Synthesis of nucleoside and nucleotide conjugates of bile acids, and polymerase construction of bile acid-functionalized DNA†

Satu Ikonen,[a,b] Hana Macíčková-Cahová,[a] Radek Pohl,[a] Miloslav Šanda[a] and Michal Hocek*[a]

Received 18th November 2009, Accepted 4th December 2009
First published as an Advance Article on the web 15th January 2010
DOI: 10.1039/b924072a

Aqueous Sonogashira cross-coupling reactions of 5-iodopyrimidine or 7-iodo-7-deazaadenine nucleosides with bile acid-derived terminal acetylenes linked *via* an ester or amide tether gave the corresponding bile acid–nucleoside conjugates. Analogous reactions of halogenated nucleoside triphosphates gave directly bile acid-modified dNTPs. Enzymatic incorporation of these modified nucleotides to DNA was successfully performed using Phusion polymerase for primer extension. One of the dNTPs (dCTP bearing cholic acid) was also efficient for PCR amplification.

Introduction

Nucleobase-functionalized nucleic acids are of great current interest due to applications in chemical biology, bioanalysis, catalysis or nanotechnology and material science.[1] Apart from classical chemical synthesis on a solid support, they can be prepared enzymatically[2] by polymerase incorporations of functionalized nucleoside triphosphates (dNTPs).[3] In addition, both approaches can be further combined with post-synthetic modifications by amide formation,[4] click chemistry[5] or the Staudinger ligation[6] to even increase the portfolio of accessible modifications. Particularly efficient is the single-step synthesis of modified dNTPs by aqueous cross-coupling reactions[7] followed by polymerase incorporation. This approach[8] we have recently used for the construction of DNA bearing amino acids,[9] ferrocene,[10] amino- and nitrophenyl,[11] Ru/Os(bpy)$_3$,[12] and tetrathiafulvalene[13] tags useful for electrochemical detection and further bioanalytical applications. 5-Substituted pyrimidine dNTPs are usually good substrates for DNA polymerases, while 8-substituted purine dNTPs were repeatedly shown[3c,9,14] to be poor substrates which should be replaced by 7-substituted 7-deazapurine dNTPs.

Bile acids are endogenous steroids which form as end products of cholesterol metabolism in the liver. They are natural substrates of enterohepatic circulation, and their main role is the digestion of lipids and lipid soluble vitamins.[15] Bile acids and their salts, as well as their derivatives, are found to have numerous applications in biology, medicine and physiology,[16] as well as in supramolecular chemistry and nanoscience.[17] Many aspects in bile acid structure and in their properties make them interesting as building blocks of functionalized DNA. The chiral, concave, steroidal scaffold of bile acids together with their tendency to self assemble might make it possible for the BA-functionalized ONs to form supramolecular constructions capable, for example, of molecular recognition similar to the entities derived from other steroids reported by Letsinger and Chaturvedi.[18] Bile acids, as natural substrates of enterohepatic circulation, are recognized and transported by specific transporter proteins. The bile acid moiety could thus be exploited in immunoassays or targeting of drug–ONs.[19]

Recently Gissot *et al.*[20] reviewed the nucleoside-, nucleotide- and oligonucleotide-based amphiphiles (ONA). Often these amphiphiles contain steroids as the hydrophobic part.[21] The appended hydrophobic segment of the ONA is hypothesized to provide an anchor for the antisense oligonucleotide into the membrane, thus facilitating internalization of the ONA. Owing to their unique facial amphiphilicity, bile acids are also expected to be potential building blocks for ONAs. There are just a few reports in the literature where bile acids have been conjugated with oligonucleotides. Furthermore, until now the existing reports mostly concern the use of the cholic acid moiety in the 5′-terminus of oligonucleotides where the functionalization is made by post-synthetic modification.[22] Surprisingly, the bile acid conjugates at the 5′-termini of ONs were found to significantly increase the thermodynamic stability of duplexes. Only two examples deal with chemical synthesis of base-functionalized ON where the (deoxy)cholic acid moiety was attached at the 5-position of the pyrimidine ring (either by post synthetic modification[23] or by solid-supported synthesis using deoxycholic-modified nucleoside phosphoramidite[24]). The steroid moiety was found to serve as a reaction enhancing cap in chemical (non-enzymatic) primer-extension[25] and to improve triplex formation.[24]

To the best of our knowledge, there are no reports about the conjugation of bile acids to nucleoside triphosphates, nor about the enzymatic incorporation of bile acids in DNA. Here, we wish to report the preparation of bile acid-functionalized deoxyribonucleosides (dN) and deoxyribonucleoside triphosphates (dNTP), and the scope of their polymerase incorporations to bile acid-functionalized DNA. Apart from the potential direct applications mentioned above (amphiphiles, supramolecular self-assembly, tissue targeting *etc.*), the bile acids serve in our study as model moieties for other biologically active steroid molecules in order to find out whether the DNA polymerases will tolerate such bulky and largely hydrophobic tags attached to nucleobases.

[a]*Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Gilead & IOCB Research Center, Flemingovo nam. 2, CZ-16610, Prague 6, Czech Republic. E-mail: hocek@uochb.cas.cz; Fax: +420 220183559; Tel: +420 220183324*
[b]*Department of Chemistry, University of Jyväskylä, P. O. Box 35, FI-40014, Jyväskylä, Finland*

† Electronic supplementary information (ESI) available: Full experimental part and spectral data, additional PAGEs, MALDI, sequencing analysis and copies of NMR spectra. See DOI: 10.1039/b924072a

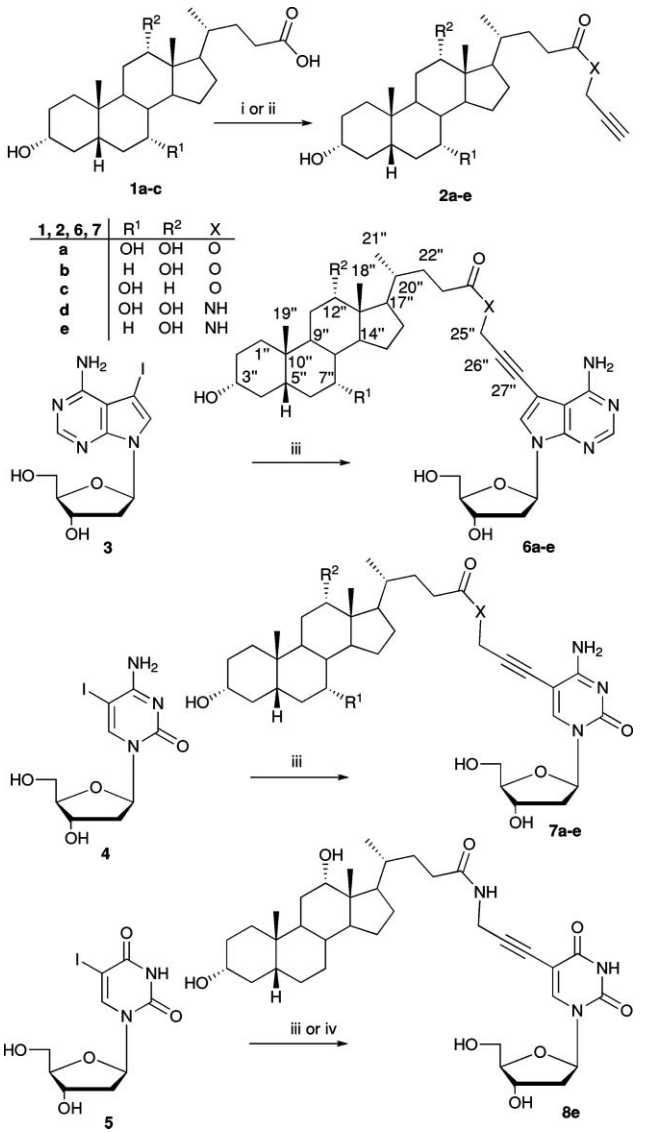
Results and discussion

Synthesis of modified nucleosides and dNTPs

Prior to the desired direct functionalization of nucleotides, the chemistry was first performed on model nucleosides in order to develop the methodology and select the most easily available and stable linkages of the use in dNTPs. The attachment of the bile acid moiety to position 5 of pyrimidine and to position 7 of 7-deazaadenine nucleosides was envisaged through the Sonogashira reaction with bile acid-derived terminal acetylenes. The synthesis of these key building blocks is shown in Scheme 1. Starting bile ester acetylenes (**2a–c**)[26] were prepared either by an acid-catalyzed esterification of bile acids with an excess of propargyl alcohol or by a base-catalyzed alkylation reaction between the bile acid (**1a–c**) and propargyl bromide. Starting amide acetylenes (**2d–e**)[27] were synthesized *via* two-step one-pot active anhydride method[28] widely used for the preparation of bile acid amides from the bile acid, without the need to protect the hydroxyl groups (**1a–b**).

