Tetrahedron 70 (2014) 8843–8851

Contents lists available at ScienceDirect

Tetrahedron

journal homepage: www.elsevier.com/locate/tet

Synthesis of new biocarrier–nucleotide systems for cellular delivery in bacterial auxotrophic strains

Swarup De [a], Elisabetta Groaz [a], Mohitosh Maiti [a], Valérie Pezo [b,c], Philippe Marlière [c], Piet Herdewijn [a,c,*]

[a] Medicinal Chemistry, Rega Institute for Medical Research, KU Leuven, Minderbroedersstraat 10, 3000 Leuven, Belgium
[b] CEA, DSV, IG, Genoscope, 2 rue Gaston Crémieux, 91057 Evry Cedex, France
[c] ISSB, Génopole genavenir6, Equipe Xénome, 5 rue Henri Desbruères, 91030 Evry Cedex, France

ARTICLE INFO

Article history:
Received 30 July 2014
Received in revised form 24 September 2014
Accepted 30 September 2014
Available online 8 October 2014

Keywords:
Thymidine monophosphate
Biotin
Kanamycin A
Carboxamide linker
Oxymethyleneoxy linker
Carbamoyl linker

ABSTRACT

In search for a delivery approach for thymidine monophosphate (TMP) in bacterial cells, we have synthesized a series of conjugates of TMP with biotin having an oxymethyleneoxy ester, a carboxy ester, and different carboxamide linkers between the carboxyl group of biotin and the 3′-OH group of TMP. The synthetic strategy starts from 5′-*O*-(dibenzylphosphate)-thymidine having the linkers already connected at the 3′-position. Likewise, kanamycin A was linked at the 3′-position of TMP using a carbamoyl or thioethyl carbamoyl group. None of the conjugates were able to sustain growth of a ΔthyA, ΔphoA *Escherichia coli* strain.

1. Introduction

Creating artificial genetic systems capable of supporting Darwinian evolution is of prominent importance to synthetic biology, not only to achieve expansion of the genetic alphabet but specially for the genetic reprogramming of engineered microorganisms.[1] The in vivo synthesis of xeno nucleic acids (XNAs), however, is thought to entail high intracellular consumption of unnatural nucleotide building blocks, which will thus need to be provided to the host cell, as most modified nucleosides are poorly phosphorylated by natural kinases. Nucleotides, however, suffer from low cell-penetrating ability and generally undergo fast enzymatic dephosphorylation before being taken up. Their bacterial uptake is further complicated by the complex multi-layered wall structure, which makes prokaryotic cells commonly more difficult to penetrate compared to eukaryotic membranes.

Thus far, efforts have focused on overexpression of endogenous kinases to broad enzyme–substrate specificity,[2] and more recently, on the transfection of *Escherichia coli (E. coli)* with a plasmid encoding an algal nucleotide triphosphate transporter to enhance triphosphates uptake.[3] Despite the ground-breaking value of this last approach, which resulted in the first in vivo replication of an unnatural DNA base pair, its universal applicability to all unnatural bases and types of backbone modifications is yet unpredictable.

In a previous paper, we discussed our interest in the development of a delivery tool for active uptake of nucleotides in bacterial cells based on a nutritional selection approach.[4] The general concept consists in linking the information system to an extracellular nutrient essential for the cell, thus taking advantage of natural transport pathways. A biocleavable covalent bond is designed to maintain the conjugate intact during cellular uptake, but allows facile release of the nucleotide intracellularly. Building on previous studies, thymidine monophosphate is selected as archetypal nucleotide system.[5] The relevant TMP conjugate is supplied within the nutrient medium to a bacterial strain deleted for the ThyA gene encoding for thymidilate synthase, which is known to unconditionally require thymine or thymidine for growth.

Because of the novelty of our delivery approach, we decided to investigate a large diversity of biological carriers, such as peptides and vitamins. We describe here the synthesis of biotin–TMP conjugates. The appeal of this choice is due to the fact that biotin-negative mutants are also well-established auxotrophs and can provide additional functional evidence.[6,7]

Expanding upon this idea, we also wish to prove the same concept using an inverse selection method. As in most cases,

* Corresponding author. E-mail address: Piet.Herdewijn@rega.kuleuven.be (P. Herdewijn).

http://dx.doi.org/10.1016/j.tet.2014.09.096
0040-4020/

antimicrobial compounds penetrate prokaryotic cell membranes effectively through relatively non-specific transport systems, we describe here also the synthesis of conjugates of nucleotides with kanamycin A. Aminoglycosides (AGs) are clinically relevant drugs, which show broad-spectrum antibacterial activity against both gram-negative and gram-positive species, binding irreversibly to the prokaryotic ribosome and mainly interfering with peptide elongation at the 30S subunit.[8,9] Although the detailed mechanism by which AGs penetrate into the bacterial cytoplasm remains unclear, their uptake is known to be a multistep energy requiring process that can be accumulated against a concentration gradient.[10,11] In this case, the observation of toxicity would ensure uptake and processing of the nucleotide by the cell.

2. Results and discussion

Our retrosynthetic analysis for the planned carrier–TMP target compounds is shown in Fig. 1. It becomes apparent that those conjugates can be accessed by either linking the carrier unit at the 3′-position of thymidine followed by 5′-O-phosphorylation (route 1), or vice versa (route 2).

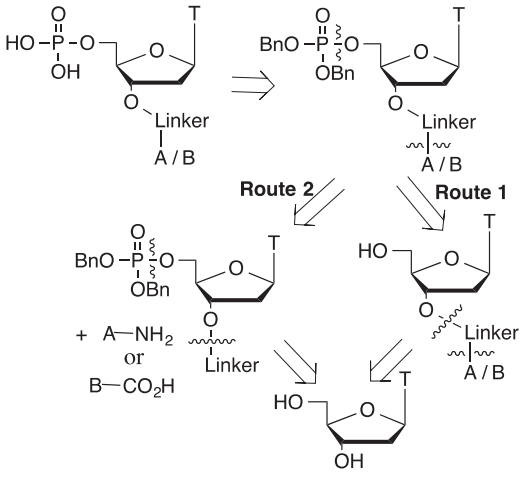


Fig. 1. Retrosynthetic analysis of carrier–TMP conjugates.

Initially, we decided to follow route 1 (Fig. 1) by applying the same optimized conditions employed for the synthesis of peptide–TMP conjugates connected though an oxymethyleneoxy ester bond.[4] Coupling of biotin at the 3′-position of methylene thiomethyl thymidine derivative **1**, obtained in turn from thymidine in 60% yield over two steps, and subsequent removal of the 5′-silyl group, proceeded efficiently giving compound **3** in good yield; however, further phosphorylation caused oxidation at the sulfur atom of biotin, during the conversion step from P^{III} to P^{V}, giving undesired sulfoxide and sulfone by-products (Scheme 1A). Various mild conditions were attempted to attain selective phosphorus oxidation, including dilute solutions of I_2,[12] *m*-chloroperbenzoic acid (MCPBA),[13] and (1*S*)-(+)-(10-camphorsulfonyl) oxaziridine[14] but none gave the desired product. Switching to the alternative disconnection, 5′-*O*-(dibenzylphosphate)-thymidine **5** was therefore prepared in four steps and subsequently subjected to coupling with biotin, leading to the formation of the desired compound **7** after deprotection (Scheme 1B). However, we found that the Pummerer rearrangement of compound **5** to **6** was less effective than in the previous route, probably due to side reaction of the electrophilic sulfur intermediate at the nucleophilic phosphate center.

During the last step, formation of side products resulting from conjugate decomposition was observed, pointing at instability of this type of linker under hydrogenation conditions.

Different biotin–TMP analogues with carboxyamido-type linkers **10** (Scheme 2) and **13** (Scheme 3) were prepared adopting previously optimized coupling and debenzylation procedures, which required access, respectively, to 3′-*O*-amino thymidylate **8** (obtained in four steps from thymidine) and 3′-*O*-aminoxymethyl thymidylate **11** (obtained in six steps from thymidine).[4] In general good to excellent isolated yields of products were obtained and linker cleavage did not occur in the last step.

The 3′-carboxy ester conjugate **15** was also prepared similarly, starting from **5** in 65% overall yield over two steps (Scheme 4).

