ARTICLE

OBC
www.rsc.org/obc

Synthesis and transcription studies on 5′-triphosphates derived from 2′-*C*-branched-uridines: 2′-homouridine-5′-triphosphate is a substrate for T7 RNA polymerase

John B. J. Pavey,[a] Anthony J. Lawrence,[a] Ian A. O'Neil,[a] Stefan Vortler[b] and Richard Cosstick *[a]

[a] *Robert Robinson Laboratories, Department of Chemistry, University of Liverpool, Crown Street, Liverpool, UK L69 7ZD. E-mail: rcosstic@liv.ac.uk*

[b] *Biochemie, Universität Bayreuth, Universitätsstrasse 30, D-95440 Bayreuth, Germany*

Received 11th November 2003, Accepted 16th January 2004
First published as an Advance Article on the web 18th February 2004

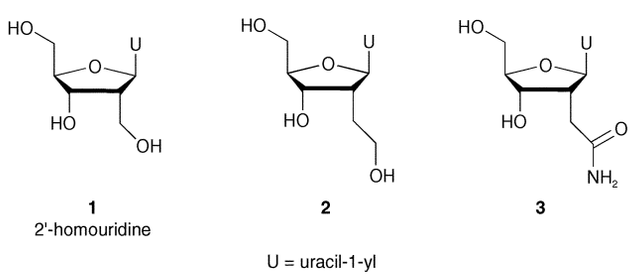
The 5′-triphosphates of 2′-hydroxymethyluridine (2′-homouridine) and 2′-hydroxyethyluridine were prepared from the corresponding acetyl-protected nucleosides by initial phosphitylation with 2-chloro-5,6-benzo-1,2,3-dioxaphosphorin-4-one. 2′-Acetamidouridine 5′-triphosphate was prepared in an analogous fashion from uridine 2′-*C*-, 3′-*O*-γ-butyrolactone, in which the 3′-hydroxyl group is internally protected as the lactone. Subsequent treatment with ammonia gave the required acetamido triphosphate. All three triphosphates were investigated as substrates for T7 RNA polymerase and a Y639F mutant of this enzyme. 2′-Homouridine triphosphate was found to be a substrate for the wild-type enzyme in the presence of manganese and was specifically incorporated into short RNA transcripts (20 and 21 nucleotides in length). The presence of the analogue within the transcripts was confirmed through its resistance to alkaline hydrolysis. Gel electrophoretic analysis also showed that 2′-homouridine could be multiply incorporated into a transcript with a length of 75 nucleotides. This is the first report of a 2′-deoxy-2′-α-*C*-branched nucleoside 5′- triphosphate acting as a substrate for T7 RNA polymerase. The 2′-hydroxyethyl- and 2′-acetamido – uridine triphosphates were not substrates for the enzymes.

Introduction

In vitro selection techniques which enable massive parallel sampling of DNA/RNA sequences have emerged as powerful tools for the generation and isolation of catalytically-active nucleic acids (DNAzymes/ribozymes).[1] Interestingly, these methods have extended the repertoire of nucleic acid-catalysed reactions beyond those achieved by natural ribozymes and include nucleic acids that are capable of accelerating nucleotide synthesis[2] and RNA polymerisation.[3] However, in comparison to proteins, the catalytic potential of nucleic acids is restricted due to the limited chemical functionality of the 4 natural nucleotide monomers. To overcome this deficiency, nucleoside triphosphates equipped with diverse functional groups have been studied as substrates for DNA and RNA polymerases. Using this approach, more than 20 nucleotides,[1a] functionalised on the non-Watson-Crick face of the bases have been incorporated into DNA[4] or RNA,[5] with in many cases, the successful outcome of enhancing their catalytic repertoire.[6]

In the case of RNA, functionalisation at the 2′-position of sugar is also an attractive site for modification for the following reasons: (i) it can be incorporated uniformly into purine and pyrimidine nucleosides alike; (ii) it is known to increase the stability of RNA to both enzymatic and chemical degradation;[7,8] (iii) it can be used to probe RNA structure-function relationships.[9,10] To date, the only sugar-modified nucleoside analogues to be successfully incorporated into RNA by template-directed RNA polymerases are deoxyribo-,[11] 2′-NH_2,[7,12,13] 2′-F,[7,12,13] 2′-OMe[13–15] 2′-N_3[14] and 2′-SH[10] derivatives. There has been no report of the enzymatic incorporation of a ribonucleotide analogue in which the functionality is attached to a 2′-carbon chain (2′-*C*-branched analogues), even though this is potentially the most versatile strategy for 2′-modification. Additionally, the 2′-*C*-branched alcohols described here, retain the functionality of the natural substrates and are interesting probes for investigating the interaction between RNA polymerases and their substrates.

We have previously reported the synthesis of many 2′-deoxy-2′-α-*C*-branched uridines including the hydroxymethyl (2′-homouridine) (**1**), the hydroxyethyl (**2**) and the acetamido (**3**) derivatives.[16,17] In model studies, dinucleoside monophosphates containing these analogues demonstrate significantly increased stability to both enzymatic and chemical degradation of the phosphodiester linkage.[17,18] We now describe the synthesis of the 5′-triphosphates derived from nucleosides **1**–**3** and studies on their incorporation into RNA transcripts. In particular, we show for the first time that a 2′-deoxy-2′-α-*C*-branched nucleoside 5′-triphosphate acts as a substrate for T7 RNA polymerase.



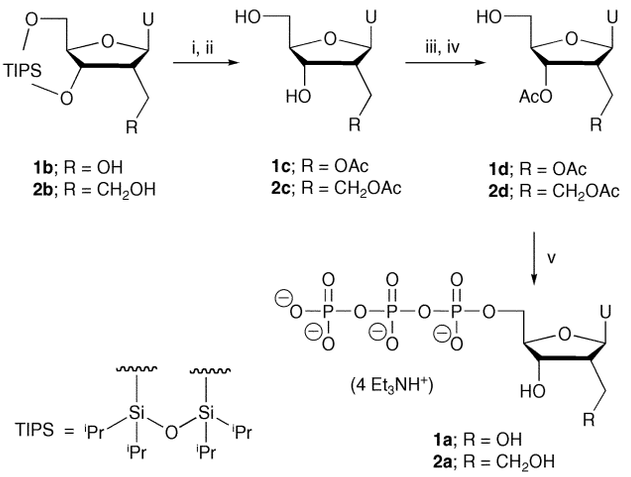
U = uracil-1-yl

Results and discussion

Synthesis of triphosphates

We have previously reported the 3′,5′-tetraisopropyldisiloxanediyl (TIPS)-protected derivatives of both 2′-deoxy-2′-α-*C*-hydroxymethyluridine (**1b**)[17] and 2′-deoxy-2′-α-*C*-(2-hydroxyethyl)uridine (**2b**)[16,18] which are now used as convenient intermediates for the synthesis of the triphosphates. Thus, hydroxy nucleosides **1b** and **2b** were protected as their acetyl derivatives and the TIPS protecting groups removed using $Et_3N{\cdot}3HF$ to yield the diols **1c** and **2c**, respectively (Scheme 1). The 5′-hydroxyl function of both **1c** and **2c** was transiently