The modified nucleosides were then prepared by single-step, aqueous-phase Sonogashira cross-coupling reactions of halogenated nucleosides (Scheme 1, Table 1) in analogy to the previously developed procedures.[29] The Pd-catalyzed, Cu-mediated cross coupling reactions of 7-iodo-7-deaza-2′-deoxyadenosine (**3**), or 5-iodo-2′-deoxycytidine (**4**) with bile acid amide acetylenes (**2d–e**) were performed in presence of $Pd(OAc)_2$ and $P(Ph–SO_3Na)_3$ (TPPTS) in water–acetonitrile (2 : 1) at 75 °C, and the reaction proceeded for 1 h giving **6d–e** and **7d–e** with 61–69% isolated yields (entries 1–4). Under the same conditions, the reaction of bile acid ester acetylenes (**2a–c**) with **3** or **4** resulted in hydrolysis of the ester bond. Therefore, the reaction temperature was decreased to 65 °C, time shortened to a max. of 20 min, solvent mixture changed to water–acetonitrile (1 : 1) for better solubility and the reaction mixture was preheated before the addition of catalyst. Under these optimized conditions, the cross-couplings proceeded reasonably well, without significant cleavage of the ester groups, to give the desired nucleoside conjugates **6a–c** and **7a–c** in 35–90% isolated yields (entries 5–10).

Reaction of 5-iodo-2′-deoxyuridine (**5**) with bile acid acetylenes under the above described conditions resulted in a complex mixture of products. The reaction with **2d** gave an unwanted furopyrimidine **9d** (see the ESI†).[5f,30] By changing the solvent to water–acetonitrile (1 : 1) and carrying out the reaction at 65 °C for only 20 min, the reaction of **5** with **2e** gave **8e** in 58% isolated



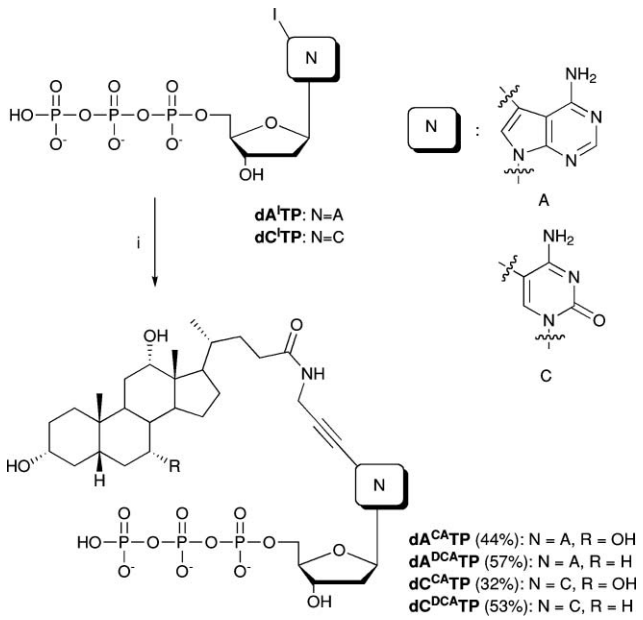
Scheme 1 Synthesis of BA-modified nucleosides. Reagents and conditions: (i) for **1a** and **1b**: (a) ethyl chloroformate, TEA, dioxane, 10–15 °C, 30 min, (b) propargyl amine, rt, 18 h; (ii) for **1d**: *p*-TSA, propargyl alcohol, 55–60 °C, 18 h, or for **1c**, and **1e**: Cs_2CO_3, DMF, propargyl bromide, rt, 18 h; (iii) **2**, CuI (10 mol%), $Pd(OAc)_2$ (5 mol%), TPPTS (5 eq. to Pd), $EtN(i\text{-}Pr)_2$ (10 eq.), $H_2O–CH_3CN$, 65–75 °C, 12–60 min; (iv) **2e**, CuI (10 mol%), $Pd(OAc)_2$ (5 mol%), TPPTS (5 eq. to Pd), $EtN(i\text{-}Pr)_2$ (10 eq.), dry DMF, 80 °C, 3 h.

Table 1 Synthesis of bile acid–nucleoside conjugates

Entry	2	dN	Solvent	Temp., time	Product	Yield (%)
1	**2d**	**3**	AN–H_2O 1 : 2	75 °C, 1 h	**6d**	69
2	**2e**	**3**	AN–H_2O 1 : 2	75 °C, 1 h	**6e**	65
3	**2d**	**4**	AN–H_2O 1 : 2	75 °C, 1 h	**7d**	66
4	**2e**	**4**	AN–H_2O 1 : 2	75 °C, 2 h	**7e**	61
5	**2a**	**3**	AN–H_2O 1 : 1	65 °C, 12 min	**6a**	69
6	**2b**	**3**	AN–H_2O 1 : 1	65 °C, 20 min	**6b**	35
7	**2c**	**3**	AN–H_2O 1 : 1	65 °C, 15 min	**6c**	45
8	**2a**	**4**	AN–H_2O 1 : 1	65 °C, 12 min	**7a**	39
9	**2b**	**4**	AN–H_2O 1 : 1	65 °C, 20 min	**7b**	88
10	**2c**	**4**	AN–H_2O 1 : 1	65 °C, 20 min	**7c**	90
11	**2e**	**5**	AN–H_2O 1 : 1	65 °C, 20 min	**8e**	58
12	**2e**	**5**	dry DMF	80 °C, 3 h	**8e**	31

yield. The same reaction was also performed under non-aqueous conditions in dry DMF, but **8e** was isolated in only 31% yield.

Having developed the methodology for the modification of nucleosides, we have proceeded with the direct functionalization of dNTPs by analogous Sonogashira cross-couplings. To avoid problems with cleavage of the esters, only selected amide acetylenes **2d–e** were used for attachment to dNTPs (Scheme 2). Thus, the bile acid acetylene amides **2d–e** were reacted with 7-iodo-7-deaza-2′-deoxyadenosine triphosphate (**dA^{I}TP**) and 5-iodo-2′-deoxycytidine triphosphate (**dC^{I}TP**). Reactions were carried out under analogous aqueous phase conditions as in the case of nucleoside conjugates. All of the reactions proceeded smoothly within 45 min to afford the corresponding conjugates **dA^{CA}TP**, **dA^{DCA}TP**, **dC^{CA}TP** and **dC^{DCA}TP**, which were isolated *via* reverse phase (RP) HPLC in good yields of 31–57% (similarly to the previous cases,[9–12] the isolated yields were lowered by partial hydrolysis of dNTPs to di- and monophosphates that are removed by HPLC).



Scheme 2 Synthesis of BA-modified dNTPs. Reagents and conditions: (i) **2**, CuI (10 mol%), $Pd(OAc)_2$ (5 mol%), TPPTS (5 eq. to Pd), $EtN(i\text{-}Pr)_2$ (10 eq.), H_2O–CH_3CN, 75 °C, 45 min.

Incorporation of modified dNTPs by DNA polymerases

The ability of DNA polymerases to incorporate the prepared modified dNTPs was studied by a primer extension (PEX)-reaction using four types of thermostable DNA polymerases—*Thermus brockianus* (DyNAzymeII), *Thermococcus litoralis* (Vent (*exo*-)), *Pyrococcus woesei* (Pwo), and Phusion® High-Fidelity DNA Polymerase, an artificial construct of a dsDNA-binding domain fused to a *Pyrococcus*-like enzyme (Phusion). Different ON templates (Table 2) were tested, including those accommodating the conjugates at separate positions, at two or three adjacent positions, or at positions alternating with another base (doublet repeat). The formation of BA-functionalized ONs in the presence of primer, modified dATP (**dA^{CA}TP** and **dA^{DCA}TP**) or dCTP (**dC^{CA}TP** and **dC^{DCA}TP**) and three additional natural dNTPs was studied by denaturing polyacrylamide gel electrophoresis (PAGE)-analysis. Each PEX experiment with a modified nucleotide was compared with a positive control experiment using all natural dNTPs, and with a negative control experiment using three natural dNTPs and absence of the particular dNTP under study. The lack of primer extension in the negative control clearly excludes any possible non-specific incorporation of natural dNTPs in place of the modified nucleotide. Such combination of experiments is self-sufficient proof for identity and sequence of the modified DNA.[3]

Table 2 Primer and templates used for PEX experiments

Primer	5′-CATGGGCGGCATGGG-3′
Temp^{md16}	5′-**CTAGCATGAGCTCAGT**CCCATGCCGCCCATG-3′
Temp^{1C}	5′-**CCCG**CCCATGCCGCCCATG-3′
Temp^{1A}	5′-**CCCT**CCCATGCCGCCCATG-3′
Temp^{AA}	5′-**GCGACGAAGAGCTT**CCCATGCCGCCCATG-3′
Temp^{3A}	5′-**TTATATTTATA**CCCATGCCGCCCATG-3′
Temp^{repA}	5′-**TATATATATAT**CCCATGCCGCCCATG-3′

Template temp^{md16} used in experiments involving magnetoseparation procedure was biotinylated at the 5′-end.