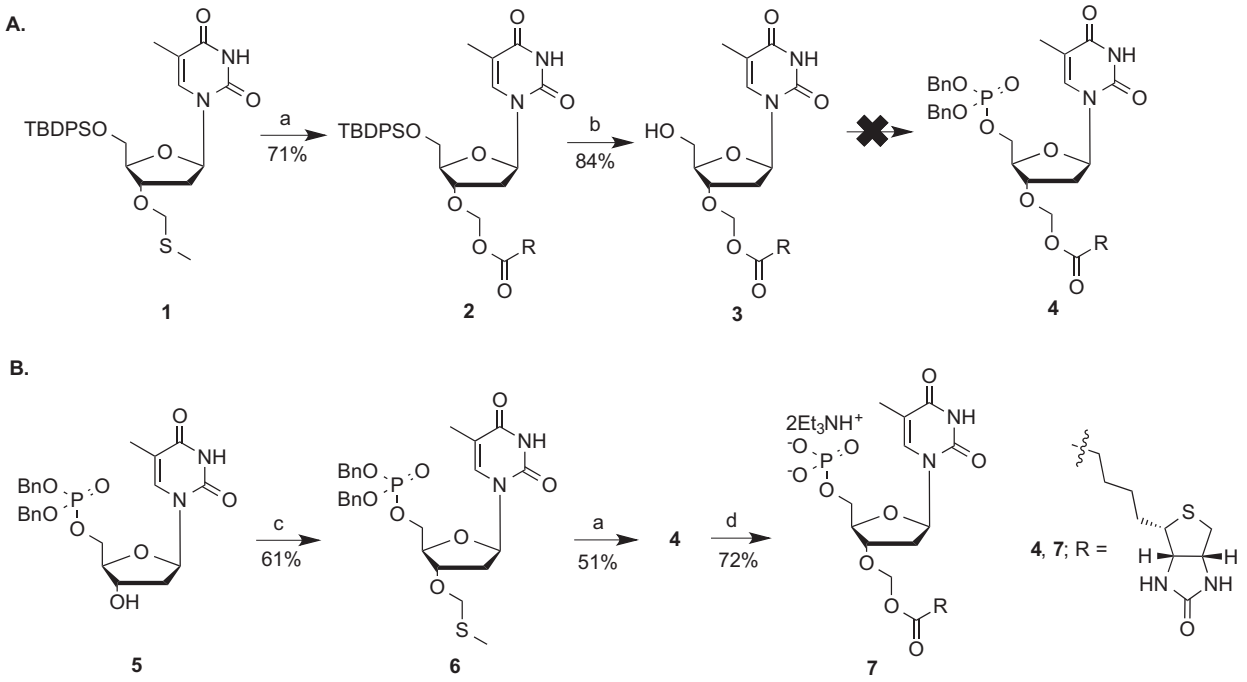
As route 2 (Fig. 1) proved to be the most profitable for generating new biotin conjugates, we chose a similar strategy to connect TMP to kanamycin A by means of a 3′-carbamoyl linker, which gave access to conjugates **22** and **24** (Fig. 2), and also led to a structural reassignment of the aminoglycosidic unit.

As cationic amino functionalities of aminoglycosides are considered to be a crucial structural requirement for enhanced uptake,[15] we sought to retain all four amino groups of rings I–III in kanamycin A. To construct the relevant conjugates efficiently, we exploited the differences in chemical reactivity between the primary hydroxy and amino functionalities present in kanamycin A. We started by replacing the primary hydroxyl group at the 6″-position of **16** with a thioethanolamino group, which would serve for the formation of the desired carbamoyl linkage with TMP. The synthesis of compound **19** has been reported by Arya et al.,[16] however, we adopted here a shorter synthetic route by modifying a more recent procedure by Tor et al.[17] (Scheme 5). To begin the synthesis, we first needed to selectively protect all four amino groups of kanamycin A. Among the several protecting strategies described in the literature, we opted for the classical Boc protection, which gave compound **17** in excellent yield. Considering the good leaving group properties of triisopropylbenzenesulfonyl chloride (TIBS-Cl), compound **17** was selectively sulphonylated at the 6″-primary hydroxyl group to its corresponding sulfonate **18**, which upon further treatment with 2-aminoethane thiol hydrochloride in the presence of cesium carbonate, gave compound **19**.

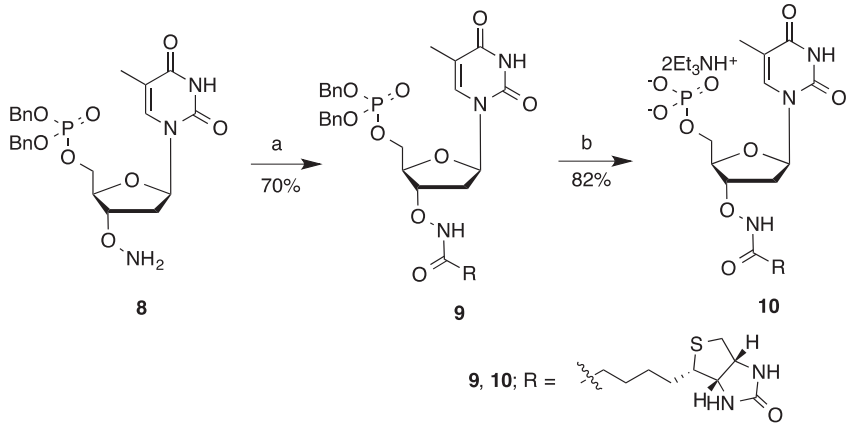
The introduction of a carbamate linker at the 3′-position of nucleosides is possible starting from commercially available isocyanates such as phenyl isocyanates,[18] however, this approach appears unsuitable for macromolecules containing unprotected polyhydroxyl and polyamino functionalities. Another commonly applied method to generate carbamates is the activation of the 3′-OH group with CDI,[19] but we felt that the strictly anhydrous conditions required by this protocol would not be compatible with the poor solubility of most aminoglycosides. A solution was found by choosing as activating/leaving group 4-nitrophenyl carbonate,[20] which is electrophilic enough to react with amino functionalities, whilst still reasonably stable in aqueous media. Compound **20** was prepared in a series of straightforward steps, involving phosphorylation at the 5′-position of thymidine after selective protection–deprotection at the 3′- and 5′-position with a MMTr and benzoyl group, respectively, and final 3′-OH activation with 4-nitrophenyl chloroformate in a controlled way in the presence of pyridine/DCM.

With both coupling units now available, we proceeded to react 3′-*O*-(4-nitrophenylcarbonate)–TMP **20** with *N*-Boc kanamycin A cysteamine **19** (Scheme 6A).

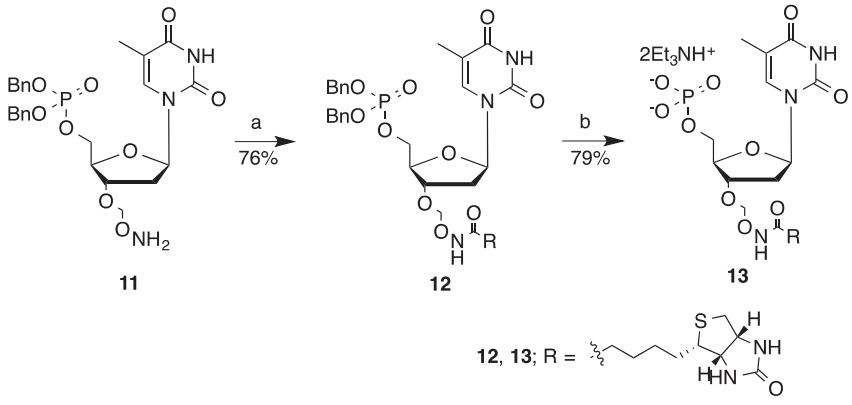
An initial attempt to synthesize **21** using previously optimized conditions in the presence of triethylamine as base and DCM/DMF (1:1) as solvent, afforded the desired conjugate in poor yield. Replacement of the solvent system with a mixture of 1,4-dioxane/water (3:1), surprisingly resulted in an improved yield (90%) (Table 1). Removal of the dibenzyl phosphate protection of **21** was



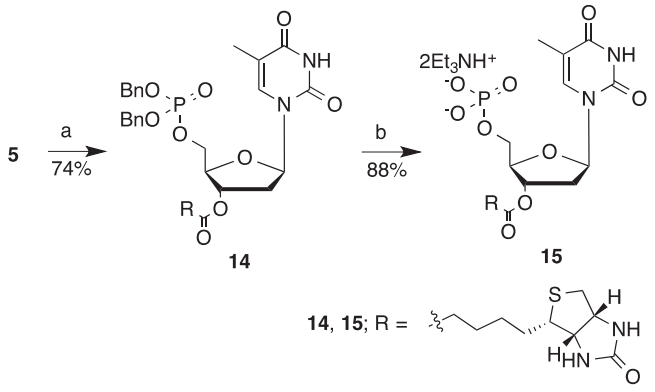
Scheme 1. Synthesis of 3′-*O*-methyleneacyloxy-TMP biotin conjugate **7**. Reagents and conditions: (a) (i) SO_2Cl_2, DCM, 0 °C to rt, 2 h, (ii) biotin, DBU, DCM, rt, 24 h; (b) $Et_3N\cdot 3HF$, THF, rt, 72 h; (c) DMSO, Ac_2O, AcOH, rt, 48 h; (d) 10% Pd/C (Degussa), Et_3N, MeOH, rt, 16 h.



Scheme 2. Synthesis of 3′-*O*-(carboxamide)-TMP biotin conjugate **10**. Reagents and conditions: (a) biotin, DCC, DMAP, DCM/DMF, rt, 24 h; (b) 10% Pd/C (Degussa), MeOH, rt, 24 h.



Scheme 3. Synthesis of 3′-*O*-methyloxycarboxamide-TMP biotin conjugate **6**. Reagents and conditions: (a) biotin, DCC, DMAP, DCM/DMF, rt, 24 h; (b) 10% Pd/C (Degussa), Et_3N, MeOH, rt, 24 h.



Scheme 4. Synthesis of 3′-ester-TMP biotin conjugate **15**. Reagents and conditions: (a) biotin, DCC, DMAP, DCM/DMF, rt, 24 h; (b) 10% Pd/C (Degussa), MeOH, rt, 24 h.