DOI: 10.1039/b314348a

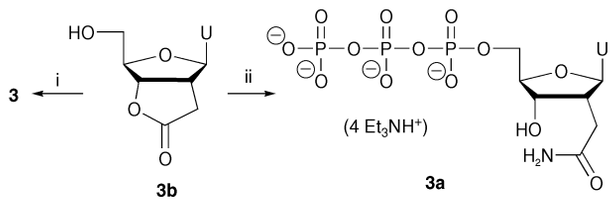


Scheme 1 *Reagents and conditions*: (i) acetic anhydride, pyridine, 6 h; (ii) $Et_3N{\cdot}3HF$, 16 h, 57% (**1c**), 80% (**2c**); (iii) dimethoxytrityl chloride, CH_2Cl_2, pyridine, 2 h, 80–85%; (iv) a) Ac_2O, pyridine, 16 h. b) AcOH, 1 h, 85% (**1d**), 94% (**2d**); (v) a) 2-chloro-5,6-benzo-1,2,3-dioxaphosphorin-4-one, pyridine, dioxane, DMF, 10 min; b) bis(tri-*n*-butylammonium) pyrophosphate, DMF, tri-*n*-butylamine, 10 min; c) I_2 (excess), pyridine, H_2O (98 : 2), 15 min; d) conc aq. ammonia, 1 h.

protected as the dimethoxytrityl (DMT) ether, to allow acetylation of the less reactive 3′-hydroxyl function. The DMT protecting group was subsequently removed from both compounds to give the 5′-hydroxy nucleosides **1d** and **2d**, which would serve as the substrates for triphosphate synthesis.

Triphosphates **1a** and **2a** were synthesised by phosphitylation with 2-chloro-5,6-benzo-1,2,3-dioxaphosphorin-4-one according to the method of Ludwig and Eckstein[19] (Scheme 1) and purified by elution from a DEAE Sephadex column, using a gradient of triethylammonium bicarbonate. Following evaporation and co-evaporation of the volatile buffer salts with water, the triphosphates, **1a** and **2a**, were isolated as their triethylammonium salts in yields of 47% and 56%, respectively. In both cases the spectroscopic data was fully consistent with the proposed structures.

We have previously reported the synthesis of lactone (**3b**) and its subsequent ring opening using concentrated ammonia to give the amide **3** (Scheme 2).[15,20] As the final step of the triphosphate synthesis involves treatment with concentrated ammonia to remove the acetyl protecting groups, it was anticipated that the lactone could be used as a masked form of the amide in the synthesis of the triphosphate **3a**. Thus, following subjection of the lactone **3b** to the triphosphate synthesis and ion-exchange chromatography, the acetamido triphosphate (**3a**) was obtained in 54% yield. To further confirm their identity, the triphosphates **1a**, **2a** and **3a** were hydrolysed back to the parent nucleosides, **1**, **2** and **3** respectively, by treatment with alkaline phosphatase and compared to authentic samples by HPLC.



Scheme 2 *Reagents and conditions*: (i) excess conc. NH_3, 15 min, 100%; (ii) see (v) Scheme 1.

The triphosphates **1a**, **2a**, **3a** had all been synthesised from uridine, and thus to remove minute quantities of uridine triphosphate (UTP) which was possibly present as a trace contaminant, the triphosphates were treated with sodium periodate[21] and subsequently repurified. Periodate treatment oxidises the ribose sugar of nucleosides/nucleotides,[22] but showed no detectable reactivity towards the 2′-*C*-branched triphosphates.

Transcription studies

Milligan *et al*. have previously reported the efficient transcription of short oligoribonucleotides from DNA templates using T7 RNA polymerase.[23] Based on Milligan's work two short DNA templates (U15 and U18, Fig. 1), each having the same T7 promoter sequence, were designed to produce transcripts of 22 and 21 nucleotides, respectively, each with a single uridine residue at the position indicated. It has been shown previously that the sequence of the first 8–10 nucleotides of the transcript has an influence on the yield of the full-length transcript, such that synthesis is preferentially aborted immediately following the incorporation of uridine.[23] To minimise prematurely aborted transcripts, the templates were designed to position the single uridine residue after the initial 10 nucleotides.

Initial transcription experiments were performed with T7 RNA polymerase under standard conditions (see Experimental section) using all four natural ribonucleosides triphosphates (NTPs), as a positive control. Reactions were also performed in the absence of UTP, which served as a negative control to assess whether base-misincorporation occurred. Transcripts were internally-labelled using [α-^{32}P]CTP and the products analysed using 20% denaturing polyacrylamide gel electrophoresis (PAGE). Lanes 1 and 2 (Fig. 2A) show the positive control (with UTP) for transcription with template U15 in the presence of Mg^{2+} and Mn^{2+}, respectively. Manganese was investigated as it has previously been shown to decrease the selectivity of the enzyme, potentially aiding the incorporation of analogues.[13,24,25] In both lanes it can be seen that a transcript of the expected length (22 nucleotides) was observed, together with longer products corresponding to extraneous addition of non-template-coded nucleotides to the 3′-end and shorter products resulting from aborted synthesis. Both these longer (up to full length transcript +5 residues) and shorter products were

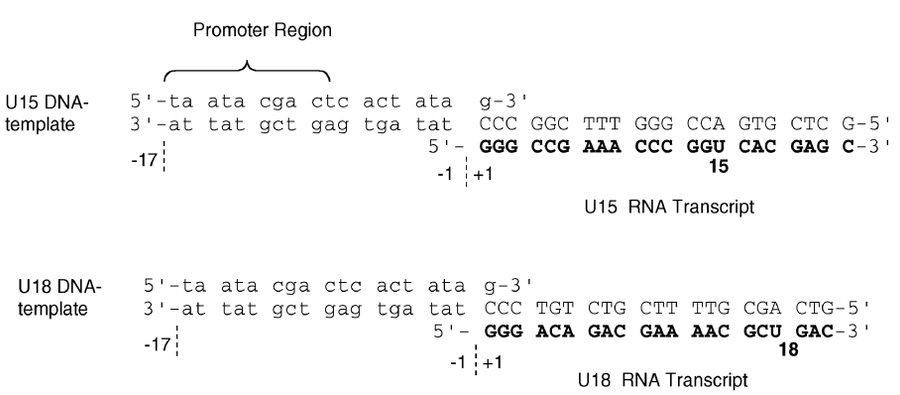


Fig. 1 Diagram of synthetic DNA templates used in transcription experiments, nucleotides in promoter region in lower case, template region and transcript in upper case. RNA transcript in bold type.