Fig. 1 shows analysis of the PEX-reaction products containing one modification synthesized on temp^{1A} and temp^{1C} templates using Phusion polymerase. In each case the primer is fully extended, proving the incorporation. On the other hand, the lack of extension in negative control experiments (–A or –C) proves that no non-specific incorporation of natural dNTPs occurred. The bands of bile acid-modified ONs are migrating more slowly compared to the natural ones, which could be explained by the increase in weight and size by attachment of the bulky bile acid moiety. Each spot of PEX product in PAGE-analysis is accompanied by another extra band of a product that is one nucleotide shorter. Since this additional spot is present even in the negative control experiments (representing here the $n - 1$ truncated primer), it is evident that it comes from partial 3′–5′-exonuclease activity of the enzyme rather than from incomplete PEX. In addition to PAGE analysis, we have also unequivocally proved the successful incorporation of **dC^{CA}TP** and **dC^{DCA}TP** by

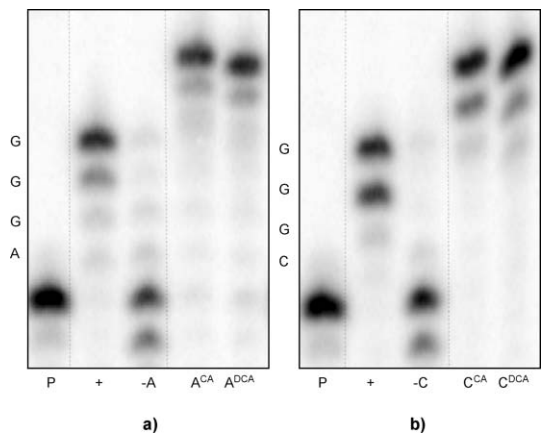


Fig. 1 Denaturing PAGE analysis of PEX products synthesized on temp^{1A} (a) and temp^{1C} (b) with Phusion DNA polymerase. 5-^{32}P-end-labeled primer-template was incubated with different combinations of natural and functionalized dNTPs. P: Primer; +: natural dNTPs; –A: dTTP, dCTP, dGTP; –C: dATP, dTTP, dGTP; A^{CA}: **dA^{CA}TP**, dTTP, dCTP, dGTP; A^{DCA}: **dA^{DCA}TP**, dTTP, dCTP, dGTP; C^{CA}: **dC^{CA}TP**, dATP, dTTP, dGTP; C^{DCA}: **dC^{DCA}TP**, dATP, dTTP, dGTP.

measurement of MALDI mass spectra of the PEX products (see ESI, Fig. S1†).

All polymerases were tested using template temp^{md16}. All conjugates were successfully incorporated through primer extension catalyzed by Vent (*exo*-), Phusion or DyNAzyme polymerases. Surprisingly, Pwo (which was successful[9,11,12] in the incorporation of most of our previous modified nucleotides) has not incorporated these bile acid dNTPs efficiently and only non-specific products were observed. Phusion DNA polymerase was the most efficient in the incorporation of all modified nucleotides to most sequences (Fig. 2a and more gels in the ESI†). The modified ONs constructed by PEX using temp^{md16} contain four modifications in one molecule resulting in a higher molecular weight and thus slower electrophoretic mobilities, which is clearly visible on the gel. In addition, the polymerase in some cases (Fig. 2a) seems to add an additional nucleotide in an untemplated fashion at the end of PEX, resulting in an extra band in PAGE analysis. Similar behavior was previously observed by us[9] and others.[31] The incorporation of the modified dATPs to sequences with two or three adjacent modifications as well as to a sequence with doublet repetition was also tested. The incorporation of both **dA^{CA}TP** and **dA^{DCA}TP** to two adjacent positions was effective (see Fig. 2); however, the incorporation of three adjacent modifications as well as the modifications to repetitive sequences was less feasible, resulting in longer, non-specific reaction products (see ESI†).

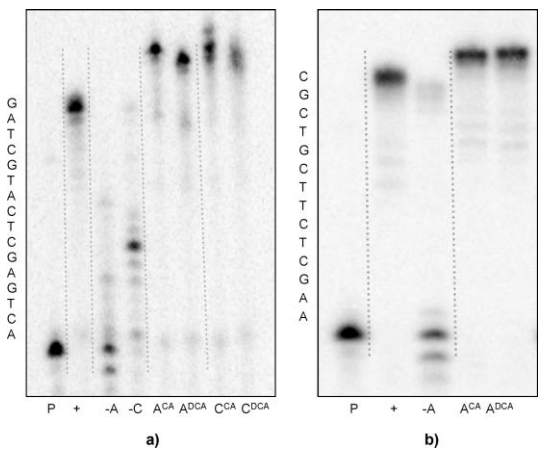


Fig. 2 Denaturating PAGE analysis of PEX products synthesized on a temp^{md16} (a) and temp^{AA} (b) with Phusion polymerase. 5-^{32}P-end-labeled primer-template was incubated with different combinations of natural and functionalized dNTPs. P: Primer; +: natural dNTPs; –A: dTTP, dCTP, dGTP; A^{CA}: **dA^{CA}TP**, dTTP, dCTP, dGTP; A^{DCA}: **dA^{DCA}TP**, dTTP, dCTP, dGTP; –C: dTTP, dATP, dGTP; C^{CA}: **dC^{CA}TP**, dTTP, dATP, dGTP; C^{DCA}: **dC^{DCA}TP**, dTTP, dATP, dGTP

With the promising results of PEX experiments, we have further tested the suitability of the bile acid dNTP conjugates in PCR reactions. The PCR experiments were performed using a 98-mer template, and 20- and 25-mer primers in the presence of Phusion DNA polymerase under different conditions. The formation of BA-functionalized DNA after 30 PCR cycles in the presence of primers, modified dATP (**dA^{CA}TP** and **dA^{DCA}TP**) or dCTP (**dC^{CA}TP** and **dC^{DCA}TP**) and three additional natural dNTPs was analyzed by agarose gel electrophoresis (Fig. 3). The best conditions for incorporation of modified nucleotides were found

Table 3 Melting temperatures of DNA duplexes[a]

	Temp^{AA}		Temp^{md16}	
	T_m/°C	ΔT_m/°C[b]	T_m/°C	ΔT_m/°C[b]
dATP	72.6	0	71.6	0
dA^{CA}TP	73.6	+0.5	74.6	+0.5
dA^{DCA}TP	73.6	+0.5	73.6	+0.8

[a] Prepared by PEX with temp^{AA}, temp^{md16}, Phusion, dCTP, dGTP, dTTP, and dATP or dA^{CA}TP, dA^{DCA}TP. [b] $\Delta T_m = (T_{mmod} - T_{mnat})/n_{mod}$.

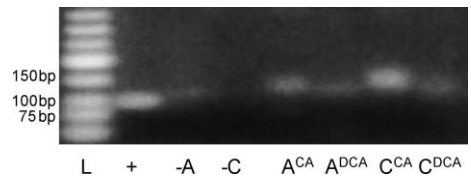


Fig. 3 Agarose gel electrophoresis analysis of PCR products synthesized on 98-mer template using Phusion polymerase. L: Ladder; +: natural dNTPs; -A: dTTP, dGTP, dCTP; -C: dATP, dTTP, dGTP; A^{CA}: **dA^{CA}TP**, dTTP, dCTP, dGTP; A^{DCA}: **dA^{DCA}TP**, dTTP, dCTP, dGTP; C^{CA}: **dC^{CA}TP**, dATP, dTTP, dGTP; C^{DCA}: **dC^{DCA}TP**, dATP, dTTP, dGTP.

with Phusion polymerase using DMSO 5%, formamide 0.25%, Betaine 37.5 mM and tetramethylammonium chloride (TMAC) 2.5 mM as additives, together with an excess of modified dNTPs and a prolonged reaction time (for exact conditions, see the Experimental part). The most efficient substrate was **dC^{CA}TP** which gave nice spot of fully extended amplified DNA, while the other nucleotides gave rather weak bands showing less efficient amplification. The bands of the modified DNA are migrating more slowly compared to the natural one, which can again be explained by the increase in the molecular weight and size.