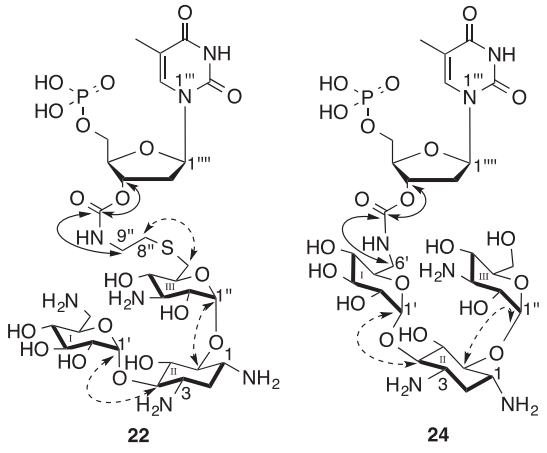
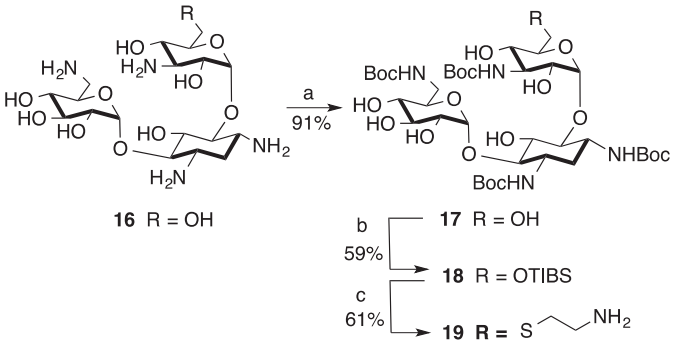


Fig. 2. Structures of kanamycin A–TMP conjugates **22** and **24** and key HMBC correlations used for structure determination. Plain arrows (↔) indicate HMBC correlations with the carbamate (OCONH) linker (cf. Fig. 3A/B) whilst dotted arrows (↔) show HMBC correlation used for other peak assignments.



Scheme 5. Synthesis of kanamycin A cysteamine **19**. Reagents and conditions: (a) Boc_2O, Et_3N, DMF/H_2O, 50 °C, 6 h; (b) TIBSCl, pyridine, rt, 16 h; (c) $HCl{\cdot}NH_2CH_2CH_2SH$, Cs_2CO_3, DMF, 25 °C, 16 h.

achieved by catalytic hydrogenolysis using an excess of 20% $Pd(OH)_2/C$ to avoid potential catalyst poisoning due to the sulfide moiety. Without further purification, the *N*-Boc protected intermediate **21** was converted to compound **22** upon treatment with aq TFA, in good yield and high purity after preparative RP-HPLC purification.

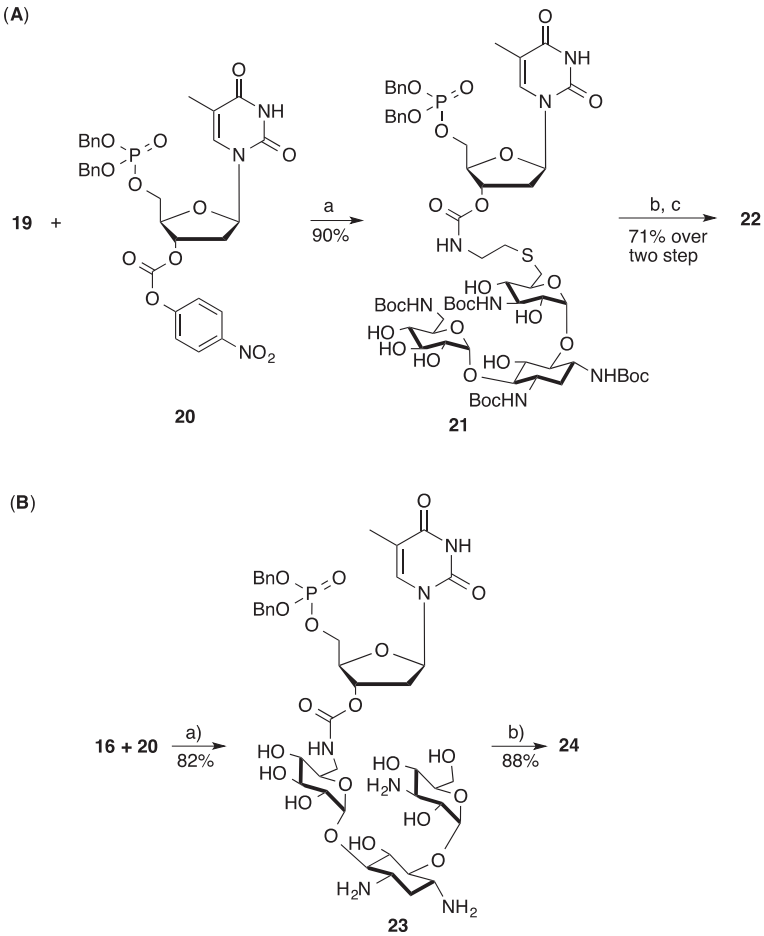
As the convergent coupling of complex fragments is important to rapidly generate a variety of conjugates, we decided to further explore the versatility of this reaction process. Since a single aminomethyl group is present in kanamycin A, the regioselective formation of a mono-carbamate linker seemed feasible without prior protection of the other amino groups. The 6′-amino group of kanamycin A is known to be the most reactive following the series 6′-NH_2>1-NH_2>3-NH_2>3″-NH_2.[21,22] When we reacted directly compound **20** with kanamycin A **16** utilizing the general protocol for carbamate formation (Scheme 6B), intermediate **23** was obtained in good yield. Subsequent hydrogenolysis in the presence of 20% $Pd(OH)_2/C$ and preparative RP-HPLC purification yielded compound **24** in good yield and high purity.

Although the reactivity order of the various hydroxyl and amino groups of kanamycin A is well documented, we thought necessary to confirm the structures of the final analogues. The peak assignment for almost all proton and carbon signals of kanamycin A–TMP conjugates **22** and **24** was unambiguously determined (Table 2, Supplementary data), as well as the alignment of the carbamoyl linkage between the 9″ and 6′ positions (Fig. 3A and B), by means of two-dimensional NMR spectroscopy in D_2O solution at 25 °C. The presence of a distinctive signal at 157.5 ppm in the carbonyl region of the carbon NMR spectrum, proton integration ratio along with HRMS mass spectrometry confirmed the formation of a single carbamate linkage at the expected positions. While performing the ^1H and ^{31}P NMR spectroscopic analysis of compound **22** (Fig. 3C and D), the appearance of a minor conformation was observed at low temperature (5 °C), which disappeared when the temperature was raised (50 °C), indicating the presence of a single conformer. This phenomenon may arise due to the formation of rotamers, as the inherent flexibility around glycosidic bonds of kanamycin A–TMP conjugates allows them to easily adopt an array of conformations.[23]

In order to test the bacterial uptake of all synthetized biotin and kanamycin A conjugates as a source of TMP, an *E. coli* W2666 thyA(−) strain [spontaneous strepto R mutant of CBO129 (F^- W1485 leu thyA dcoB or C supE)] was employed. Its genotype is *E. coli* W2666 ΔthyA ΔphoA. This strain is deleted for thyA and phoA genes, which, respectively, encode for thymidylate synthase and alkaline phosphatase, in order to block all pathways that could lead to the formation or degradation of TMP inside the bacterial cells. All compounds were preventively submitted to LC/MS analysis to exclude potential sample contamination, but no traces of thymine could be observed below detection limit. LB medium and LB agar were sterilized by autoclaving (120 °C, 30 min). Thymine (positive control), TMP (negative control), and all compounds were dissolved in MilliQ water and sterilized through a 0.2 μm sterile syringe filter (VWR). A single colony of the mutant strain was grown overnight at 37 °C in LB medium (3 mL). Cells were then harvested by centrifugation, washed and resuspended in fresh LB medium. The bacterial cells were then diluted 1:1000 in fresh LB medium before streaking on LB agar plates containing 20 μg/mL of each compound and grown overnight at 37 °C. Unfortunately, no colony formation (compounds **7**, **10**, **13**, **15**) or toxicity (compounds **22**, **24**) was observed.

3. Conclusion

In summary, we have accomplished the synthesis of a variety of biotin–TMP and kanamycin A–TMP conjugates in a straightforward manner starting from previously optimized 3′-modified TMP intermediates and featuring in turn an oxymethyleneoxy ester, a carboxy ester, a carboxamide, a carbamoyl, and a thioethyl carbamoyl group. On the basis of preliminary biological screening none of the conjugates showed functional activity in *E. coli* mutants lacking thymine. However, this might be due to several factors, for instance high concentrations needed for substantial growth, which are currently under further study. As additional evidence, biotin–TMP conjugates are undergoing biological evaluation using *E. coli* mutants lacking biotin.



Scheme 6. Synthesis of 3′-carbamoyl kanamycin A–TMP conjugates **22** and **24**. Reagents and conditions: (a) Et_3N, 1,4-dioxane/water, 0 °C to rt, 18 h; (b) 20% $Pd(OH)_2/C$, H_2, EtOH/H_2O (8:2), rt, 24 h; (c) TFA, thioanisole, H_2O, rt, 5 h.

Table 1
Reaction circumstances for the synthesis of compounds **21** and **23** from **20**

Entry	Solvent	Temp	Time (h)	Yield of **21**[a] (%)	Yield of **23**[a] (%)
1	DCM/DMF (1:1)	rt	60	48	—
2	DMF	rt	48	55	42
3	Dioxane/H_2O (3:1)	0 °C to rt	16	90	82

[a] Isolated yield.