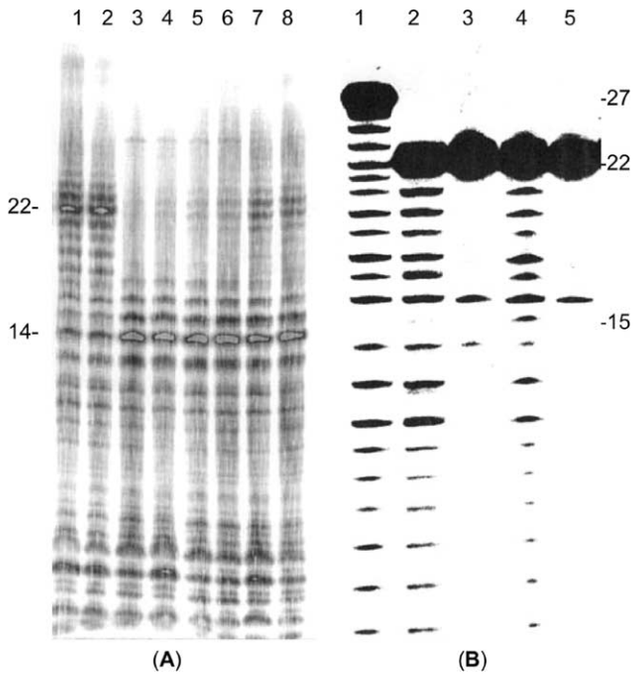


Fig. 2 (**A**) Autoradiograph of PAGE of [α-^{32}P]CTP-internally-labelled transcripts from the U15 template. Lanes 1 and 2; transcription with all four NTPs (positive control) in the presence of Mg^{2+} and Mn^{2+}, respectively. Lanes 3 and 4; transcription without UTP (negative control) in the presence of Mg^{2+} and Mn^{2+}, respectively. Lanes 5 and 6, transcription with 2′-homouridine triphosphate (2 mM, lane 5 and 4 mM, lane 6) in the presence of Mg^{2+}. Lanes 7 and 8, transcription with 2′-homouridine triphosphate (2 mM, lane 7 and 4 mM, lane 8) in the presence of Mn^{2+}. (**B**) Autoradiograph of PAGE of ^{32}P end-labelled transcripts of the U15 template. Lane 1, partial alkaline hydrolysis of *n*+5 transcript, obtained using 2′-homouridine triphosphate. Lane 2, partial alkaline hydrolysis of full-length transcript, obtained using 2′-homouridine triphsphate. Lane 3, marker for position 15, generated by iodine treatment of full-length transcript obtained using [α-S]UTP in place of UTP. Lane 4, marker lane generated from partial alkaline hydrolysis of full-length transcript, obtained with unmodified UTP. Lane 5, full length transcript produced using [α-^{32}P]UTP prior to treatment with iodine.

expected based on the previous studies of Milligan,[23] but the mechanism by which the additional nucleotides are added to the 3′-terminus is not understood. Negative controls (no UTP) performed with both Mg^{2+} and Mn^{2+} (Fig. 2A, lanes 3 and 4, respectively) showed major products corresponding to termination immediately prior to the site of uridine incorporation (position 14) and negligible production of full-length transcripts; indicating that misincorporation was not occurring to a significant extent.

Transcription with the 2′-*C*-branched nucleosides initially concentrated on 2′-homouridine triphosphate (**1a**) as this represents the minimum perturbation from the natural substrate. Reactions were run in the presence of both Mg^{2+} and Mn^{2+} and with varying concentrations (see legend to Fig. 2A) of 2′-homouridine triphosphate. In the presence of Mg^{2+} only trace amounts of full-length transcript were observed and the majority of the product corresponded to termination at position 14 (Fig. 2A, lane 5 and 6). However, when the magnesium was replaced with 2.5 mM Mn^{2+}, the quantity of full-length product was greatly increased (Fig. 2A, lanes 7 and 8), indicating that 2′-homouridine was successfully incorporated into the transcript. This is in accord with previous observations that the use of Mn^{2+} decreases the selectivity of the enzyme. The yield of the full-length transcript containing 2′-homouridine was estimated to be about 10–20% (based on intensity of gel bands) of that obtained with UTP. Essentially identical results were obtained when the transcription studies were performed with the U18 template, although these results have not been shown.

Our previous studies have established that 2′-homouridine, when chemically incorporated into a RNA oligomer, does not behave like a classic ribonucleoside, and a dinucleoside monophosphate containing this residue was shown to be almost 2000-fold more resistant to base-catalysed cleavage than the parent dimer UpU.[17] The increased stability of a phosphodiester bond adjacent to the hydroxymethyl group, suggested a method to verify that 2′-homouridine had been successfully incorporated into the U15 transcript. Partial alkaline hydrolysis of an end-labelled U15 transcript, followed by analysis by PAGE, should produce a hydrolysis ladder with a band missing at the site of incorporation of 2′-homouridine.

To execute this analysis, unlabelled transcripts from the U15 template were prepared using both homouridine triphosphate and, for comparison, unmodified UTP, and the full-length products were purified by PAGE. The 5′-terminal triphosphate group from both transcripts was removed using calf intestine alkaline phosphatase and the resulting products end-labelled using polynucleotide kinase and [γ-^{32}P]ATP.[26] Using this method it was possible to isolate end-labelled transcripts corresponding to the full-length transcript (n), which was contaminated by the product with an additional unidentified nucleotide added at the 3′-end (n + 1 transcript). Surprisingly, a transcript with 5 additional nucleotides at the 3′-end (n + 5 transcript) was also isolated. Alkaline hydrolysis of the various labelled transcripts was performed in 50 mM $NaHCO_3$ for 5 min at 95 °C. The resulting hydrolysates were analysed by PAGE (Fig. 2B). Both the full length (lane 2) and n + 5 (lane 1) modified transcripts showed a missing band at the 15-position indicating that 2′-homouridine had been successfully incorporated. In comparison, the partial hydrolysis of the unmodified transcript (Fig. 2B, lane 4) shows the expected band at position 15.

Further evidence that the missing band was at the required position was obtained through an experiment similar to the positive control, in which 0.1 mM [α-*S*]-uridine thiotriphosphate was included with the transcription reaction. This results in a small proportion (~10%) of the transcripts having a phosphorothioate linkage to the 5′-side of the uridine at the 15-position. As a small proportion (about 5%) of the phosphorothioate linkage undergoes oxidative cleavage in the presence of iodine,[27] the end-labelled phosphorothioate transcript treated with iodine, gave a weak cleavage band adjacent (position 14) to the "missing" band (lane 3). Coincidentally, before iodine treatment the labelled phosphorothioate transcript was shown to have an unintentional degradation band at position 16 (lane 5). The relative positions of the phosphorothioate marker band and the "missing" band in the 2′-homouridine transcript can be explained by the fact that the phosphorothioate linkage is to the 5′-side of the uridine whereas the 2′-modified nucleoside stabilises the oligonucleotide to the 3′-side.