The PCR products with natural dCTP and **dC^{CA}TP** were also partly characterized by sequencing analysis. The products of PCR with either natural or modified nucleotides were purified and amplified in subsequent PCR with natural dNTPs (see ESI, Figure S2†). The amplified natural DNA was then sequenced with a 20-mer primer. In this way, we were able to determine part of the sequence (46–52 nt) showing good fidelity of the PCR (ESI, Figure S3†).

Thermal denaturation studies

Thermal denaturation studies were used to assess the stability of DNA duplexes containing the modification in one strand. The ONs for these measurements were prepared by PEX on large scale using Phusion polymerase, templates temp^{AA}, and temp^{md16}, dCTP, dGTP, dTTP and dATP (positive control) or **dA^{TCA}TP** or **dA^{DCA}TP** as surrogates of natural dATP. A higher concentration of the template and primers were used to get pure products. The melting temperatures of the natural and modified DNAs are summarized in the Table 3. Based on the melting temperatures, the presence of the bile acid modification does not significantly affect the stability of the duplexes. There is only a very small stabilizing effect, ~0.5 °C/modification that may have been caused by the effect of acetylene group reported previously for 7-alkyne-modified 7-deazaadenine containing ONs.[32]

Conclusions

Conjugates of nucleosides and nucleoside triphosphates with bile acids have been successfully prepared by aqueous cross-coupling reactions of halogenated nucleosides or dNTPs with ester- or amide-linked bile acid terminal acetylenes. The amide tether is more useful since the ester linker is partially cleaved during the aqueous cross-coupling in presence of base. The bile acid-modified dNTPs were successfully incorporated in DNA by primer extension using several polymerases (the best one was Phusion, while Pwo was not efficient). One of the dNTPs (**dC^{CA}TP**) was also efficient for PCR amplification. The results show that the DNA polymerases can tolerate the presence of very bulky and hydrophobic steroid moiety in position 5 of pyrimidine or position 7 of 7-deazaadenine dNTPs. The bile acid-modified DNA could serve as tools in a number of applications in supramolecular chemistry and chemical biology (targeting liver tissues, immunoassays *etc.*). Moreover, the study has shown the feasibility of the presented methodology for possible attachment of other types of bioactive steroids (*e.g.* corticoids, neurosteroids, hormones *etc.*) to nucleic acids for a plethora of potential applications.

Experimental part

All cross-coupling reactions were performed in degassed solvents under argon atmosphere. Preparative HPLC separations were performed on a column packed with 10 μm C18 reversed phase (Phenomenex, Luna C18(2)). Typical experimental procedures and representative examples of characterization of compounds are given below. A complete detailed experimental part including full characterization data for all compounds is in the ESI.†

Synthesis of modified nucleosides—Sonogashira cross-coupling—general procedure

Water–acetonitrile mixture (1.5 mL) was added through a septum to an argon-purged vial containing a halogenated nucleoside **3–5**, an acetylene **2** (1.2–1.5 equiv. to nucleoside), and CuI (10 mol%). In a separate vial, $Pd(OAc)_2$ (5 mol%) and TPPTS (5 equiv. to Pd) were combined under argon with 0.5 mL of H_2O–CH_3CN. After dissolution of the solids, the catalyst solution was added to the reaction mixture followed by addition of $EtN(i\text{-}Pr)_2$ (10 equiv.). The reaction mixture was stirred under argon at 65–75 °C for 12–120 min and concentrated. Products were purified by column chromatography and dried *in vacuo* to give **6a**–**8e**.

7-{[(3α,7α,12α-trihydroxy-24-oxo-5β-cholan-24-yl)oxy]prop-1-yn-1-yl}-7-deaza-2′-deoxyadenosine (6a). Yellowish solid, yield 69%. ^{1}H NMR (500.0 MHz, CD_3OD): 0.47 (s, 3H, CH_3-18″); 0.77 (qd, 1H, J = 12.2, 5.9, H-15″α); 0.87 (s, 3H, CH_3-19″); 0.95 (td, 1H, J = 14.4, 3.5, H-1″β); 1.09 (d, 3H, J = 6.4, CH_3-21″); 1.20-1.67 and 1.73-1.95 (steroid envelope); 2.19 (td, 1H, J = 12.5, 4.6, H-9″α); 2.25 (td, 1H, J = 13.2,11.3, H-4″α); 2.32 (ddd, 1H, J_{gem} = 13.4, $J_{2'b,1'}$ = 5.9, $J_{2'b,3'}$ = 2.6, H-2′b); 2.36 (ddd, 1H, J = 14.9, 9.1, 7.3, H-23″b); 2.44 (ddd, 1H, J = 14.9, 7.6, 5.1, H-23″a); 2.61 (ddd, 1H, J_{gem} = 13.4, $J_{2'a,1'}$ = 8.2, $J_{2'a,3'}$ = 5.8, H-2′a); 3.35 (tt, 1H, J = 11.3, 4.5, H-3″β); 3.70 (q, 1H, J = 3.1, H-7″β); 3.72, 3.80 (2 × dd, 2 × 1H, J_{gem} = 12.1, $J_{5',4'}$ = 3.4, H-5′); 3.86 (t, 1H, J = 3.1, H-12″β); 4.01 (td, 1H, $J_{4',5'}$ = 3.4, $J_{4',3'}$ = 2.6, H-4′); 4.51 (dt, 1H, $J_{3',2'}$ = 5.8, 2.6 $J_{3',4'}$ = 2.6, H-3′); 4.90, 4.94 (2 × d, 2 × 1H, J = 15.8, H-25″); 6.47 (dd, 1H, $J_{1',2'}$ = 8.2, 5.9, H-1′); 7.66 (s, 1H, H-8); 8.09 (s, 1H, H-2). ^{13}C NMR (125.7 MHz, CD_3OD): 12.84 (CH_3-18″); 17.52 (CH_3-21″); 23.19 (CH_3-19″); 24.14 (CH_2-15″); 27.78 (CH-9″); 28.72 (CH_2-16″); 29.46 (CH_2-11″); 31.14 (CH_2-2″); 32.18 (CH_2-23″); 32.54 (CH_2-22″); 35.80 (CH_2-6″); 35.86 (C-10″); 36.38 (CH-20″); 36.44 (CH_2-1″); 40.42 (CH_2-4″); 40.90 (CH-8″); 41.76 (CH_2-2′); 42.85 (CH-14″); 43.15 (CH-5″); 47.42 (C-13″); 48.12 (CH-17″); 53.74 (CH_2-25″); 63.64 (CH_2-5′); 68.99 (CH-7″); 72.85 (CH-3″); 73.08 (CH-3′); 73.94 (CH-12″); 80.14 (C-27″); 86.91 (CH-1′); 88.09 (C-26″); 89.34 (CH-4′); 96.30 (C-7); 104.99 (C-5); 128.84 (CH-8); 149.98 (C-4); 153.32 (CH-2); 159.21 (C-6); 175.68 (C-24″). ν(KBr)/cm^{-1}: 3434, 2936, 2868, 2231, 1733, 1626, 1590, 1571, 1538, 1464, 1450, 1377, 1300, 1196, 1169, 1092, 1077, 1043, 980, 914, 797; MS (ES$^+$): found m/z: 695.4 ([M + H]$^+$), 717.4 ([M + Na]$^+$), 1411.0 ([2M + Na]$^+$); HRMS (ES$^-$): found m/z: 695.4018 ($C_{38}H_{55}O_8N$ requires 695.4014).