4. Experimental section

4.1. General

For all reactions, analytical grade solvents were used. Kanamycin A·H_2SO_4 was purchased from Sigma–Aldrich (containing kanamycin B as a 5% impurity). Kanamycin A·H_2SO_4 salt was transformed into its free amino form by treatment with Amberlite IRA-400 ion exchange resin. All moisture-sensitive reactions were carried out using oven-dried glassware (135 °C) under a nitrogen or argon atmosphere. Reaction temperatures are reported as bath temperatures. Pre-coated aluminum sheets (254 nm) were used for

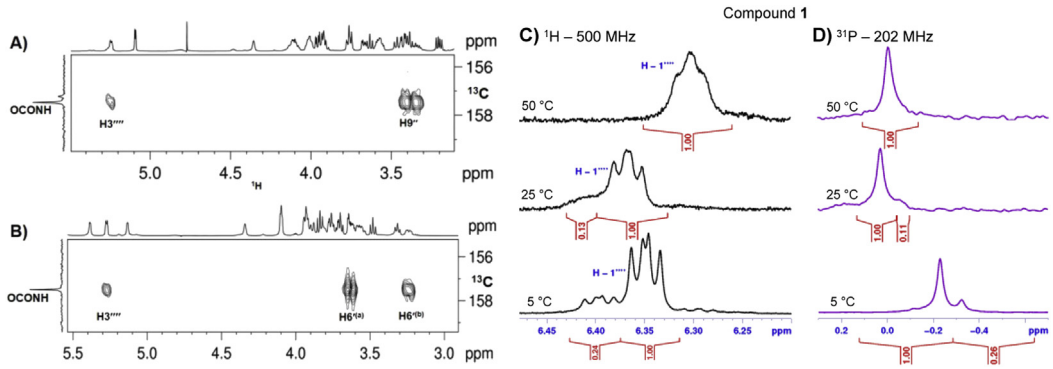


Fig. 3. (A) Key HMBC correlation for compound **22**; (B) key HMBC correlation NMR spectrum of compound **24**; (C) and (D) temperature-dependent NMR (^1H and ^{31}P) study of compound **22**. All NMR spectra were recorded in D_2O.

TLC. Compounds were visualized with UV light (λ=254 nm). Products were purified by flash chromatography on ICN silica gel 63–200, 60 Å. All final compounds were purified by preparative RP-HPLC using a gradient of H_2O and MeCN, both contain either 50 mmol TEAB or 0.1% TFA as eluant buffer as mentioned. ^1H, ^{13}C, and ^{31}P NMR spectra were recorded on 300, 500, and 600 MHz Bruker spectrometers. Final compounds were characterized using 2D NMR (H-COSY, HSQC, HMBC, TOCSY, NOESY) techniques. The ^1H and ^{13}C chemical shifts were referenced to residual solvent signals relative to TMS as internal standard wherever applied. Coupling constants *J* [Hz] were directly taken from the spectra. Splitting patterns are designated as s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), and br (broad). High-resolution mass spectra were acquired on a quadrupole orthogonal acceleration time-of-flight mass spectrometer (Synapt G2 HDMS, Waters, Milford, MA). Samples were infused at 3 μL/min and spectra were obtained in positive (or negative) ionization mode with a resolution of 15,000 (FWHM) using leucine enkephalin as lock mass.

4.2. 5′-*O*-(*tert*-Butyl diphenyl silane)-3′-*O*-(biotinmethyloxyester)-thymidine (2)

A 1 M solution of SO_2Cl_2 in CH_2Cl_2 (2.07 mL, 2.07 mmol) was added to a stirred solution of **1** (0.935 g, 0.107 mmol) in dry CH_2Cl_2 (20 mL) at 0 °C. The reaction mixture was allowed to warm slowly to room temperature over 2 h, then it was concentrated using a rotary evaporator (bath temperature 10–15 °C) under reduced pressure to give a 3′-*O*-chloromethylthymidine derivative as a light yellowish foam, which was used without further purification.

In a separate round-bottomed flask, biotin (0.634 g, 2.593 mmol) was suspended in dry CH_2Cl_2 (20 mL), and DBU (0.38 mL, 2.51 mmol) was then added. After 15 min, the solution was added to the crude 3′-*O*-chloromethylthymidine derivative (redissolved in dry CH_2Cl_2, 20 mL), and the reaction mixture was kept stirring at room temperature for 24 h. The mixture was then diluted with CH_2Cl_2 (100 mL), and 1 M aq acetic acid (4 mL) was added with vigorous stirring. The aqueous layer was discarded, and the organic layer was washed with saturated aq $NaHCO_3$ (2×50 mL) and brine, dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude residue was purified by column chromatography on silica gel (gradient CH_2Cl_2/MeOH, 99:1, v/v; 98:2, v/v; 97:3, v/v) to give **2** (0.91 g, 71%) as a colorless foam. ^1H NMR (300 MHz, $CDCl_3$) δ=11.9 (s, 1H, N*H*-Thy), 7.68–7.65 (m, 4H, Ar–*H* Ph), 7.51 (s, 1H, H-6), 7.46–7.37 (m, 6H, Ar–*H* Ph), 6.94 (s, 1H, N*H* Bio), 6.91 (s, 1H, N*H* Bio), 6.27 (dd, *J*=8.6, 5.2 Hz, 1H, H-1′), 5.71 (m, 1H, H-g Bio), 4.92 (m, 1H, H-f Bio), 4.59–4.54 (m, 1H, H-3′), 4.38–4.35 (m, 2H, H-5′ and H-5″), 4.12–4.10 (m, 1H, H-4′), 4.03–3.82 (m, 2H, OC*H*$_2$O), 3.21–3.14 (m, 1H, H-e Bio), 2.96–2.82 (m, 2H, H-h Bio and H-h′ Bio), 2.74–2.68 (m, 1H, H-2′), 2.49–2.28 (m, 2H, H-a Bio), 2.14–2.05 (m, 1H, H-2″), 1.83–1.65 (m, 4H, H-b and H-d Bio), 1.62 (s, 3H, C*H*$_3$), 1.55–1.40 (m, 2H, H-c Bio), 1.11 (s, 9H, ^{t}Bu); ^{13}C NMR (75 MHz, $CDCl_3$) δ=172.6 (OCO), 165.0 (C-4), 164.3 (NHCONH), 151.4 (C-2), 135.4 (Ar–C), 135.2 (Ar–C), 134.3 (C-6), 132.6 (Ar–C), 132.2 (Ar–C), 130.1 (Ar–C), 130.0 (Ar–C), 128.0 (Ar–C), 127.9 (Ar–C), 111.2 (C-5), 88.1 (OCH_2O), 85.1 (C-1′), 84.5 (C-4′), 81.6 (C-3′), 64.0 (C-5′), 62.1 (C-f Bio), 60.3 (C-g Bio), 55.4 (C-e Bio), 40.7 (C-h, h′ Bio), 38.9 (C-2′), 33.8 (C-a Bio), 28.0 (C-c Bio), 27.9 (C-d Bio), 26.9 (^{t}Bu), 24.5 (C-b Bio), 19.2 (1C ^{t}Bu), 12.1 (CH$_3$–Thy); HRMS for $C_{37}H_{48}N_4O_8SSi$ $[M+H]^+$ calcd: 737.3035, found: 737.3045.

4.3. 3′-*O*-(Biotinmethyloxyester)-thymidine (3)

Triethylamine trihydrofluoride (0.76 mL, 4.67 mmol) was added to a stirred solution of **2** (0.86 g, 1.17 mmol) in THF (20 mL) in a Falcon tube at room temperature. The reaction mixture was stirred at room temperature for 72 h, and then it was concentrated under reduced pressure. The crude residue was purified by column chromatography on silica gel (gradient CH_2Cl_2/MeOH, 99:1, v/v; 96:4, v/v; 93:7, v/v) to give **3** (0.49 g, 84%) as a white solid. ^1H NMR (300 MHz, DMSO-d_6) δ=11.34 (s, 1H, N*H*-Thy), 7.68 (s, 1H, H-6), 6.42 (s, 1H, N*H* Bio), 6.37 (s, 1H, N*H* Bio), 6.01 (app t, *J*=7.0 Hz, 1H, H-1′), 5.30 (dd, *J*=20.2, 6.6 Hz, 2H, OCH_2O), 5.13 (t, *J*=5.1 Hz, 1H, 5′-OH), 4.39–4.36 (m, 1H, H-3′), 4.33–4.28 (m, 1H, H-g Bio), 4.15–4.10 (m, 1H, H-f Bio), 3.93–3.90 (m, 1H, H-4′), 3.61–3.56 (m, 2H, H-5′ and H-5″), 3.13–3.06 (m, 1H, H-e Bio), 2.82 (dd, *J*=12.4, 5.1 Hz, 1H, H-h Bio), 2.57 (d, *J*=12.4 Hz, 1H, H-h′ Bio), 2.36 (t, *J*=7.4 Hz, 1H, and H-a Bio), 2.27–2.18 (m, 2H, H-2′ and H-2″), 1.77 (s, 3H, C*H*$_3$), 1.66–1.45 (m, 4H, H-b and H-d Bio), 1.42–1.29 (m, 2H, H-c Bio); ^{13}C NMR (75 MHz, DMSO-d_6) δ=172.4 (OCO), 163.6 (C-4), 162.7 (NHCONH), 150.5 (C-2), 135.9 (C-6), 109.5 (C-5), 87.4 (OCH_2O), 84.8 (C-1′), 83.6 (C-4′), 79.4 (C-3′), 61.1 (C-5′), 61.0 (C-f Bio), 59.2 (C-g Bio), 55.3 (C-e Bio), 40.2 (C-h, h′ Bio, merged with DMSO), 36.9 (C-2′), 33.3 (C-a Bio), 28.0 (C-c Bio), 27.9 (C-d Bio), 24.2 (C-b Bio), 12.2 (CH$_3$–Thy); HRMS for $C_{21}H_{30}N_4O_8S$ $[M-H]^-$ calcd: 497.1711, found: 497.1715.