Having established that the hydroxymethyl derivative had been successfully incorporated into the transcript, studies were performed to further optimise the transcription efficiency for this analogue. It was found that the most efficient transcription was observed in the presence of 2.5 mM Mn^{2+} with a DNA template concentration of 1 μM, 10 units per μl of T7 RNA polymerase and 2–4 mM 2′-homouridine triphosphate. Using these optimised conditions the multiple incorporation of 2′-homouridine into a longer transcript was investigated using a plasmid containing the *E. coli* $tRNA^{Asp}$ gene as a template, which produces a transcript containing a total of 75 nucleotides including 17 uridine residues.[12] Using this template, transcription was examined over a range of 2′-homouridine triphosphate concentrations (Fig. 3, lanes 4–7). The full-length transcript can clearly be seen in lanes 5 and 6, (concentrations of 2 and 4 mM 2′-homouridine triphosphate, respectively), although the amount of full-length product is very low, about 5–10% of that obtained for the UTP positive control (lane 1). The major products from transcription with 2′-homouridine

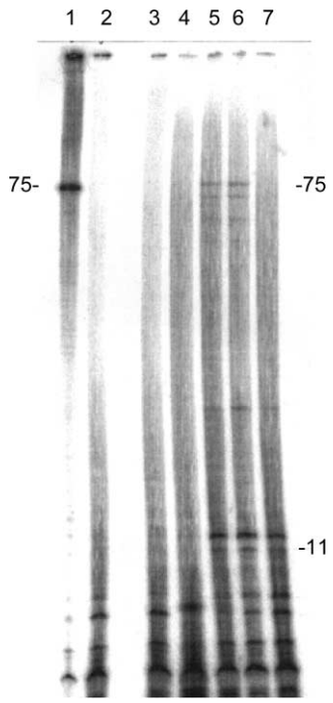


Fig. 3 Autoradiograph of PAGE of [α-^{32}P]CTP-internally-labelled transcripts from the tRNA^{Asp} template. Lane 1; positive control in the presence of Mg^{2+}. Lane 2, negative control in the presence of Mg^{2+}. Lane 3, negative control in the presence of Mn^{2+}. Lanes 4–7, transcription performed with varying concentrations of 2′-homouridine (1, 2, 4 and 8 mM, respectively) in the presence of Mn^{2+}.

triphosphate result from abortion of transcription, particularly at positions 11 and 12 where there are two consecutive uridine residues. Additionally, a band just below the 75-nucleotide full-length product is clearly present and most probably results from termination caused by two consecutive uridine residues at positions 69 and 70. Once again the optimum concentration for the modified triphosphate was shown to be 2–4 mM: at 1 mM (Fig. 3, lane 4) no discernable full-length product was present, whilst at 8 mM (Fig. 3, lane 7), the analogue appeared to inhibit the enzyme.

The optimised transcription conditions were also used to investigate the incorporation of the hydroxyethyl and acetamido triphosphates with the U15 and U18 templates. When analysed, transcription reactions with these analogues gave PAGE autoradiographs which were essentially indistinguishable from the negative controls (no UTP). Additionally, a mutant T7 RNA polymerase in which tyrosine 639 is replaced by phenylalanine (Y639F) and that utilises either ribo- or deoxyribonucleoside triphosphates to synthesise RNA, DNA or 'transcripts' of mixed dNMP/rNMP composition,[11] was unable to utilise any of the 2′-*C*-branched uridine triphosphates.

It is interesting to speculate as to which structural features are most important in determining the suitability of a 2′-modified triphosphate as a substrate T7 RNA polymerase. Studies with the Y639F mutant clearly indicate that the hydroxyl group of the tyrosine residue 639 is involved in discriminating between ribo- and deoxyribo- sugars, probably through hydrogen-bonding to the 2′-OH[28] or by providing a steric gate.[29] In this respect it was appealing to compare the incorporation efficiency of the hydroxymethyl and hydroxyethyl analogues to that of UTP. Given that the hydroxyethyl analogue could not be incorporated, it seems apparent that the size of the 2′-substiuent has a major influence on determining substrate suitability, either through a steric blocking effect or by incorrectly positioning the hydroxyl group on the 2′-*C*-alkyl chain for interaction with the tyrosine residue. As an additional consideration, the 2′-substituent also has an important effect on the equilibrium between the north and south conformations of the ribose sugar. It is known that for ribonucleosides, the sugar pucker is predominantly populated by the north ($C_{3'}$-*endo*/$C_{2'}$-*exo*) conformation to the extent of about 60% and that this geometry favours an A-type helix.[30] For all practical purposes the percentage of the south conformer present in nucleoside phosphates in aqueous solution can be calculated using the simple equation %S = $10J_{1'-2'}$.[31] The large $J_{1'-2'}$ coupling constants (range 8.6–9.5 Hz) for the 2′-*C*-branched triphosphates establish that in these compounds, the sugar pucker is very largely populated (~90%) by the south conformation, which is very different from the natural substrates. Multiple incorporation of 2′-*C*-branched nucleotides would disfavour the A-type helix for the transcript-template duplex and may account for the termination of synthesis observed where consecutive uridine residues are present in the tRNA^{Asp} transcript.[13] However, previous studies indicate that sugar conformation alone, is not crucial in determining substrate suitability and 2′-aminouridine triphosphate, which predominantly adopts the south conformation, is a better substrate than the predominantly north 2′-fluorouridine.

In conclusion, our results clearly demonstrate that the homouridine is incorporated into RNA transcripts using T7 RNA polymerase and is the first report of a 2′-*C*-branched nucleotide acting as a substrate for T7 RNA polymerase. The results also suggest that the size of the 2′-substiuent is more important than its functionality. In comparison to uridine, 2′-homouridine has a primary and potentially more reactive hydroxyl group, but produces a vicinal phosphodiester bond that is much more resistant to both enzyme- and base-catalysed cleavage. This analogue could be of considerable utility in probing specific interactions involving the 2′-OH groups of RNA. Very recently, the introduction of a second mutation into the Y639F background, in which histidine 784 is replace by alanine, has been shown to enhance the utilisation of nucleoside triphosphates with bulkier 2′-substituents, allowing the efficient incorporation of both 2′-*O*-methyl and 2′-azido analogues.[14] Although we have not been able to study this mutant with our substrates, these results suggest that it should be possible to prepare transcripts with a wider range of 2′-funtionalised nucleosides in the future.