7-{[(3α,12α-dihydroxy-24-oxo-5β-cholan-24-yl)amino]prop-1-yn-1-yl}-7-deaza-2′-deoxyadenosine (6e). Yellowish solid, yield 65%. ^{1}H NMR (500.0 MHz, CD_3OD): 0.52 (s, 3H, CH_3-18″); 0.80 (qd, 1H, J = 11.9, 5.8, H-15″α); 0.89 (s, 3H, CH_3-19″); 0.96 (td, 1H, J = 14.2, 3.4, H-1″β); 1.01 (d, 3H, J = 6.5, CH_3-21″); 1.11 (qd, 1H, J = 13.8, 3.7, H-7″α); 1.20-1.62 and 1.71-1.90 (steroid envelope); 2.19 (dt, 1H, J = 13.7, 8.3, H-23″b); 2.29 (ddd, 1H, J = 13.7, 8.4, 5.1, H-23″a); 2.31 (ddd, 1H, J_{gem} = 13.4, $J_{2'b,1'}$ = 5.9, $J_{2'b,3'}$ = 2.6, H-2′b); 2.62 (ddd, 1H, J_{gem} = 13.4, $J_{2'a,1'}$ = 8.3, $J_{2'a,3'}$ = 5.9, H-2′a); 3.51 (tt, 1H, J = 11.2, 4.6, H-3″β); 3.72, 3.79 (2 × dd, 2 × 1H, J_{gem} = 12.2, $J_{5',4'}$ = 3.4, H-5′); 3.89 (t, 1H, J = 3.1, H-12″β); 4.01 (td, 1H, $J_{4',5'}$ = 3.4, $J_{4',3'}$ = 2.6, H-4′); 4.11, 4.16 (2×d, 2×1H, J = 17.6, H-25″); 4.51 (dt, 1H, $J_{3',2'}$ = 5.9, 2.6 $J_{3',4'}$ = 2.6, H-3′); 6.46 (dd, 1H, $J_{1',2'}$ = 8.3, 5.9, H-1′); 7.57 (s, 1H, H-8); 8.09 (s, 1H, H-2). ^{13}C NMR (125.7 MHz, CD_3OD): 13.11 (CH_3-18″); 17.69 (CH_3-21″); 23.72 (CH_3-19″); 24.78 (CH_2-15″); 27.40 (CH_2-7″); 28.38 (CH_2-6″); 28.72 (CH_2-16″); 29.80 (CH_2-11″); 30.62 (CH_2-25″); 31.03 (CH_2-2″); 33.34 (CH_2-22″); 33.99 (CH_2-23″); 34.73 (CH-9″); 35.27 (C-10″); 36.39 (CH_2-1″,CH-20″); 37.16 (CH_2-4″); 37.34 (CH-8″); 41.66 (CH_2-2′); 43.58 (CH-5″); 47.49 (C-13″); 48.14 (CH-17″); 49.29 (CH-14″); 63.66 (CH_2-5′); 72.51 (CH-3″); 73.09 (CH-3′); 74.00 (CH-12″); 75.88 (C-27″); 86.84 (CH-1′); 89.28 (CH-4′); 89.74 (C-26″); 97.07 (C-7); 104.97 (C-5); 127.98 (CH-8); 149.84 (C-4); 153.17 (CH-2); 159.19 (C-6); 177.08 (C-24″). ν(KBr)/cm^{-1}: 3430, 2934, 2866, 2230, 1626, 1591, 1571, 1533, 1461, 1450, 1377, 1301, 1198, 1174, 1090, 1076, 1043, 981, 914, 858, 797; MS (ES$^-$): found m/z: 676.3 ([M − H]$^-$); HRMS (ES$^-$): found m/z: 676.4067 ($C_{38}H_{54}N_5O_6$ requires 676.4074).

5-{[(3α,7α-dihydroxy-24-oxo-5β-cholan-24-yl)oxy]prop-1-yn-1-yl}-2′-deoxycytidine (7c). White solid, yield 90%. ^{1}H NMR (500.0 MHz, CD_3OD): 0.64 (s, 3H, CH_3-18″); 0.92 (s, 3H, CH_3-19″); 0.96 (d, 3H, J = 6.6, CH_3-21″); 0.98 (td, 1H, J = 14.1, 3.2, H-1″β); 1.04 (qd, 1H, J = 11.7, 6.3, H-15″α); 1.16 (q, 1H, J = 9.8, H-17″α); 1.18 (td, 1H, J = 12.0, 4.0, H-12″α); 1.24-1.74 and 1.81-2.01 (steroid envelope); 2.13 (dt, 1H, J_{gem} = 13.6, $J_{2'b,1'}$ = $J_{2'b,3'}$ = 6.3, H-2′b); 2.25 (td, 1H, J = 13.3, 11.5, H-4″α); 2.33 (dt, 1H, J = 15.3, 8.2, H-23″b); 2.39 (ddd, 1H, J_{gem} = 13.6, $J_{2'a,1'}$ = 6.3, $J_{2'a,3'}$ = 3.6, H-2′a); 2.43 (ddd, 1H, J = 15.3, 8.5, 5.2, H-23″a); 3.37 (tt, 1H, J = 11.2, 4.5, H-3″β); 3.74 (dd, 1H, J_{gem} = 12.1, $J_{5'b,4'}$ = 3.6, H-5′b); 3.78 (q, 1H, J = 7.8, H-7″β); 3.83 (dd, 1H, J_{gem} = 12.1, $J_{5'a,4'}$ = 3.1, H-5′a); 3.95 (td, 1H, $J_{4',5'}$ = 3.6, 3.1, $J_{4',3'}$ = 3.6, H-4′); 4.37 (dt, 1H, $J_{3',2'}$ = 6.2, 3.6 $J_{3',4'}$ = 3.6, H-3′); 4.89, 4.94

(2 × d, 2 × 1H, J = 15.9, H-25″); 6.20 (t, 1H, $J_{1',2'}$ = 6.3, H-1′); 8.40 (s, 1H, H-6). ^{13}C NMR (125.7 MHz, CD_3OD): 12.22 (CH_3-18″); 18.74 (CH_3-21″); 21.75 (CH_2-11″); 23.40 (CH_3-19″); 24.67 (CH_2-15″); 29.29 (CH_2-16″); 31.32 (CH_2-2″); 31.99 (CH_2-23″); 32.29 (CH_2-22″); 34.01 (CH-9″); 35.86 (CH_2-6″); 36.20 (C-10″); 36.52 (CH_2-1″); 36.55 (CH-20″); 40.44 (CH_2-4″); 40.70 (CH-8″); 40.98 (CH_2-12″); 42.47 (CH_2-2′); 43.13 (CH-5″); 43.68 (C-13″); 51.52 (CH-14″); 53.53 (CH_2-25″); 57.30 (CH-17″); 62.41 (CH_2-5′); 69.06 (CH-7″); 71.73 (CH-3′); 72.83 (CH-3″); 78.39 (C-27″); 88.03 (CH-1′); 89.12 (CH-4′); 91.20 (C-26″); 91.66 (C-5); 146.75 (CH-6); 156.68 (C-2); 166.48 (C-4); 175.36 (C-24″). ν(KBr)/cm^{-1}: 3425, 2936, 2868, 2232, 1732, 1649, 1602, 1505, 1447, 1416, 1377, 1306, 1259, 1236, 1190, 1153, 1092, 1078, 1045, 1000, 982, 951, 914, 858, 786; MS (ES$^+$): found m/z: 656.2 ([M + H]$^+$), 678.4 ([M + Na]$^+$), 694.3 ([M + K]$^+$), 1311.4 ([2M + H]$^+$), 1334.3 ([2M + Na]$^+$); HRMS (ES$^+$): found m/z: 656.3909 ($C_{36}H_{54}O_8N_3$ requires 656.3905).