4.4. 5′-*O*-(Dibenzylphosphate)-3′-*O*-(methylthiomethyl)-thymidine (6)

Acetic anhydride (10.9 mL) and acetic acid (3.5 mL) were added to a stirred solution of **5** (2.40 g, 4.78 mmol) in DMSO (15.5 mL). The reaction mixture was stirred at room temperature for 30 h, and then it was concentrated under reduced pressure. The residue was neutralized with saturated aq $NaHCO_3$ (230 mL), and the mixture was extracted with ethyl acetate (3×170 mL). The combined organic extracts were washed with brine, dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude residue was purified by column chromatography on silica gel (gradient hexane/EtOAc, 2:1, v/v; 1:1, v/v; 1:3, v/v) to give **6** (1.64 g, 61%) as a colorless foam. ^1H NMR (300 MHz, $CDCl_3$) δ=8.22 (br s, 1H, N*H*), 7.37–7.35 (m, 11H, H-6 and Ar–*H* OBn), 6.26 (dd, *J*=8.2, 5.8 Hz, 1H, H-1′), 5.11–5.00 (m, 4H, 2×OCH_2Ph), 4.64–4.51 (m, 2H, OCH_2S), 4.35–4.31 (m, 1H, H-3′), 4.17–4.09 (m, 3H, H-4′, H-5′ and H-5″), 2.34–2.26 (m, 1H, H-2′), 2.11 (s, 3H, SCH_3), 1.92–1.81 (m, 4H, H-2″ and CH_3–Thy); ^{13}C NMR (75 MHz, $CDCl_3$) δ=163.4 (C-4), 150.2 (C-2), 135.6 (d, $^3J_{C,P}$=6.1 Hz, 1C of OCH_2Ph), 135.3 (C-6), 129.0 (Ar–C), 128.9 (Ar–C), 128.2 (Ar–C), 111.5 (C-5), 85.0 (C-1′), 82.8 (d, $^3J_{C,P}$=8.2 Hz, C-4′), 76.3 (C-3′), 74.1 (OCH_2S), 69.9–69.8 (2×d, $^2J_{C,P}$=3.5 Hz, 2×OCH_2Ph), 67.0 (d, $^2J_{C,P}$=5.7 Hz, C-5′), 37.5 (C-2′), 14.0 (SCH_3), 12.5 (CH$_3$–Thy); ^{31}P NMR (121 MHz, $CDCl_3$) δ=−0.5; HRMS for $C_{26}H_{31}N_2O_8PS$ $[M-H]^-$ calcd: 561.1466, found: 561.1461.

4.5. 5′-*O*-(Dibenzylphosphate)-3′-*O*-(biotinmethyloxyester)-thymidine (4)

Following a similar procedure as the one used for the synthesis of **2**, compound **4** was obtained (41.0 mg, 51%) as a colorless foam, starting from **6** (60.0 mg, 0.107 mmol), a 1 M solution of SO_2Cl_2 in CH_2Cl_2 (0.13 mL, 0.128 mmol) in dry CH_2Cl_2 (4 mL), and biotin (39.1 mg, 0.160 mmol), DBU (0.023 mL, 0.155 mmol) in dry CH_2Cl_2 (15 mL). The crude residue was purified by column chromatography on silica gel (gradient CH_2Cl_2/MeOH, 99:1, v/v; 97:3, v/v; 93:7, v/v). ^1H NMR (300 MHz, $CDCl_3$) δ=9.76 (s, 1H, N*H*–Thy), 7.35–7.37 (m, 11H, H-6 and Ar–*H* OBn), 6.81 (br s, 1H, N*H* Bio), 6.31 (app t, *J*=6.6 Hz, 1H, H-1′), 6.18 (br s, 1H, N*H* Bio), 5.13–4.99 (m, 6H, OCH_2O and 2×OCH_2Ph), 4.53–4.49 (m, 1H, H-3′), 4.38–4.35 (m, 1H, H-g Bio), 4.26–4.12 (m, 3H, H-f Bio, H-5′ and H-5″), 4.05–4.03 (m, 1H, H-4′), 3.16–3.12 (m, 1H, H-e Bio), 2.89 (dd, *J*=12.2, 7.2 Hz, 1H, H-h Bio), 2.79 (d, *J*=12.2 Hz, 1H, H-h′ Bio), 2.58–2.52 (m, 1H, H-2′), 2.40–2.27 (m, 3H, H-2″ and H-a Bio), 1.80 (s, 3H, C*H*$_3$), 1.76–1.69 (m, 4H, H-b and H-d Bio), 1.52–1.43 (m, 2H, H-c Bio); ^{13}C NMR (75 MHz, $CDCl_3$) δ=172.7 (OCO), 164.7 (C-4), 164.2 (NHCONH), 150.7 (C-2), 135.5 (d, $^3J_{C,P}$=8.0 Hz, 1C of OCH_2Ph), 134.5 (C-6), 128.8 (Ar–C),

128.7 (Ar–C), 128.1 (Ar–C), 111.3 (C-5), 88.1 (OCH_2O), 84.8 (C-1′), 83.1 (d, $^3J_{C,P}$=8.1 Hz, C-4′), 81.3 (C-3′), 69.9–69.8 (2×d, $^2J_{C,P}$=5.2 Hz, 2×OCH_2Ph), 67.0 (d, $^2J_{C,P}$=5.9 Hz, C-5′), 62.3 (C-f Bio), 60.2 (C-g Bio), 55.6 (C-e Bio), 40.2 (C-h, h′ Bio), 38.6 (C-2′), 33.9 (C-a Bio), 28.2 (C-c Bio), 28.0 (C-d Bio), 24.8 (C-b Bio), 12.5 (CH_3–Thy); ^{31}P NMR (121 MHz, $CDCl_3$) δ=−0.4; HRMS for $C_{35}H_{43}N_4O_{11}PS$ $[M+Na]^+$ calcd: 781.2279, found: 781.2288.

4.6. 3′-*O*-(Biotinmethyloxyester)-thymidine monophosphate triethylammonium salt (7)

Pd/C (10%) (Degussa, 15.0 mg, 50% w/w) was added to a degassed stirred solution of **4** (30.0 mg, 0.039 mmol) and Et_3N (0.011 mL, 0.079 mmol) in MeOH (10 mL), and the mixture was subjected to hydrogenation at atmospheric pressure using a balloon filled with H_2 for 24 h. The catalyst was removed by filtration through a cellulose filter (0.45 μm) and the filtrate was concentrated under reduced pressure (bath temperature 10–15 °C). The resulting crude residue was purified by RP-HPLC [TEAB (triethylammonium bicarbonate; 50 mmol) in H_2O/MeCN, 98:2; and TEAB (50 mmol) in H_2O/MeCN, 50:50]. The collected eluate was freeze-dried repeatedly until constant mass to obtain the triethylammonium salt of **7** (22.2 mg, 72%) as a white solid. ^{1}H NMR (500 MHz, D_2O) δ=7.80 (s, 1H, H-6), 6.28 (app t, *J*=7.3 Hz, 1H, H-1′), 5.46–5.32 (m, 2H, O*CH*$_2$O), 4.63–4.62 (m, 1H, H-3′), 4.56–4.54 (m, 1H, H-g Bio), 4.38–4.36 (m, 1H, H-f Bio), 4.25 (br s, 1H, H-4′), 3.94–3.88 (m, 2H, H-5′ and H-5″), 3.29–3.25 (m, 1H, H-e Bio), 2.92 (dd, *J*=13.0, 5.0 Hz, 1H, H-h), 2.70 (d, 1H, *J*=13.0 Hz, H-h′ Bio), 2.45 (t, 2H, *J*=7.2 Hz, H-a Bio), 2.42–2.40 (m, 1H, H-2′ and H-2″), 1.88 (s, 3H, C*H*$_3$), 1.73–1.52 (m, 4H, H-b and H-d Bio), 1.43–1.36 (m, 2H, H-c Bio); ^{13}C NMR (125 MHz, D_2O) δ=175.5 (OCO), 168.3 (C-4), 164.8 (NHCONH), 152.9 (C-2), 137.0 (C-6), 111.4 (C-5), 88.1 (OCH_2O), 84.3 (C-1′), 83.8 (d, $^3J_{C,P}$=9.2 Hz, C-4′), 80.7 (C-3′), 63.3 (d, $^2J_{C,P}$=3.9 Hz, C-5′), 61.5 (C-f Bio), 59.7 (C-g Bio), 54.8 (C-e Bio), 39.2 (C-h, h′ Bio), 36.6 (C-2′), 33.1 (C-a Bio), 27.4 (C-c), 27.1 (C-d Bio), 23.4 (C-b Bio), 11.5 (CH_3–Thy); ^{31}P NMR (202 MHz, D_2O) δ=3.4; HRMS for $C_{21}H_{31}N_4O_{11}PS$ $[M-H]^-$ calcd: 577.1375, found: 577.1375.