Experimental

FAB mass spectra were recorded on a VG Analytical 7070E mass spectrometer operating with a PDP 11/250 data system and an Ion Tech FAB ion gun working at 8 Kv. High resolution FAB mass spectra were obtained on either the above instrument or a VG ZAB/E spectrometer at the EPSRC Mass Spectrometry Service Centre (Swansea UK). 3-Nitrobenzyl alcohol was used as a matrix unless stated otherwise. NMR spectra were recorded at the field strength indicated and chemical shifts are given in ppm down field from an internal standard of tetramethylsilane. For TLC analysis, nucleosides were visualised either as a black spot by spraying with a solution of 5% (v/v) sulfuric acid and 3% (w/v) phenol in ethanol and charring at 120 °C. Pyridine was dried by heating to reflux over calcium hydride for 2–3 h and was then distilled under atmospheric pressure. THF and dioxane were dried by heating to reflux with sodium benzophenone until the purple colouration persisted and then distilled under atmospheric pressure. Dimethylformamide was purchased anhydrous from Aldrich in Sure-seal™ bottles.

HPLC was conducted using a Varian Star 9010 liquid chromatograph equipped with a Varian Star 9050 variable wavelength UV-VIS detector recording at 260nm. Unless stated otherwise, analyses were performed on a Nucleosil C_{18} reverse phase column, with a flow rate of 1 mL/min. Data were recorded using a Varian 4400 recording integrator. Nucleosides were analysed using a gradient of 0 ⟶ 10% MeCN in 50 mM potassium phosphate (pH 6.5) over 20 min. Nucleoside triphosphates were analysed using a gradient of 0 ⟶ 30% MeOH

in 50 mM potassium phosphate (pH 6.5), containing 8 mM tetrabutylammonium hydrogen sulfate (ion-pair buffer) over 20 min.

The U15 and U18 primer-templates were synthesised on an Applied Biosystems DNA synthesiser, annealed (by heating to 80 °C in10 mM Tris-HCl, pH 8.0, 1 mm EDTA and 50 mM NaCl and slow cooling to room temperature) and stored at a final concentration of 50 μM. T7 RNA polymerase,[12] the Y639F mutant T7 RNA polymerase [14] and plasmid AspUC19 [12] containing the *E. coli* $tRNA^{Asp}$ gene were used in the laboratory of F. Eckstein (Göttingen, Germany) and were isolated as previously described. T4 Polynucleotide kinase (PNK) (30 units μL^{-1}) was purchased from U.S. Biochemicals and supplied with reaction buffer. Calf intestinal alkaline phosphatase was purchased from New England BioLabs (10 units μL^{-1}). Microcon microconcentrators were purchased from Amicon. Gel Loading Buffer contained 8 M urea, 0.09 M Tris-borate at pH 8.3, 2.5 mM Na_2EDTA, 0.1% xylene cyanol FF and 0.1% bromophenol blue. Imaging of gels was performed on a Fuji BAS 2000 Bio-Imaging analyzer, using Fuji Bas-II plates. The iodine induced cleavage of transcripts containing phosphorothioate linkages was performed as previously described.[27]

2′-Deoxy-2′-α-*C*-(acetoxymethyl)uridine (1c) and 2′-deoxy-2′-α-*C*-(2-acetoxyethyl)uridine (2c). Hydroxy nucleoside **1b** or **2b** (2.3 mmol), in anhydrous pyridine (20 mL) was treated with acetic anhydride (2.16 mL, 23 mmol). After 16 h the solvent was removed *in vacuo* and residual pyridine removed by coevaporation with toluene (2 × 20 mL). The residue was taken up in CH_2Cl_2 (50 mL) and washed with saturated aqueous sodium bicarbonate solution (2 × 30 mL) and brine (2 × 20 mL). The aqueous washings were combined and back extracted with CH_2Cl_2 (2 × 20 mL). The combined organic extracts were dried ($MgSO_4$) and the solvent removed *in vacuo*. Column chromatography of the residue over silica gel eluting with CH_2Cl_2 and methanol (0–5%) afforded the acetylated nucleoside as a white amorphous solid. To remove the TIPS group, the nucleoside (0.45 mmol) was dissolved in dry THF (30 mL) and treated with $NEt_3 \cdot 3HF$ (0.08 mL, 5 mmol), the solution was left for 8–16 h and monitored by TLC eluting with CH_2Cl_2–MeOH (10%). The solution was concentrated *in vacuo*, diluted with methanol (20 mL) and evaporated onto silica gel, the product was subsequently isolated by column chromatography (CH_2Cl_2 containing an increasing gradient of MeOH from 0–10%) as a colourless oil or white foam.

2′-Deoxy-2′-α-*C*-(acetoxymethyl)uridine (1c). 45% Yield; anal. found C, 47.68, H, 5.79, N, 8.77, $C_{12}H_{16}N_2O_7.1/2\ CH_3OH$ requires C, 47.45, H, 5.74, N, 8.86%. Found FAB HRMS *m/z* $(M + H)^+$, 301.1034. $C_{12}H_{17}N_2O_7$ requires $(M + H)^+$, 301.1035; ^{1}H NMR (400 MHz; D_3COD) δ 7.97 (1H, d, J = 8.1 Hz, H6), 6.15 (1H, d, J = 8.6 Hz, H1′), 5.72 (1H, d, J = 8.1 Hz, H5), 4.38–4.34 (2H, m, H3′ and H6′), 4.22 (1H, dd, J = 8.1, 11.4 Hz, H6″), 3.99–3.98 (1H, m, H4′), 3.74–3.70 (2H, m, 2 × H5′), 2.66–2.58 (2H, m, H2′), 1.93 (3H, s, CH_3); ^{13}C NMR (75.5 MHz; D_3COD) δ 172.55 (CO), 166.03 (C4), 152.34 (C2), 142.50 (C6), 103.07 (C5), 89.10 (C1′), 88.56 (C4′), 73.37 (C3′), 63.19 (C5′), 62.11 (C6′), 49.17 (C2′), 20.67 (CH_3); *m/z* (FAB^+) 301 ($[M + H]^+$, 20%), 189 ($[M -$ uracil$]^+$, 22), 129 (100).

2′-Deoxy-2′-α-*C*-(2-acetoxyethyl)uridine (2c). 77% Yield; found FAB HRMS *m/z* $(M + H)^+$, 315.1202 $C_{13}H_{19}N_2O_7$ requires $(M + H)^+$, 315.1192; ^{1}H NMR (400 MHz; $CDCl_3$) δ 7.99 (1H, d, J = 8.1 Hz, H6), 6.06 (1H, d, J = 9.1 Hz, H1′), 5.74 (1H, d, J = 8.1 Hz, H5), 4.28 (1H, dd, J = 4.9, 1.0 Hz, H3′), 4.27–4.07 (2H, m, 2 × H7′), 4.00 (1H, dt, J = 3.3, 1.4 Hz, H4′), 3.30 (2H, s, 2 × H5′), 2.42–2.38 (1H, m, H2′), 2.08–2.01 (1H, m, H6′), 2.01 (3H, s, CH_3), 1.74–1.69 (1H, m, H6″); ^{13}C NMR (75.5 MHz; $CDCl_3$) δ 172.77 (CO), 165.77 (C4), 152.48 (C2), 142.42 (C6), 103.23 (C5), 89.51 (C1′), 88.84 (C4′), 73.66 (C3′), 63.66, 63.34 (C7′ + C5′), 46.60 (C2′), 24.36 (C6′), 20.75 (CH_3); *m/z* (FAB^+) 315 ($[M + H]^+$, 20%), 102 (100).