5-[(3α,12α-dihydroxy-24-oxo-5β-cholan-24-yl)amino]prop-1-yn-1-yl}-2′-deoxycytidine (7e). Yellowish solid, 61%. ^{1}H NMR (500.0 MHz, CD_3OD): 0.66 (s, 3H, CH_3-18″); 0.92 (s, 3H, CH_3-19″); 0.98 (td, 1H, J = 14.1, 3.5, H-1″β); 1.02 (d, 3H, J = 6.4, CH_3-21″); 1.03 (m, 1H, H-15″α); 1.14 (qd, 1H, J = 13.4, 4.2, H-7″α); 1.22-1.63 and 1.74-1.92 (steroid envelope); 2.12 (dt, 1H, J_{gem} = 13.6, $J_{2'b,1'}$ = $J_{2'b,3'}$ = 6.2, H-2′b); 2.16 (ddd, 1H, J = 13.8, 8.8, 7.6, H-23″b); 2.28 (ddd, 1H, J = 13.8, 9.2, 5.2, H-23″a); 2.39 (ddd, 1H, J_{gem} = 13.6, $J_{2'a,1'}$ = 6.2, $J_{2'a,3'}$ = 3.9, H-2′a); 3.52 (tt, 1H, J = 11.0, 4.7, H-3″β); 3.73 (dd, 1H, J_{gem} = 12.0, $J_{5'b,4'}$ = 3.6, H-5′b); 3.81 (dd, 1H, J_{gem} = 12.0, $J_{5'a,4'}$ = 3.2, H-5′a); 3.94 (t, 1H, J = 3.0, H-12″β); 3.95 (ddd, 1H, $J_{4',3'}$ = 3.9, $J_{4',5'}$ = 3.6, 3.2, H-4′); 4.12, 4.15 (2 × d, 2 × 1H, J = 17.7, H-25″); 4.36 (dt, 1H, $J_{3',2'}$ = 6.2, 3.9, $J_{3',4'}$ = 3.9, H-3′); 6.20 (t, 1H, $J_{1',2'}$ = 6.2, H-1′); 8.31 (s, 1H, H-6). ^{13}C NMR (125.7 MHz, CD_3OD): 13.25 (CH_3-18″); 17.70 (CH_3-21″); 23.72 (CH_3-19″); 24.91 (CH_2-15″); 27.46 (CH_2-7″); 28.40 (CH_2-6″); 28.70 (CH_2-16″); 29.89 (CH_2-11″); 30.54 (CH_2-25″); 31.07 (CH_2-2″); 33.20 (CH_2-22″); 33.92 (CH_2-23″); 34.81 (CH-9″); 35.30 (C-10″); 36.42 (CH_2-1″); 36.58 (CH-20″); 37.20 (CH_2-4″); 37.43 (CH-8″); 42.44 (CH_2-2′); 43.62 (CH-5″); 47.57 (C-13″); 48.12 (CH-17″); 49.28 (CH-14″); 62.52 (CH_2-5′); 71.82 (CH-3′); 72.53 (CH-3″); 74.04 (CH-12″); 74.47 (C-27″); 87.98 (CH-1′); 89.11 (CH-4′); 92.36 (C-5); 92.86 (C-26″); 145.83 (CH-6); 156.71 (C-2); 166.56 (C-4); 176.85 (C-24″). ν(KBr)/cm^{-1}: 3419, 2829, 2866, 2359, 2341, 2233, 1732, 1649, 1598, 1506, 1448, 1416, 1376, 1308, 1260, 1163, 1091, 1078, 1050, 1000, 979, 955, 900, 860, 785; MS (ES$^+$): found m/z: 655.3 ([M + H]$^+$), 677.4 ([M + Na]$^+$), 1309.4 ([2M + H]$^+$), 1331.4 ([2M + Na]$^+$); HRMS (ES$^-$): found m/z: 653.3919 ($C_{36}H_{54}N_4O_7$ requires 653.3914).

5-[(3α,12α-dihydroxy-24-oxo-5β-cholan-24-yl)amino]prop-1-yn-1-yl}-2′-deoxyuridine (8e). White solid, yield 58% of **7a** as a white solid. ^{1}H NMR (500.0 MHz, CD_3OD): 0.69 (s, 3H, CH_3-18″); 0.93 (s, 3H, CH_3-19″); 0.98 (td, 1H, J = 14.2, 3.5, H-1″β); 1.02 (d, 3H, J = 6.4, CH_3-21″); 1.07 (qd, 1H, J = 11.8, 5.7, H-15″α); 1.17 (qd, 1H, J = 13.7, 4.6, H-7″α); 1.24-1.64 and 1.74-1.93 (steroid envelope); 2.13 (ddd, 1H, J = 14.1, 9.5, 7.2, H-23″b); 2.21 (ddd, 1H, J_{gem} = 13.6, $J_{2'b,1'}$ = 7.1, $J_{2'b,3'}$ = 6.1, H-2′b); 2.27 (ddd, 1H, J = 14.1, 9.7, 4.8, H-23″a); 2.31 (ddd, 1H, J_{gem} = 13.6, $J_{2'a,1'}$ = 6.3, $J_{2'a,3'}$ = 3.6, H-2′a); 3.52 (tt, 1H, J = 11.1, 4.6, H-3″β); 3.73 (dd, 1H, J_{gem} = 12.0, $J_{5'b,4'}$ = 3.5, H-5′b); 3.81 (dd, 1H, J_{gem} = 12.0, $J_{5'a,4'}$ = 3.1, H-5′a); 3.94 (ddd, 1H, $J_{4',5'}$ = 3.5, 3.1, $J_{4',3'}$ = 3.2, H-4′); 3.95 (t, 1H, J = 2.8, H-12″β); 4.13 (s, 2H, H-25″); 4.40 (ddd, 1H, $J_{3',2'}$ = 6.1, 3.6, $J_{3',4'}$ = 3.2, H-3′); 6.24 (dd, 1H, $J_{1',2'}$ = 7.1, 6.3, H-1′); 8.32 (s, 1H, H-6). ^{13}C NMR (125.7 MHz, CD_3OD): 13.25 (CH_3-18″); 17.69 (CH_3-21″); 23.71 (CH_3-19″); 24.92 (CH_2-15″); 27.47 (CH_2-7″); 28.40 (CH_2-6″); 28.67 (CH_2-16″); 29.87 (CH_2-11″); 30.36 (CH_2-25″); 31.05 (CH_2-2″); 33.15 (CH_2-22″); 33.91 (CH_2-23″); 34.79 (CH-9″); 35.30 (C-10″); 36.42 (CH_2-1″); 36.72 (CH-20″); 37.17 (CH_2-4″); 37.44 (CH-8″); 41.73 (CH_2-2′); 43.61 (CH-5″); 47.56 (C-13″); 48.07 (CH-17″); 49.26 (CH-14″); 62.62 (CH_2-5′); 72.12 (CH-3′); 72.53 (CH-3″); 74.05 (CH-12″); 75.03 (C-27″); 87.03 (CH-1′); 89.18 (CH-4′); 90.20 (C-26″); 99.99 (C-5); 145.45 (CH-6); 151.17 (C-2); 164.61 (C-4); 176.43 (C-24″). ν(KBr)/cm^{-1}: 3422, 3066, 2935, 2864, 2363, 2343, 1691, 1534, 1459, 1420, 1377, 1355, 1281, 1194, 1091, 1042, 986, 945, 922, 851, 774, 758; MS (ES$^+$): found m/z: 678.3 ([M + Na]), 1333.3 ([2M + Na]$^+$); HRMS (ES$^+$): found m/z: 656.3907 ($C_{36}H_{54}N_3O_8$ requires 656.3905).

Synthesis of modified dNTPs—Sonogashira cross-coupling—general procedure

Water–acetonitrile mixture (1.5 mL) was added through a septum to an argon-purged vial containing halogenated nucleoside triphosphate **dA^I**TP or **dC^I**TP, acetylene **2** (1.5 equiv.), and CuI (10 mol%). In a separate vial, $Pd(OAc)_2$ (5 mmol%) and TPPTS (5 equiv. to Pd) were combined under argon with 0.5 mL of H_2O–CH_3CN. After dissolution of the solids, the catalyst solution was added to the reaction mixture followed by addition of EtN(i-Pr)$_2$ (10 equiv.). The reaction mixture was stirred under argon at 75 °C for 45 min. The product was isolated from the crude reaction mixture by HPLC on a C18 column using two-step linear gradient from 0.1 M TEAB (triethylammonium bicarbonate) in H_2O to 0.1 M TEAB in H_2O–MeOH (1 : 1) and from 0.1 M TEAB in H_2O–MeOH (1 : 1) to MeOH as an eluent. Several co-distillations from water and conversion to sodium salt form (Dowex 50WX8 in Na^+ cycle) followed by freeze drying gave products as white solids.

7-{[(3α,7α,12α-trihydroxy-24-oxo-5β-cholan-24-yl)amino]prop-1-yn-1-yl}-7-deaza-2′-deoxyadenosinetriphosphate (dA^{CA}TP). Prepared from **dA^I**TP and **2d**, yield 44%.

MS (ES$^-$): found m/z: 465.8 ([M – 2H]$^{2-}$); HRMS (ES$^-$): found m/z: 465.6470 ($C_{38}H_{56}N_5O_{16}P_3$, z = 2 requires 465.6473). ^{31}P NMR (202.3 MHz, D_2O, pD = 7.1, ref (phosphate buffer) = 2.35 ppm): -21.19 (t, J = 20, Pβ); -9.92 (d, J = 20, Pα); -7.29 (d, J = 20, Pγ).