4.7. 5′-*O*-(Dibenzylphosphate)-3′-*O*-(biotincarboxamide)-thymidine (9)

DCC (127.4 mg, 0.618 mmol) was added to a stirred solution of **8** (200.0 mg, 0.386 mmol), biotin (108.6 mg, 0.444 mmol), and DMAP (2.0 mg, 0.016 mmol) in a mixture of dry DMF (2.0 mL) and dry CH_2Cl_2 (10 mL) at 0 °C. The reaction mixture was allowed to warm slowly to room temperature and stirred for 24 h. All the volatiles were removed under reduced pressure and the resulting residue was purified by column chromatography on silica gel (gradient CH_2Cl_2/MeOH, 99:2, v/v; 97:4, v/v; 94:7, v/v) to give **9** (201.0 mg, 70%) as a colorless foam. ^{1}H NMR (300 MHz, $CDCl_3$+CD_3OD) δ=7.41 (d, *J*=1.0 Hz, 1H, H-6), 7.39–7.34 (m, 10H, Ar–*H* OBn), 6.27 (dd, *J*=8.9, 5.6 Hz, 1H, H-1′), 5.12–5.00 (m, 4H, 2×O*CH*$_2$Ph), 4.55–4.48 (m, 2H, H-3′ and H-g Bio), 4.33–4.29 (m, 2H, H-4′ and H-f Bio), 4.23–4.22 (m, 2H, H-5′ and H-5″), 3.21–3.15 (m, 1H, H-e Bio), 2.91 (dd, *J*=12.8, 4.9 Hz, 1H, H-h Bio), 2.73 (d, *J*=12.8 Hz, 1H, H-h′ Bio), 2.54–2.48 (m, 1H, H-2′), 2.14 (t, *J*=6.7 Hz, 3H, H-a Bio), 1.94–1.84 (m, 1H, H-2″), 1.81 (d, *J*=0.9 Hz, 3H, C*H*$_3$), 1.75–1.61 (m, 4H, *H*-b and *H*-d Bio), 1.51–1.43 (m, 2H, H-c Bio); ^{13}C NMR (75 MHz, $CDCl_3$+CD_3OD) δ=172.8 (ONHCO), 165.1 (C-4), 164.9 (NHCONH), 151.3 (C-2), 136.0 (C-6), 135.6 (d, $^3J_{C,P}$=5.7 Hz, 1C of OCH_2Ph), 129.4 (Ar–C), 129.1 (Ar–C), 128.5 (Ar–C), 111.8 (C-5), 85.4 (C-3′), 85.3 (C-1′), 81.6 (d, $^3J_{C,P}$=8.3 Hz, C-4′), 70.5–70.4 (d, $^2J_{C,P}$=5.6 Hz, 2×OCH_2Ph), 68.1 (d, $^2J_{C,P}$=5.8 Hz, C-5′), 62.4 (C-f Bio), 60.6 (C-g Bio), 55.9 (C-e Bio), 40.6 (C-h, h′ Bio), 36.4 (C-2′), 32.8 (C-a Bio), 28.7 (C-c Bio), 28.5 (C-d Bio), 25.5 (C-b Bio), 12.4 (CH_3–Thy); ^{31}P NMR (121 MHz, $CDCl_3$) δ=−1.0; HRMS for $C_{34}H_{42}N_5O_{10}PS$ $[M+Na]^+$ calcd: 766.2282, found: 766.2302.

4.8. 3′-*O*-(Biotincarboxamide)-thymidine monophosphate triethylammonium salt (10)

Following a similar procedure as the one used for the synthesis of **7**, the triethylammonium salt of compound **10** was obtained as a white solid (169.0 mg, 82%), starting from **9** (200.0 mg, 0.269 mmol), Et_3N (0.075 mL, 0.538 mmol), and 10% Pd/C Degussa (100 mg, 50% w/w) in MeOH (20 mL). ^{1}H NMR (500 MHz, D_2O) δ=7.90 (s, 1H, H-6), 6.40 (dd, *J*=9.1, 5.7 Hz, 1H, H-1′), 4.73–4.72 (m, 1H, H-3′), 4.56–4.54 (m, 1H, H-g Bio), 4.38–4.35 (m, 1H, H-f Bio), 4.32 (br s, 1H, H-4′), 3.92–3.91 (m, 2H, H-5′ and H-5″), 3.32–3.28 (m, 1H, H-e Bio), 2.93 (dd, *J*=13.0, 4.9 Hz, 1H, H-h), 2.69 (d, 1H, *J*=13.0 Hz, H-h′ Bio), 2.48–2.44 (m, 1H, H-2′), 2.40–2.34 (m, 1H, H-2″), 2.15 (t, 2H, *J*=7.0 Hz, H-a Bio), 1.89 (s, 3H, C*H*$_3$), 1.72–1.67 (m, 1H, H-b Bio), 1.65–1.60 (m, 2H, H-d Bio), 1.59–1.51 (m, 1H, H-b′ Bio), 1.40–1.34 (m, 2H, H-c Bio); ^{13}C NMR (125 MHz, D_2O) δ=172.8 (ONHCO), 166.9 (C-4), 164.8 (NHCONH), 151.8 (C-2), 137.3 (C-6), 111.4 (C-5), 84.9 (C-3′), 84.3 (C-1′), 82.1 (d, $^3J_{C,P}$=8.9 Hz, C-4′), 63.6 (d, $^2J_{C,P}$=3.9 Hz, C-5′), 61.4 (C-f Bio), 59.7 (C-g Bio), 54.7 (C-e Bio), 39.2 (C-h, h′ Bio), 34.8 (C-2′), 31.8 (C-a Bio), 27.3 (C-c Bio), 27.0 (C-d Bio), 24.5 (C-b Bio), 11.3 (CH_3–Thy); ^{31}P NMR (202 MHz, D_2O) δ=3.4; HRMS for $C_{20}H_{30}N_5O_{10}PS$ $[M-H]^-$ calcd: 562.1378, found: 562.1382.

4.9. 5′-*O*-(Dibenzylphosphate)-3′-*O*-(biotinmethyloxycarboxamide)-thymidine (12)

Following a similar procedure as the one used for the synthesis of **9**, compound **12** was obtained as a colorless foam (214.0 mg, 76%), starting from **11** (200 mg, 0.365 mmol), biotin (102.6 mg, 0.420 mmol), DCC (120.6 mg, 0.584 mmol), and DMAP (1.8 mg, 0.0146 mmol) in a mixture of dry DMF (2.0 mL) and dry CH_2Cl_2 (10 mL). ^{1}H NMR (300 MHz, DMSO-d_6) δ=11.34 (br s, 1H, N*H*), 11.00 (br s, 1H, N*H*), 7.50 (s, 1H, H-6), 7.35–7.35 (m, 10H, Ar–*H* OBn), 6.41 (br s, 1H, N*H* Bio), 6.35 (br s, 1H, N*H* Bio), 6.15 (dd, *J*=8.2, 5.9 Hz, 1H, H-1′), 5.06 (d, *J*=8.0 Hz, 4H, 2×O*CH*$_2$Ph), 4.87–4.81 (m, 2H, O*CH*$_2$O), 4.51–4.50 (m, 1H, H-3′), 4.33–4.29 (m, 3H, H-g Bio, H-f Bio and H-4′), 4.19–4.10 (m, 2H, H-5′ and H-5″), 3.10–3.04 (m, 1H, H-e Bio), 2.81 (dd, *J*=12.6, 5.2 Hz, 1H, H-h Bio), 2.57 (d, *J*=12.6 Hz, 1H, H-h′ Bio), 2.35–2.28 (m, 1H, H-2′), 2.17–2.08 (m, 1H, H-2″), 2.00 (t, *J*=7.0 Hz, 3H, H-a Bio), 1.69 (s, 3H, C*H*$_3$), 1.63–1.42 (m, 4H, H-b and H-d Bio), 1.36–1.27 (m, 2H, H-c Bio); ^{13}C NMR (75 MHz, DMSO-d_6) δ=169.6 (ONHCO), 163.6 (C-4), 162.7 (NHCONH), 150.4 (C-2), 135.9 (d, $^3J_{C,P}$=6.9 Hz, 1C of OCH_2Ph), 135.6 (C-6), 128.5 (Ar–C), 128.4 (Ar–C), 127.8 (Ar–C), 109.9 (C-5), 97.0 (OCH_2O), 84.0 (C-1′), 82.3 (d, $^3J_{C,P}$=7.6 Hz, C-4′), 77.1 (C-3′), 68.7 (d, $^2J_{C,P}$=5.4 Hz, 2×OCH_2Ph), 67.1 (d, $^2J_{C,P}$=5.0 Hz, C-5′), 61.0 (C-f Bio), 59.2 (C-g Bio), 55.3 (C-e Bio), 39.5 (C-h, h′ Bio), 36.2 (C-2′), 32.0 (C-a Bio), 28.1 (C-c Bio), 28.0 (C-d Bio), 24.9 (C-b Bio), 12.0 (CH_3–Thy); ^{31}P NMR (121 MHz, $CDCl_3$) δ=−0.9; HRMS for $C_{35}H_{44}N_5O_{11}PS$ $[M+Na]^+$ calcd: 796.2388, found: 796.2403.

4.10. 3′-*O*-(Biotinmethyloxycarboxamide)-thymidine monophosphate triethylammonium salt (13)

Following a similar procedure as the one used for the synthesis of **7**, the triethylammonium salt of compound **13** was obtained as a white solid (170.6 mg, 79%), starting from **12** (210 mg, 0.271 mmol), Et_3N (0.075 mL, 0.543 mmol), and 10% Pd/C Degussa (105 mg, 50% w/w) in MeOH (20 mL). ^{1}H NMR (500 MHz, D_2O) δ=7.85 (d, *J*=1.1 Hz, 1H, H-6), 6.31 (app t, *J*=7.3 Hz, 1H, H-1′), 4.98 (s, 2H, O*CH*$_2$O), 4.71–4.70 (m, 1H, H-3′), 4.57–4.54 (m, 1H, H-g Bio), 4.39–4.37 (m, 1H, H-f Bio), 4.28–4.27 (m, 1H, H-4′),

3.97–3.90 (m, 2H, H-5′ and H-5″), 3.31–3.27 (m, 1H, H-e Bio), 2.94 (dd, J=13.0, 5.0 Hz, 1H, H-h), 2.71 (d, 1H, J=13.0 Hz, H-h′ Bio), 2.42–2.39 (m, 1H, H-2′ and H-2″), 2.20 (t, 2H, J=7.0 Hz, H-a Bio), 1.89 (d, J=1.1 Hz, 3H, CH_3), 1.70–1.54 (m, 1H, H-b and H-d Bio), 1.39–1.37 (m, 2H, H-c Bio); ^{13}C NMR (125 MHz, D_2O) δ=172.9 (ONHCO), 166.2 (C-4), 164.8 (NHCONH), 151.3 (C-2), 137.2 (C-6), 111.4 (C-5), 96.8 (OCH$_2$O), 84.3 (C-1′), 83.8 (d, $^3J_{C,P}$=8.7 Hz, C-4′), 78.4 (C-3′), 63.4 (d, $^2J_{C,P}$=3.7 Hz, C-5′), 61.5 (C-f Bio), 59.7 (C-g Bio), 54.8 (C-e Bio), 39.2 (C-h, h′ Bio), 36.1 (C-2′), 31.5 (C-a Bio), 27.2 (C-c Bio), 27.1 (C-d Bio), 24.2 (C-b Bio), 11.2 (CH$_3$–Thy); ^{31}P NMR (202 MHz, D_2O) δ=2.8; HRMS for $C_{21}H_{32}N_5O_{11}PS$ [M−H]$^-$ calcd: 592.1484, found: 592.1480.

