 E. Mayer-Enthart and H. A. Wagenknecht, *Angew. Chem., Int. Ed.*, 2006, **45**, 3372–3375.
22 S. Jäger and M. Famulok, *Angew. Chem., Int. Ed.*, 2004, **43**, 3337–3340.
23 V. R. Sirivolu, P. Chittepu and F. Seela, *ChemBioChem*, 2008, **9**, 2305–2316.
24 J. He and F. Seela, *Nucleic Acids Res.*, 2002, **30**, 5485–5496.
25 L. H. Thoresen, G. S. Jiao, W. C. Haaland, M. L. Metzker and K. Burgess, *Chem,–Eur. J.*, 2003, **9**, 4603–4610.
26 P. Capek, R. Pohl and M. Hocek, *Org. Biomol. Chem.*, 2006, **4**, 2278–2284.
27 M. Vrabel, M. Hocek, L. Havran, M. Fojta, I. Votruba, B. Klepetarova, R. Pohl, L. Rulisek, L. Zendlova, P. Hobza, I. H. Shih, E. Mabery and R. Mackman, *Eur. J. Inorg. Chem.*, 2007, 1752–1769.
28 H. Weizman and Y. Tor, *J. Am. Chem. Soc.*, 2002, **124**, 1568–1569.
29 V. Borsenberger, N. Mitchell and S. Howorka, *J. Am. Chem. Soc.*, 2009, **131**, 7530–7531.
30 M. Yoshikawa, T. Kato and T. Takenishi, *Tetrahedron Lett.*, 1967, 5065–5068.
31 M. Yoshikawa, T. Kato and T. Takenishi, *Bull. Chem. Soc. Jpn.*, 1969, **42**, 3505–3508.
32 K. Burgess and D. Cook, *Chem. Rev.*, 2000, **100**, 2047–2059.
33 J. L. Ruth and Y. C. Cheng, *Mol. Pharmacol.*, 1981, **20**, 415–422.
34 J. Ludwig, in *Biophosphates and their analogues- Synthesis, structure, metabolism and activity*, eds. K. S. Bruzik and W. J. Stec, Elesevier, Amsterdam, first edn., 1987, pp. 131–133.
35 T. Sowa and S. Ouchi, *Bull. Chem. Soc. Jpn.*, 1975, **48**, 2084–2090.
36 W. H. Dawson, R. L. Cargill and R. B. Dunlap, *J. Carb. Nucleos. Nucl.*, 1977, **4**, 363–375.
37 A. Collier and G. Wagner, *Org. Biomol. Chem.*, 2006, **4**, 4526–4532.
38 J. M. Campagne, J. Coste and P. Jouin, *J. Org. Chem.*, 1995, **60**, 5214–5223.
39 I. Diala, S. Murao and M. Fujii, *Nucleic Acids Symposium Series*, 2008, **52**, 679–680.
40 S. Doublie, S. Tabor, A. M. Long, C. C. Richardson and T. Ellenberger, *Nature*, 1998, **391**, 251–258.
41 M. C. Franklin, J. M. Wang and T. A. Steitz, *Cell*, 2001, **105**, 657–667.
42 U. Hubscher, G. Maga and S. Spadari, *Annu. Rev. Biochem.*, 2002, **71**, 133–163.
43 K. Kincaid, J. Beckman, A. Zivkovic, R. L. Halcomb, J. W. Engels and R. D. Kuchta, *Nucleic Acids Res.*, 2005, **33**, 2620–2628.
44 J. Ludwig and F. Eckstein, *J. Org. Chem.*, 1988, **54**, 631–635.