Preparation of 2′-deoxy-2′-α-*C*-(acetoxymethyl)-3′*O*-acetyluridine (1d) and 2′-deoxy-2′-α-*C*-(2-acetoxyethyl)-3′*O*-acetyluridine (2d). The diol **1c** or **2c** was converted to its dimethoxytrityl derivative according to the previously described procedure.[18] Following tritylation the nucleoside (0.23 mmol) was dissolved in pyridine (3.0 mL) and treated with acetic anhydride (0.428 mL, 4.65 mmol). After 16 h the solvent was removed *in vacuo* and residual pyridine removed by coevaporation with ethanol (2 × 30 mL). The residue was taken up in CH_2Cl_2 (10 mL) and washed with saturated aqueous bicarbonate solution (2 × 10 mL) and brine (2 × 5 mL). The organic layer was separated, dried ($MgSO_4$) and the solvent removed *in vacuo*. The residue was dissolved in 80% aqueous acetic acid (2 mL). After a further 16 h the solvent was removed *in vacuo* and residual acetic acid removed by coeveporation with ethanol (3 × 5 mL). Column chromatography over silica gel eluting with CH_2Cl_2–methanol (0–7%) afforded the required product.

2′-Deoxy-2′-α-*C*-(acetoxymethyl)-3′*O*-acetyluridine (1d). White hygroscopic foam (79% yield); Found FAB HRMS *m/z* $(M + H)^+$, 343.1138. $C_{14}H_{19}N_2O_8$ requires $(M + H)^+$, 343.1141; ^{1}H NMR (400 MHz; $CDCl_3$) δ 9.42 (1H, br s, NH), 7.83 (1H, d, J = 8.1 Hz, H6), 6.16 (1H, d, J = 8.8 Hz, H1′), 5.80 (1H, d, J = 8.1 Hz, H5), 5.43 (1H, dd, J = 1.5, 6.0 Hz, H3′), 4.36 (1H, dd, J = 6.6, 11.5 Hz, H6′), 4.21 (1H, dd, J = 7.4, 11.5 Hz, H6″), 4.12–4.10 (1H, m, H4′), 3.93 (2H, d, 2 × H5′), 2.13 (3H, s, CH_3), 1.98 (3H, s, CH_3), 2.97–2.93 (1H, m, H2′); ^{13}C NMR (75.4 MHz; $CDCl_3$) δ 170.64, 170.49 (2 × CO), 163.59 (C4), 150.69 (C2), 140.80 (C6), 102.94 (C5), 87.72 (C1′), 85.26 (C4′), 74.75 (C3′), 62.33 (C6′), 59.41 (C5′), 46.00 (C2′), 20.57, 20.41 (2 × CH_3); *m/z* (FAB^+) 343 ($[M + H]^+$, 26%), 231 ($[M -$ uracil$]^+$, 15), 81 (100).

2′-Deoxy-2′-α-*C*-(2-acetoxyethyl)-3′*O*-acetyluridine (2d). White amorphous solid (87% yield); found: C, 50.53; H, 5.70; N, 7.89. $C_{15}H_{20}O_8N_2$ requires C, 50.56; H, 5.66; N, 7.86; ^{1}H NMR (400 MHz; $CDCl_3$) δ 9.64 1H, br s, NH), 7.98 (1H, d, J = 8.1 Hz, H6), 6.09 (1H, d, J = 9.4 Hz, H1′), 5.81 (1H, d, J = 8.0 Hz, H5′), 5.35 (1H, d, J = 5.5 Hz, H3′), 4.07–3.93 (5H, m, 2 × H5′, 2 × H7′, H4′), 2.63–2.58 (1H, m, H2′), 2.15 (3H, s, CH_3), 2.04–2.00 (4H, m, CH_3 and H6′), 1.79–1.69 (1H, m, H6″); ^{13}C NMR (75.4 MHz; $CDCl_3$) δ 171.06, 170.64 (2 × CO), 163.42 (C4), 150.92 (C2), 140.64 (C6), 103.19 (C5), 88.79 (C1′), 85.41 (C4′), 75.57 (C3′), 62.67 (C5′), 62.18 (C7′), 44.37 (C2′), 23.07 (C6′), 20.65, 20.60 (2 × CH_3); *m/z* (FAB^+) 379 ($[M + Na]^+$, 88%), 357 ($[M + H]^+$, 100).

Synthesis of 2′-modified nucleoside 5′-triphosphates. Triphosphates were synthesised by phosphitylation with 2-chloro-5,6-benzo-1,2,3-dioxaphosphorin-4-one using a 100 μmoles of nucleoside **1d**, **2d**, or **3b**.[19] Prior to the transcription studies, the triphosphates were treated with sodium periodate as previously described [21] and repurified.

2′-Deoxy-2′-α-*C*-hydroxymethyluridine-5′-*O*-triphosphate (1a). Clear amorphous solid, yield 47%; ^{1}H NMR (400 MHz; $CDCl_3$) δ 7.82 (1H, d, J = 8.1 Hz, H6), 6.04 (1H, d, J = 8.6 Hz, H1′), 5.84 (1H, d, J = 8.1 Hz, H5), 4.54 (1H, d, J = 5.8 Hz, H3′), 4.11–4.03 (3H, m, 2 × H5′ + H4′), 3.81 (1H, dd, J = 6.7, 11.5 Hz, H6′), 3.60 (1H, dd, J = 6.7, 11.5 Hz, H6″), 2.55–2.52 (1H, m, H2′); ^{31}P NMR (101.3 MHz; D_2O) δ −6.38 (d, J = 20.3 Hz), −11.52 (d, J = 19.3), −22.90 (dd, J = 19.3, 20.3 Hz); *m/z* (ES^-) 497 ($[M - H]^-$, 6%), 417 ($[M - H - HPO_3]^-$, 100), 337 ($[M - H - H_2P_2O_6]^-$, 34), 248 ($[M - H]^-$, 72%); retention time (ion pair buffers) 16.16 min; Digest with alkaline phosphatase gave a single product, retention time 10.01 min, coelutes with compound **1**.