4.11. 5′-*O*-(Dibenzylphosphate)-3′-*O*-(biotinester)-thymidine (14)

Following a similar procedure as the one used for the synthesis of **9**, compound **14** was obtained (370.0 mg, 74%) as a colorless foam, starting from **5** (345 mg, 0.686 mmol), biotin (201.0 mg, 0.823 mmol), DCC (198.0 mg, 0.960 mmol), and DMAP (8.0 mg, 0.068 mmol) in a mixture of dry DMF (3 mL) and dry CH_2Cl_2 (15 mL). ^{1}H NMR (300 MHz, $CDCl_3$) δ=10.95 (br s, 1H, N*H* Thy), 7.40 (s, 1H, H-6), 7.34–7.33 (m, 10H, Ar–*H* OCH$_2$Ph), 6.76 (br s, 1H, N*H* Bio), 6.50 (br s, 1H, N*H* Bio), 6.26 (app t, J=6.7 Hz, 1H, H-1′), 5.13–5.00 (m, 5H, H-3′ and OCH_2Ph), 4.51–4.50 (m, 1H, H-g Bio), 4.30–4.22 (m, 1H, H-f Bio, H-5′ and H-5″), 4.11 (br s, 1H, H-4′), 3.17–3.15 (m, 1H, H-e Bio), 2.91–2.88 (m, 1H, H-h Bio), 2.78–2.74 (m, 1H, H-h′ Bio), 2.35–2.25 (m, 3H, H-a Bio and H-2′), 1.97–1.94 (m, 1H, H-2″), 1.81 (s, 3H, CH_3), 1.69–1.60 (m, 4H, H-b and H-d Bio), 1.46–1.44 (m, 2H, H-c Bio); ^{13}C NMR (75 MHz, $CDCl_3$) δ=172.7 (OCO), 164.3 (C-4), 164.2 (NHCONH), 150.5 (C-2), 135.1 (d, $^3J_{C,P}$=6.1 Hz, 1C of OCH$_2$Ph), 134.7 (C-6), 128.5 (Ar–C), 128.3 (Ar–C), 127.7 (Ar–C), 111.2 (C-5), 84.1 (C-1′), 82.4 (d, $^3J_{C,P}$=8.1 Hz, C-4′), 74.1 (C-3′), 69.4 (app t, $^2J_{C,P}$=4.7 Hz, 2×OCH$_2$Ph), 66.8 (d, $^2J_{C,P}$=4.7 Hz, C-5′), 61.7 (C-f Bio), 60.0 (C-g Bio), 55.2 (C-e Bio), 40.2 (C-h, h′ Bio), 36.8 (C-2′), 33.4 (C-a Bio), 28.0 (C-c Bio), 27.8 (C-d Bio), 24.3 (C-b Bio), 12.0 (CH$_3$–Thy); ^{31}P NMR (121 MHz, $CDCl_3$) δ=−0.7; HRMS for $C_{34}H_{41}N_4O_{10}PS$ [M+H]$^+$ calcd: 729.2353, found: 729.2354.

4.12. 3′-*O*-(Biotinester)-thymidine monophosphate triethylammonium salt (15)

Following a similar procedure as the one used for the synthesis of **7**, the triethylammonium salt of compound **15** was obtained as a white solid (335.0 mg, 88%), starting from **14** (370.0 mg, 0.686 mmol), Et_3N (0.141 mL, 1.015 mmol), and 10% Pd/C (Degussa, 185.0 mg, 50% w/w) in MeOH (30 mL). ^{1}H NMR (600 MHz, DMSO-d_6) δ=11.34 (br s, 1H, N*H* Thy), 7.91 (s, 1H, H-6), 6.51 (br s, 1H, N*H* Bio), 6.40 (br s, 1H, N*H* Bio), 6.23 (dd, J=9.0, 5.8 Hz, 1H, H-1′), 5.28–5.27 (m, 1H, H-3′), 4.32–4.30 (m, 1H, H-g Bio), 4.15–4.12 (m, 1H, H-f Bio), 4.07 (br s, 1H, H-4′), 3.90–3.83 (m, 2H, H-5′ and H-5″), 3.13–3.10 (m, 1H, H-e Bio), 2.83–2.82 (m, 1H, H-h Bio merged with triethylamine), 2.58 (d, 1H, J=12.4 Hz, H-h′ Bio), 2.36 (t, 2H, J=7.1 Hz, H-a Bio), 2.33–2.29 (m, 1H, H-2′), 2.21–2.17 (m, 1H, H-2″), 1.81 (s, 3H, CH_3), 1.66–1.62 (m, 1H, H-b Bio), 1.61–1.55 (m, 2H, H-d Bio), 1.51–1.44 (m, 1H, H-b′ Bio), 1.39–1.32 (m, 2H, H-c Bio); ^{13}C NMR (150 MHz, DMSO-d_6) δ=172.6 (OCO), 163.9 (C-4), 162.9 (NHCONH), 150.7 (C-2), 136.2 (C-6), 110.3 (C-5), 83.7 (C-1′), 83.5 (d, $^3J_{C,P}$=7.4 Hz, C-4′), 75.5 (C-3′), 64.2 (d, $^2J_{C,P}$=6.0 Hz, C-5′), 61.1 (C-f Bio), 59.3 (C-g Bio), 55.4 (C-e Bio), 39.6 (C-h,h′ Bio merged with DMSO), 36.5 (C-2′), 33.4 (C-a Bio), 28.0 (C-c and C-d Bio), 24.5 (C-b Bio), 12.2 (CH$_3$–Thy); ^{31}P NMR (202 MHz, DMSO-d_6) δ=−0.1; HRMS for $C_{20}H_{29}N_4O_{10}PS$ [M−H]$^-$ calcd: 547.1269, found: 547.1279.

4.13. *N*-Boc kanamycin A (17)

To a stirred solution of kanamycin A **16** (1.8 g, 3.71 mmol) in a mixture of DMF/H_2O (4:1, 48 mL), was added di-*tert*-butyldicarbonate (5.12 mL, 22.29 mmol) and the solution was heated at 60 °C for 5 h and then cooled to room temperature. After removal of all the volatiles, the resulting residue was washed with hexane (2×100 mL) and then with water (3×100 mL), and finally dried in vacuo to afford **17** (2.99 g, 91%) as a white solid. R_f=0.4 (10% MeOH in DCM). ^{1}H NMR (300 MHz, DMSO-d_6) δ=6.89 (br s, 1H), 6.58 (br s, 1H), 6.50 (d, J=8.6 Hz, 1H), 6.34 (br s, 1H), 5.35 (d, J=3.8 Hz, 1H), 5.26 (s, 1H), 4.89 (m, 4H), 4.68 (d, J=6.2 Hz, 1H), 4.19 (t, J=5.4 Hz, 1H), 3.82 (m, 1H), 3.61–3.35 (m, 9H), 3.30–3.17 (m, 6H), 3.07 (m, 1H), 1.80 (m, 1H), 1.40–1.35 (m, 37H); ^{13}C NMR (75 MHz, DMSO-d_6) δ=156.3, 156.1, 155.3, 154.9, 101.1, 97.8, 84.0, 80.3, 77.8, 77.7, 77.2, 75.0, 72.9, 72.7, 72.1, 70.5, 70.3, 70.1, 67.4, 60.3, 55.9, 50.0, 49.1, 41.4, 34.7, 28.3, 28.2, 28.1; HRMS [ESI$^+$] for $C_{38}H_{68}N_4O_{19}$ [M+H]$^+$ calcd: 885.4550, found: 885.4552.

4.14. *O*-TIBS-*N*-Boc kanamycin A (18)

To a stirred solution of **17** (2.48 g, 2.81 mmol) in dry pyridine (50 mL) was added 2,4,6-triisopropylbenzenesulfonyl chloride (5.96 g, 19.67 mmol) and the solution was stirred at 25 °C for 16 h. The reaction mixture was neutralized by adding 1 N HCl and diluted with water. The aqueous layer was extracted with ethyl acetate (3×400 mL). The collected organic fractions were washed with brine, dried over Na_2SO_4, and concentrated in vacuo. The crude residue was purified by column chromatography over silica gel (3% MeOH in DCM) to afford **18** (1.91 g, 59%) as a white solid. R_f=0.48 (10% MeOH in DCM). ^{1}H NMR (300 MHz, CD_3OD) δ=7.28 (s, 1H), 6.58 (br s, 1H), 5.35 (d, J=3.8 Hz, 1H), 5.04 (m, 2H), 4.38 (m, 2H), 4.14 (m, 3H), 3.71–3.31 (m, 12H), 3.18 (t, J=9.4 Hz, 1H), 2.95 (m, 1H), 2.02 (m, 1H), 1.45–1.41 (m, 37H), 1.28–1.25 (m, 18H); ^{13}C NMR (75 MHz, CD_3OD) δ=159.4, 159.2, 157.9, 157.7, 155.3, 152.3, 130.7, 124.9, 102.8, 99.8, 85.8, 80.9, 80.6, 80.4, 80.2, 76.8, 74.5, 73.9, 72.3, 71.9, 71.7, 71.6, 69.2, 57.4, 52.2, 50.9, 41.9, 36.0, 35.5, 30.8, 28.9, 28.8, 28.7, 25.2, 25.1, 23.9; HRMS [ESI$^+$] for $C_{53}H_{90}N_4O_{21}S$ [M+H]$^+$ calcd: 1151.5890, found: 1151.5895.