5-{[(3α,12α-dihydroxy-24-oxo-5β-cholan-24-yl)amino]prop-1-yn-1-yl}-2′-deoxycytidinetriphosphate (dC^{DCA}TP). Prepared from **dA^I**TP and **2d**, yield 53%.

MS (ES$^-$): found m/z: 406.3 ([M-PO$_3$-2H]$^{2-}$), 446.3 (100%, [M – 2H]$^{2-}$); HRMS (ES$^-$): found m/z: 446.1413 ($C_{36}H_{55}N_4O_{16}P_3$, z = 2 requires 446.1418). ^{31}P NMR (202.3 MHz, D_2O, pD = 7.1, ref (phosphate buffer) = 2.35 ppm): -21.34 (bdd, J = 19, 15, Pβ); -10.37 (d, J = 19, Pα); -7.23 (bd, J = 15, Pγ).

Primer extension, purification and analysis of the PEX products

Materials. Synthetic ONs were purchased from either VBC genomics (Austria) or from Sigma Aldrich. Primer: 5′-CAT GGG CGG CAT GGG-3′; templates: 5′-**CTA GCA TGA GCT CAG T***CC CAT GCC GCC CAT G*-3′ (temp^{md16}), 5′-**CCC G***CC CAT GCC GCC CAT G*-3′ (temp^{1C}), 5′-**CCC T***CC CAT GCC GCC CAT*

G-3′ (temp^{1A}), 5′-**GCG ACG AAG AGC TT***C CCA TGC CGC CCA TG*-3′ (temp^{AA}), 5′-**TTA TAT TTA TA***C CCA TGC CGC CCA TG*-3′ (temp^{3A}), 5′-**TAT ATA TAT AT***C CCA TGC CGC CCA TG*-3′ (temp^{rep}) (segments forming a duplex with the primer are in italics, the replicated segments are in bold). Templates used in experiments involving the DBstv magnetoseparation procedure were biotinylated at their 5′ ends. Dynabeads M-270 streptavidin (DBstv) were obtained from Dynal A.S. (Norway), DyNAzyme II and Phusion DNA Polymerases from Finnzymes (Finland), Pwo DNA Polymerase from PeqLab (Germany), unmodified nucleoside triphosphates (dATP, dTTP, dCTP, and dGTP) from Fermentas (USA), Vent (*exo*-) DNA Polymerase and T4 polynucleotide kinase from New England Biolabs (Great Britain) and γ-^{32}P-ATP from Izotop, Institute of isotopes Co, Ltd. (Hungary).

Primer extension experiments. The reaction mixture (20 μL) contained DNA polymerase (2 U), dNTPs (either natural or modified, 200 μM; composition of the dNTP is specified in the text and figure legends for individual experiments), primer (150 nM) and ON template (225 nM) in the reaction buffer. For polyacrylamide gel electrophoresis (PAGE) experiments, the primer was labelled using [γ^{32}P]-ATP according to standard techniques. Reaction mixtures were incubated at 60 °C for 30 min and at 95 °C for 5 min in a thermal cycler, and were stopped by addition of stop solution (40 μL, 80% [v/v] formamide, 20 mM EDTA, 0.025% [w/v] bromophenol blue, 0.025% [w/v] xylene cyanol). The reaction mixture was subjected to gel electrophoresis in 12.5% denaturing polyacrylamide gel containing 1xTBE buffer (pH 8) and 7% urea at 50–60 W for ~60 min. Gels were dried, and phosphoroimaged.

Polymerase chain reactions. The PCR reaction mixture (20 μL) contained DNA polymerase Phusion (2 U, Finnzymes, Finland), dNTPs (either natural (200 μM) or modified (600 μM)), DMSO (5%), formamide (0.25%), betaine (37.5 mM) and tetramethylammonium chloride (2.5 mM) primers LT25TH (400 nM, 5′-CAAGGACAAAATACCTGTATT CCTT-3′) and L20- (400 nM, 5′-GACATCATGAGAGACA TCGC-3′), template (25 nM, 5′-GACATCATGAGAGACAT CGCCTCTGGGCTAATAGGACTACTTCTAATCTGTAAGA-GCAGATCCCTGGACAGGCAAGGAATACAGGTATTTTG-TCCTTG-3′), in the Phusion reaction buffer HF 5× supplied by the manufacturer. 30 PCR cycles were run under the following conditions: denaturation for 1 min at 94 °C, annealing for 1 min at 55 °C, extension for 1.5 min at 72 °C, followed by final extension step of 5 min at 72 °C. PCR products were analyzed on a 2% agarose gel in 0.5 × TBE buffer, followed by staining with GelRed™.

Thermal denaturation studies. Double stranded ONs were prepared by PEX-reaction at a 500 μL scale using primer (2 μM), template (2 μM), Phusion DNA polymerase (0.2 U), and dNTPs (200 μM) in the reaction buffer. The PEX-products were purified using QIAquick Nucleotide Removal Kit Protocol (Qiagen). The melting temperature of oligonucleotides containing modified base or bases hybridized with natural DNA template were measured. The melting temperature of the control unmodified duplex was also determined in each case. The DNA duplexes were dissolved in 160 μL of phosphate buffer (10 mM) and 1 M NaCl (pH 7) and further diluted to final duplex concentrations of 0.73 ± 0.03 μM with the buffer. Melting curves were recorded on a Cary 100 bio UV/Vis instrument with temperature controller (Varian). Melting temperatures (T_m values in °C) were obtained by plotting temperature *versus* absorbance and by applying a sigmoidal curve fit and results are averages of 4–6 measurements.

Acknowledgements

This work is a part of the research projects Z4 055 0506 supported by the Academy of Sciences of the Czech Republic. It was specifically supported by the Ministry of Education (LC512), Czech Science Foundation (203/09/0317) and by Gilead Sciences, Inc. (Foster City, CA, U. S. A.). S. I. acknowledges the support from the Finnish Ministry of Education (The National Graduate School of Organic Chemistry and Chemical Biology) and Finnish Cultural Foundation.

Notes and references

1 (*a*) M. Famulok, *Curr. Opin. Mol. Ther.*, 2005, **7**, 137–143; (*b*) M. Famulok, J. S. Hartig and G. Mayer, *Chem. Rev.*, 2007, **107**, 3715–3743; (*c*) A. Peracchi, *ChemBioChem*, 2005, **6**, 1316–1322; (*d*) A. Condon, *Nat. Rev. Genet.*, 2006, **7**, 565–575; (*e*) F. E. Alemdaroglu and A. Herrmann, *Org. Biomol. Chem.*, 2007, **5**, 1311–1320; (*f*) U. Feldkamp and C. M. Niemeyer, *Angew. Chem., Int. Ed.*, 2006, **45**, 1856–1876.

2 Review: S. H. Weisbrod and A. Marx, *Chem. Commun.*, 2008, 5675–5685.

3 Recent examples: (*a*) T. Gourlain, A. Sidorov, N. Mignet, S. J. Thorpe, S. E. Lee, J. A. Grasby and D. M. Williams, *Nucleic Acids Res.*, 2001, **29**, 1898–1905; (*b*) L. H. Thoresen, G.-S. Jiao, W. C. Haaland, M. L. Metzker and K. Burgess, *Chem.–Eur. J.*, 2003, **9**, 4603–4610; (*c*) S. Jäger, G. Rasched, H. Kornreich-Leshem, M. Engeser, O. Thum and M. Famulok, *J. Am. Chem. Soc.*, 2005, **127**, 15071–15082; (*d*) M. Kuwahara, J. Nagashima, M. Hasegawa, T. Tamura, R. Kitagata, K. Hanawa, S. Hososhima, T. Kasamatsu, H. Ozaki and H. Sawai, *Nucleic Acids Res.*, 2006, **34**, 5383–5394; (*e*) A. Shoji, T. Hasegawa, M. Kuwahara, H. Ozaki and H. Sawai, *Bioorg. Med. Chem. Lett.*, 2007, **17**, 776–779; (*f*) S. Obeid, M. Yulikow, G. Jeschke and A. Marx, *Angew. Chem., Int. Ed.*, 2008, **47**, 6782–6785; (*g*) B. Holzberger and A. Marx, *Bioorg. Med. Chem.*, 2009, **17**, 3653–3658; (*h*) V. Borsenberger and S. Howorka, *Nucleic Acids Res.*, 2009, **37**, 1477–1485; (*i*) V. Borsenberger, M. Kukwikila and S. Howorka, *Org. Biomol. Chem.*, 2009, **7**, 3826–3835.