2′-Deoxy-2′-α-*C*-(2-hydroxyethyl) uridine-5′-*O*-triphosphate (2a). Clear amorphous solid, yield 56%; ^{1}H NMR (400 MHz; D_2O) δ 7.81 (1H, d, J = 8.2 Hz, H6), 5.96 (1H, d, J = 9.5, H1′), 5.86 (1H, d, J = 8.2 Hz, H5), 4.43 (1H, d, J = 5.4, H3′), 4.13 (1H, s, H4′), 4.06–4.01 (2H, m, 2 × H5′), 3.49 (2H, t, J = 6.5 Hz, H7′), 3.08–3.03 (24H, m, NCH_2), 2.39–2.36 (1H, m, H2′), 1.77–1.74 (1H, m, H6′), 1.20–1.17 (36H, m, CH_3), 1.50–1.54 (1H, m, H6″); ^{31}P NMR (101.3 MHz; D_2O) δ −6.25 (1P, d, J = 21.9 Hz), −11.18 (1P, d, J = 19.5 Hz), −22.25 (dd, J = 19.5, J' = 21.9 Hz); *m*/*z* (FAB$^-$) 511 ([*M*-H]$^-$, 52%), 239 ([$H_2P_3O_9$]$^-$, 100); HPLC retention time (ion pair buffers) 15.98 min; Digestion with alkaline phosphatase gave a single product, retention time 9.35 min, coeluted with a sample of compound **2**.

2′-Deoxy-2′-α-*C*-acetamidouridine-5′-*O*-triphosphate (3a). Clear amorphous solid; yield 54%; ^{1}H NMR (400 MHz; D_2O) δ 7.74 (1H, d, J = 8.2 Hz, H6), 5.97 (1H, d, J = 9.2 Hz, H1′), 5.83 (1H, d, J = 8.1 Hz, H5), 4.42 (1H, d, J = 5.4 Hz, H3′), 4.11 (1H, s, H4′), 4.07–4.01 (2H, m, H5′), 3.31–3.29 (24H, m, NCH_2), 2.71–2.67 (1H, m, H2′), 2.55 (1H, dd, J = 5.8, 15.1 Hz, H6′), 2.27 (1H, dd, J = 8.8, 11.1 Hz, H6″), 1.16–1.12 (36H, m, CH_3); ^{31}P NMR (101.3 MHz; D_2O) δ −6.22 (1P, d, J = 21.9 Hz), −11.15 (1P, d, J = 19.5 Hz), −22.17 (dd, J = 19.5, 21.9 Hz); *m*/*z* (FAB$^-$) 511 (63%, [*M* − H]$^-$), 507 (100%); HPLC retention time (ion pair buffers) 15.57 min; digestion with alkaline phosphatase gave a single product, retention time 9.03 min, coeluted with a sample of compound **3**.

General procedure for transcription reactions

Analytical scale transcription reactions (total volume 20μL) were performed in 2 mL microfuge tubes with reagents at the following final concentrations; Tris–HCl buffer (60 mM, pH 8.0), spermidine (1 mM), $MgCl_2$ (12 mM) or $MnCl_2$ (2.5 mM), DTT (5 mM), rATP (1 mM), rGTP (1 mM), rCTP (1.0 mM) containing [α-^{32}P] CTP (0.5 μCi), T7 RNA polymerase (200 units), DNA template (1.0 μM), UTP (1 mM for positive control reactions), modified UTP (1–8 mM). Negative control reactions were performed with no UTP present. The transcription reactions were incubated at 37 °C for 3 h, treated with gel loading buffer (10 μL) and analysed by PAGE on either a 12 or 20% denaturing gel, depending on the length of the transcript. The labelled transcripts were visualised by autoradiography.

Procedure for 5′-end labelling of the unmodified U15 transcript

A solution of tris–HCl buffer (120 mM, pH 8.0), spermidine (1 mM), $MnCl_2$ (2.5 mM), DTT (5 mM), rUTP (0.9 mM), rUαSTP (0.1 mM), rATP (1 mM), rGTP (1 mM), rCTP (1 mM), T7 RNA polymerase (900 units), DNA template (1 μM), was made up to a total volume of 100 μL and incubated at 37 °C for 3 h. The mixture was centrifuged at 14000 rpm for 5 min and the supernatant was concentrated using a Microcon 3 filter centrifuged at 10000 rpm for 20 min. Material was recovered from the filter by inversion and centrifugation into a 1 mL microfuge tube at 1200 rpm for 5 min. The filter was washed with 15 μL water and the washing added to the previous material, treated with gel loading buffer (15 μL) and loaded onto a 20% denaturing gel alongside a sample of internally labelled transcript. After elution of the gel and imaging film, the unlabelled transcript was cut, using the labelled transcript as a length marker. The gel piece was crushed and treated with 1 M NaOAc (200 μL), frozen in liquid nitrogen and allowed to warm to 0 °C on ice. The mixture was centrifuged and the aqueous layer separated. The gel was treated with further NaOAc (2 × 150 μL), vortexed, cooled to −196 °C and allowed to warm to 0 °C. The combined aqueous extracts were treated with glycogen (1 μL) and absolute ethanol (1 mL). The mixture was precipitated at −80 °C for 2 h and then centrifuged at 14000 rpm for 5 min. The mother liquor was removed and the remaining pellet was washed with 80% aqueous ethanol (200 μL) and then dried under reduced pressure. The sample was redissolved in water (15 μL) and treated with CIAP (5 μL), RNase (0.5 μL) inhibitor and dephosphorylation buffer (2.3 μL). The mixture was incubated at 37 °C for 30 min and then water (77 μL) was added followed by extraction with aqueous phenol (100 μL) and chloroform (2 × 100 μL). The organic extracts were discarded and the aqueous layer treated with 3 M NaOAc (10 μL), glycogen (1 μL) and ethanol (250 μL) and precipitated at −20 °C overnight. The mixture was centrifuged at 14000 rpm for 5 min. The mother liquor was removed and the remaining pellet was washed with 80% aqueous ethanol (200 μL) and then dried under reduced pressure. The pellet was redissolved in a solution of [γ-^{32}P]ATP (9.5 μL, approx. 20 μCi) and treated with PNK buffer (1.5 μL), RNase inhibitor (0.5 μL) and T4 PNK (3.5 μL). The mixtures were incubated at 37 °C for 30 min and then treated with 1 mM ATP (10 μL). After a further 5 min at 37 °C the mixtures were treated with 10 μL loading buffer and subjected to 20% polyacrylamide gel electrophoresis. After elution the gel was imaged, the required band cut and the labelled oligonucleotide extracted from the gel and precipitated by the procedure outlined above.

Procedure for 5′-end labelling of modified U15 transcript

Four separate solutions of tris–HCl buffer (120 mM, pH 8.0), spermidine (1.0 mM), $MnCl_2$ (2.5 mM), DTT (5 mM), rU^{HM}TP (1.0 mM), rATP (1.0 mM), rGTP (1.0 mM), rCTP (1.0 mM), T7 RNA polymerase (1800 units), DNA template (1.0 μM), were each made up to a total volume of 200 μL and incubated at 37 °C for 3 h. The mixtures were centrifuged at 14000 rpm for 5 min and the supernatant from each reaction was combined and concentrated using a Microcon 3 filter. From this point the end labelling procedure described above was followed.

Procedure for partial alkaline hydrolysis of 5′-end labelled transcripts

5′-End-labelled transcripts were treated with 50 mM aqueous $NaHCO_3$ (12 μL) for 6 min at 95 °C. The samples were cooled on ice and treated with gel loading buffer (10 μL) and the products separated by PAGE on 20% denaturing gel and visualised by autoradiography.

Acknowledgements

We would like to thank the BBSRC for support (grant BO8483) and the EPSRC for studentships to AJL and JBJP. JBJP was also the recipient of a short term EMBO Fellowship. We are also grateful to Allan Mills (Liverpool) and the EPSRC service at Swansea for obtaining mass spectra and Dr Paul Leonard (Liverpool) for assistance with NMR spectroscopy. We would like to thank Professor F. Eckstein (Göttingen, Germany) in whose laboratory the transcription studies were performed and Volker Fehring for assistance with transcription studies.

References

1 For reviews see: (*a*) J. A. Bittker, K. J. Phillips and D. R. Liu, *Curr. Opin. Chem. Biol.*, 2002, **6**, 367; (*b*) D. S. Wilson and J. W. Szostak, *Ann. Rev. Biochem.*, 1999, **68**, 611; (*c*) R. R. Breaker, *Chem. Rev.*, 1997, **97**, 371.
2 P. J. Unrau and D. P. Bartel, *Nature*, 1998, **395**, 260.
3 W. K. Johnston, P. J. Unrau, M. S. Lawrence, M. E. Glasner and D. P. Bartel, *Science*, 2001, **292**, 1319.
4 (*a*) H. A. Held and S. A. Benner, *Nucleic Acids Res.*, 2002, **30**, 3857; (*b*) H. Sawai, A. N. Ozaki, F. Satoh, T. Ohbayashi, M. M. Masud and H. Ozaki, *Chem. Commun.*, 2001, 2604; (*c*) T. Gourlain, A. Sidorov, N. Mignet, S. J. Thorpe, S. E. Lee, J. A. Grasby and D. M. Williams, *Nucleic Acids Res.*, 2001, **29**, 1898; (*d*) O. Thum, S. Jager and M. Famulok, *Angew. Chem., Int. Ed.*, 2001, **40**, 3990; (*e*) K. Sakthivel and C. F. Barbas, *Angew. Chem., Int. Ed.*, 1998, **37**, 2872.

5 N. K. Vaish, A. W. Fraley, J. W. Szostak and L. W. McLaughlin, *Nucleic Acids Res.*, 2000, **28**, 3316.
6 L. Lermer, Y. Roupioz, R. Ting and D. M. Perrin, *J. Am. Chem. Soc.*, 2002, **124**, 9960; (*a*) D. M. Perrin, T. Garestier and C. Helene, *J. Am. Chem. Soc.*, 2001, **123**, 1556; (*b*) S. W. Santoro, G. F. Joyce, K. Sakthivel, S. Gramatikova and C. F. Barbas, *J. Am. Chem. Soc.*, 2000, **122**, 2433; (*c*) T. M. Tarasow, S. L. Tarasow and B. E. Eaton, *Nature*, 1997, **389**, 54.
7 (*a*) S. P. Zinnen, K Domenico, M. Wilson, B. A. Dickinson, A. Beaudry, V. Mokler, A. T. Daniher, A. Burgin and L. Beigelman, *RNA*, 2002, **8**, 214; (*b*) M. Sioud and D. R. Sørensen, *Nature Biotech.*, 1998, **16**, 556.
8 (*a*) A. De Mesmaeker, R. Haner, P. Martin and H. E. Moser, *Acc. Chem. Res.*, 1995, **28**, 366; (*b*) B. E. Eaton and W. A. Pieken, *Ann. Rev. Biochem.*, 1995, **64**, 837.
9 S. P. Ryder, L. Ortoleva-Donnolly, A. B. Kosek and S. A. Strobel, *Methods Enzymol.*, 2000, **317**, 92.
10 C. Raines and P. A. Gottleib, *RNA*, 1998, **4**, 340.
11 R. Sousa and R. Padilla, *EMBO J.*, 1995, **14**, 4609.
12 H. Aurup, D. M. Williams and F. Eckstein, *Biochemistry*, 1992, **31**, 9636.
13 R. Padilla and R. Sousa, *Nucleic Acids Res.*, 1999, **27**, 1561.
14 R. Padilla and R. Sousa, *Nucleic Acids Res.*, 2002, **30**, e138.
15 F. Conrad, A. Hanne, R. K. Gaur and G. Krupp, *Nucleic Acids Res.*, 1995, **23**, 1845.
16 A. J. Lawrence, J. B. J. Pavey, I. A. O'Neil and R. Cosstick, *Tetrahedron Lett.*, 1995, **36**, 6341.
17 J. B. J. Pavey, A. J. Lawrence, A. J. Potter, R. Cosstick and I. A. O'Neil, *Tetrahedron Lett.*, 1998, **39**, 6967.
18 A. J. Lawrence, J. B. J. Pavey, I. A. O'Neil and R. Cosstick, *J. Org. Chem.*, 1996, **61**, 9213.
19 J. Ludwig and F. Eckstein, *J. Org. Chem.*, 1989, **54**, 631.
20 A. J. Lawrence, J. B. J. Pavey, M.-Y. Chan, R. Fairhurst, S. Collingwood, J. Fisher, Cosstick and I. A. O'Neil, *J. Chem. Soc., Perkin Trans. 1*, 1997, 2761.
21 J. P. Richard and P. A. Frey, *J. Am. Chem. Soc.*, 1982, **104**, 3476.
22 D. M. Brown, M. Fried and A. R. Todd, *Chem. Ind. (London)*, 1953, 352.
23 J. F. Milligan, D. R. Groebe, G. W. Witherell and O. Uhlenbeck, *Nucleic Acids Res.*, 1987, **15**, 8783.
24 Y. Huang, A. Beaudry, J. McSwiggen and R. Sousa, *Biochemistry*, 1997, **36**, 13718.
25 S. Tabor and C. Richardson, *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 4076.
26 J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2 ed, Cold Spring Harbor Laboratory Press, NY, 1989.
27 G. Gish and F. Eckstein, *Science*, 1988, **240**, 1520.
28 Y. Huang, F. Eckstein, R. Padilla and R. Sousa, *Biochemistry*, 1997, **36**, 8231.
29 C. M. Joyce, *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 1619.
30 W. Saenger, in *Principles of Nucleic acid Structure*, ed. W. Saenger, Springer, New York, 1984.
31 C. Altona and M. Sundaralingam, *J. Am. Chem. Soc.*, 1973, **93**, 2333.