4 (*a*) J. D. Kahl and M. M. Greenberg, *J. Am. Chem. Soc.*, 1999, **121**, 597–604; (*b*) S. Kohgo, K. Shinozuka, H. Ozaki and H. Sawai, *Tetrahedron Lett.*, 1998, **39**, 4067–4070; (*c*) M. M. Masud, A. Ozaki-Nakamura, M. Kuwahara, H. Ozaki and H. Sawai, *ChemBioChem*, 2003, **4**, 584–588.

5 Review:(*a*) P. M. E. Gramlich, C. T. Wirges, A. Manetto and T. Carell, *Angew. Chem., Int. Ed.*, 2008, **47**, 8350–8358. Recent examples: (*b*) J. Gierlich, K. Gutsmiedl, P. M. E. Gramlich, A. Schmidt, G. A. Burley and T. Carell, *Chem.–Eur. J.*, 2007, **13**, 9486–9494; (*c*) 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; (*d*) P. M. E. Gramlich, S. Warncke, J. Gierlich and T. Carell, *Angew. Chem., Int. Ed.*, 2008, **47**, 3442–3444; (*e*) F. Seela, V. R. Sirivolu and P. Chittepu, *Bioconjugate Chem.*, 2008, **19**, 211–224; (*f*) F. Seela and V. R. Sirivolu, *Org. Biomol. Chem.*, 2008, **6**, 1674–1687.

6 (*a*) S. H. Weisbrod and A. Marx, *Chem. Commun.*, 2007, 1828–1830; (*b*) A. Baccaro, S. Weisbrod and A. Marx, *Synthesis*, 2007, 1949–1954.

7 (*a*) E. C. Western, J. R. Daft, E. M. Johnson, II, P. M. Garnnett and K. H. Shoughnessy, *J. Org. Chem.*, 2003, **68**, 6767–6774; (*b*) P. Čapek, R. Pohl and M. Hocek, *Org. Biomol. Chem.*, 2006, **4**, 2278–2284; (*c*) A. Collier and G. Wagner, *Org. Biomol. Chem.*, 2006, **4**, 4526–4532; (*d*) T. Pesnot and G. K. Wagner, *Org. Biomol. Chem.*, 2008, **6**, 2884–2891.

8 Recent review: M. Hocek and M. Fojta, *Org. Biomol. Chem.*, 2008, **6**, 2233–2241.

9 P. Čapek, H. Cahová, R. Pohl, M. Hocek, C. Gloeckner and A. Marx, *Chem.–Eur. J.*, 2007, **13**, 6196–6203.

10 P. Brázdilová, M. Vrábel, R. Pohl, H. Pivoňková, L. Havran, M. Hocek and M. Fojta, *Chem.–Eur. J.*, 2007, **13**, 9527–9533.

11 H. Cahová, L. Havran, P. Brázdilová, H. Pivoňková, R. Pohl, M. Fojta and M. Hocek, *Angew. Chem., Int. Ed.*, 2008, **47**, 2059–2062.
12 M. Vrábel, P. Horáková, H. Pivoňková, L. Kalachova, H. Černocká, H. Cahová, R. Pohl, P. Šebest, L. Havran, M. Hocek and M. Fojta, *Chem.–Eur. J.*, 2009, **15**, 1144–1154.
13 J. Riedl, P. Horáková, P. Šebest, R. Pohl, L. Havran, M. Fojta and M. Hocek, *Eur. J. Org. Chem.*, 2009, 3519–3525.
14 (*a*) H. Cahová, R. Pohl, L. Bednárová, K. Nováková, J. Cvačka and M. Hocek, *Org. Biomol. Chem.*, 2008, **6**, 3657–3660; (*b*) C. Lam, C. Hipolito and D. M. Perrin, *Eur. J. Org. Chem.*, 2008, 4915–4923; (*c*) H. Macíčková-Cahová and M. Hocek, *Nucleic Acids Res.*, 2009, **37**, 7612–7622.
15 H. Danielsson, in *The Bile Acids: Chemistry, Physiology and Metabolism*, ed. P. P. Nair and D. Kritchevsky, Plenum Press, New York, 1973, vol. 2, pp. 1–32.
16 (*a*) A. Enhsen, W. Kramer and G. Wess, *Drug Discovery Today*, 1998, **3**, 409–418; (*b*) S. Mukhopadhyay and U. Maitra, *Curr. Sci.*, 2004, **87**, 1666–1683; (*c*) A. Hofmann and R. L. Hagey, *Cell. Mol. Life Sci.*, 2008, **65**, 2461–2483.
17 (*a*) J. Tamminen and E. Kolehmainen, *Molecules*, 2001, **6**, 21–46; (*b*) E. Virtanen and E. Kolehmainen, *Eur. J. Org. Chem.*, 2004, 3385–1683; (*c*) A. P. Davis, *Molecules*, 2007, **12**, 2106–2122; (*d*) Nonappa and U. Maitra, *Org. Biomol. Chem.*, 2008, **6**, 657–669; (*e*) *Bile Acids: Toxicology and Bioactivity*, ed. G. J. Jenkins and L. J. Hardie, Royal Society of Chemistry, Cambridge, 2008.
18 R. L. Letsinger and S. Chaturvedi, *Bioconjugate Chem.*, 1998, **9**, 826–830.
19 (*a*) W. Kramer and G. Wess, *Eur. J. Clin. Invest.*, 1996, **26**, 715–732; (*b*) E. Sievänen, *Molecules*, 2007, **12**, 1859–1889.
20 A. Gissot, M. Camplo, M. W. Grinstaff and P. Barthélémy, *Org. Biomol. Chem.*, 2008, **6**, 1324–1333.
21 (*a*) M. Manoharan, L. K. Johnson, C. F. Bennet, T. A. Vickers, D. J. Ecker, L. M. Cowsert, S. M. Freier and P. D. Cook, *Bioorg. Med. Chem. Lett.*, 1994, **4**, 1053–1060; (*b*) M. Manoharan, G. Inamati, K. L. Tivel, B. Conklin, B. S. Ross and P. D. Cook, *Nucleosides, Nucleotides Nucleic Acids*, 1997, **16**, 1141.
22 (*a*) C. F. Bleczinski and C. Richert, *Rapid Commun. Mass Spectrom.*, 1998, **12**, 1737–1743; (*b*) C. F. Bleczinski and C. Richert, *J. Am. Chem. Soc.*, 1999, **121**, 10889–10894; (*c*) J. Tuma and C. Richert, *Biochemistry*, 2003, **42**, 8957–8965.
23 P. Baumhof, N. Griesang, M. Bächle and C. Richert, *J. Org. Chem.*, 2006, **71**, 1060–1067.
24 D. Bhatia, L. Yue-Ming and K. N. Ganesh, *Bioorg. Med. Chem. Lett.*, 1999, **9**, 1789–1794.
25 J. A. Rojas Stütz and C. Richert, *J. Am. Chem. Soc.*, 2001, **123**, 12718–12719.
26 N. G. Aher, V. S. Pore and S. P. Patil, *Tetrahedron*, 2007, **63**, 12927–12934.
27 N. S. Vatmurge, B. G. Hazra, V. S. Pore, F. Shirazi, P. S. Chavan and M. V. Deshpande, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 2043–2047.
28 A. Norman, *Ark. Kemi*, 1955, **8**, 331–342.
29 (*a*) M. Vrábel, R. Pohl, I. Votruba, M. Sajadi, S. A. Kovalenko, N. P. Ernsting and M. Hocek, *Org. Biomol. Chem.*, 2008, **6**, 2852–2860; (*b*) L. Kalachova, R. Pohl and M. Hocek, *Synthesis*, 2009, 105–112.
30 For a related examples of furopyrimidine formation, see: (*a*) C. J. Yu, H. Yowanto, Y. J. Wan, T. J. Meade, Y. Chong, M. Strong, L. H. Donilon, J. F. Kayyem, M. Gozin and G. F. Blackburn, *J. Am. Chem. Soc.*, 2000, **122**, 6767–6768; (*b*) A. R. Pike, L. C. Ryder, B. R. Horrocks, W. Clegg, M. R. J. Elsegood, B. A. Connolly and A. Houlton, *Chem.–Eur. J.*, 2002, **8**, 2891–2899.
31 J. M. Clark, *Nucleic Acids Res.*, 1988, **16**, 9677–9686.
32 F. Seela and M. Zulauf, *Helv. Chim. Acta*, 1999, **82**, 1878–1898.