4.15. *N*-Boc kanamycin A cysteamine (19)

To a stirred solution of **18** (0.9 g, 0.782 mmol) in DMF (30 mL) was added 2-mercaptoethylamine hydrochloride (0.67 g, 5.86 mmol) followed by cesium carbonate (2.55 g, 7.82 mmol) and the solution was stirred at 25 °C for 16 h. DMF was partly removed in vacuo and then partitioned between water (150 mL) and ethyl acetate (300 mL). The aqueous layer was extracted with ethyl acetate (2×150 mL). The collected organic fractions were washed with brine, dried over Na_2SO_4, and concentrated in vacuo. The crude residue was purified by column chromatography over silica gel (5% MeOH in DCM) to afford **19** (0.45 g, 61%) as an off-white solid. R_f=0.38 (10% MeOH in DCM). ^{1}H NMR (300 MHz, CD_3OD) δ=5.13 (d, J=3.5 Hz, 1H), 5.08 (d, J=3.3 Hz, 1H), 4.17 (t, J=6.7 Hz, 1H), 3.72–3.52 (m, 6H), 3.47–3.38 (m, 7H), 3.21 (m, 3H), 3.05 (m, 1H), 2.90–2.81 (m, 3H), 2.68 (dd, J=14.5, 6.8 Hz, 1H), 2.12 (m, 1H), 1.46–1.43 (m, 37H); ^{13}C NMR (75 MHz, CD_3OD) δ=159.4, 159.2, 158.1, 157.7, 102.6, 100.6, 85.2, 83.2, 80.6, 80.4, 80.2, 71.2, 74.5, 74.0, 73.8, 72.4, 72.1, 72.0, 71.8, 57.0, 52.0, 50.8, 41.8, 40.0, 35.9, 34.7, 31.4, 30.4, 28.9, 28.8, 28.8, 28.7, 24.2; HRMS [ESI$^+$] for $C_{40}H_{73}N_5O_{18}S$ [M+H]$^+$ calcd: 944.4744, found: 944.4733.

4.16. 5′-*O*-(Dibenzylphosphate)-3′-*O*-(*N*-Boc kanamycin A cysteamine carbamate)-thymidine (21)

To a stirred solution of compound **19** (353 mg, 0.374 mmol) in 1,4-dioxane/water (3:1, 8 mL) was added **20** (250 mg, 0.374 mmol)

followed by triethylamine (0.052 mL, 0.374 mmol) at 0 °C and then the solution was stirred at room temperature for 16 h. After completion of the reaction, the mixture was evaporated to dryness in vacuo and the residue was purified by column chromatography (4% MeOH in DCM) to obtain compound **21** (495 mg, 90%) as a white solid. R_f=0.45 (10% MeOH in DCM). ^{1}H NMR (500 MHz, CD_3OD) δ=7.47 (s, 1H), 7.36 (s, 10H), 6.25 (dd, J=8.6, 5.8 Hz, 1H), 5.14–5.07 (m, 6H), 5.05 (br s, 1H), 4.29 (m, 2H), 4.20 (m, 2H), 3.71–3.60 (m, 4H), 3.55–3.48 (m, 4H), 3.42–3.32 (m, 6H), 3.19 (t, J=9.3 Hz, 1H), 3.02 (dd, J=13.9, 2.1 Hz, 1H), 2.75–2.62 (m, 3H), 2.33 (dd, J=14.2, 5.5 Hz, 1H), 2.07 (m, 1H), 1.76 (s, 3H), 1.44–1.42 (m, 37H); ^{13}C NMR (125 MHz, CD_3OD) δ=166.18, 159.5, 159.2, 158.0, 157.7, 152.3, 137.2, 137.1 (d, $^3J_{C,P}$=2.3 Hz), 137.0 (d, $^3J_{C,P}$=2.3 Hz), 130.0, 129.8, 129.3, 112.2, 102.7, 100.2, 86.3, 84.4 (d, $^3J_{C,P}$=7.8 Hz), 80.6, 80.5, 80.4, 80.2, 77.2, 76.0, 74.7, 74.1, 73.7, 72.5, 72.4, 72.3, 71.8, 71.2 (2×d, $^2J_{C,P}$=5.8 Hz), 68.8 (d, $^2J_{C,P}$=6.1 Hz), 57.2, 50.9, 41.7, 38.2, 35.0, 33.9, 28.9, 28.9, 28.9, 28.8, 28.8, 12.6; ^{31}P NMR (202 MHz, CD_3OD) δ=−0.03; HRMS [ESI$^+$] for $C_{65}H_{98}N_7O_{27}PS$ [M+Na]$^+$ calcd: 1494.5861, found: 1494.5873.

4.17. 3′-*O*-(Kanamycin A cysteamine carbamate)–TMP (22)

To a stirred solution of **21** (200 mg, 0.136 mmol, 1 equiv) in EtOH/H_2O (8:2, 15 mL), was added 20% $Pd(OH)_2$/C (160 mg, 0.8 equiv w/w) and evacuation was then carried out with hydrogen atmosphere replacements (3×). The reaction mixture was stirred at room temperature for 24 h under an atmospheric pressure of hydrogen. After completion of the reaction, the catalyst was removed by filtration through a cellulose filter (0.45 μm) and the filtrate was concentrated under reduced pressure to obtain crude 3′-*O*-(*N*-Boc kanamycin A cysteamine carbamate)–TMP (∼175 mg, quant.). The so obtained residue was treated without further purification with 90% aq TFA (5 mL) at 0 °C and then the solution was stirred at room temperature for 5 h. After completion of the reaction, the mixture was evaporated to dryness in vacuo. The residue was coevaporated three times with toluene to remove residual TFA and to give crude 3′-*O*-(kanamycin A cysteamine carbamate)–TMP. The crude residue was purified by preparative RP-HPLC (0.1% TFA in 98% H_2O+2% ACN and 0.1% TFA in 98% ACN+2% H_2O). The collected eluates were freeze-dried repeatedly until constant mass to afford compound **21** (86 mg, 71% over two steps) as trifluoroacetic salt. ^{31}P NMR (202 MHz, D_2O) δ=−0.02; HRMS [ESI$^-$] for $C_{31}H_{54}N_7O_{19}PS$ [M−H]$^-$ calcd: 890.2860, found: 890.2852.

4.18. 5′-*O*-(Dibenzylphosphate)-3′-*O*-(kanamycin A carbamate)-thymidine (23)

Compound **23** was obtained according to the procedure for the preparation of compound **21** starting from kanamycin A **16** (100 mg, 0.206 mmol), compound **20** (138 mg, 0.206 mmol), and triethylamine (29 μL, 0.206 mmol). The crude residue was purified by column chromatography (IPA/H_2O/Et_3N 10:2:1) to give **23** as a white solid (171 mg, 82%). R_f=0.45 (IPA/H_2O/Et_3N 6:2:2). ^{1}H NMR (300 MHz, D_2O) δ=7.16 (s, 1H), 7.09 (s, 10H), 6.1 (app t, J=6.0 Hz, 1H), 5.21 (d, J=5.5 Hz, 1H), 5.01 (m, 2H), 4.84 (m, 4H), 4.19–3.89 (m, 4H), 3.81–3.08 (m, 14H), 2.99 (t, J=10.4 Hz, 1H), 2.82 (t, J=11.3 Hz, 1H), 2.31 (m, 1H), 2.04–1.80 (m, 2H), 1.52 (s, 3H), 1.33–1.19 (m, 1H); ^{13}C NMR (75 MHz, D_2O) δ=165.3, 156.8, 151.0, 135.7, 134.8 (d, $^3J_{C,P}$=5.3 Hz), 128.5, 128.3, 127.7, 127.6, 127.1, 110.8, 99.9, 99.7, 86.8, 86.1, 84.6, 82.2, 74.2, 73.7, 72.4, 72.0, 71.5, 71.2, 70.6, 70.4, 69.6, 67.8, 67.0, 59.8, 54.2, 50.0, 48.9, 41.0, 36.2, 33.5, 11.5; ^{31}P NMR (121 MHz, D_2O) δ=−0.89; HRMS [ESI$^+$] for $C_{43}H_{61}N_6O_{20}P$ [M+H]$^+$ calcd: 1013.3751, found: 1013.3745.

4.19. 3′-*O*-(Kanamycin A carbamate)–TMP (24)

Compound **24** was obtained as a white solid (86.7 mg, 88%) starting from compound **23** (120 mg, 0.118 mmol) following the same hydrogenation procedure described for compound **22**, using 20% $Pd(OH)_2$/C (24 mg, 0.2 equiv w/w). ^{31}P NMR (202 MHz, D_2O) δ=−0.02; HRMS [ESI$^-$] for $C_{29}H_{49}N_6O_{20}P$ [M−H]$^-$ calcd: 831.2666, found: 831.2666.

Acknowledgements

We would like to thank Prof. Abram Aertsen and Anirban Ghosh for kindly providing the bacterial strain, Prof. Jef Rozenski for performing LC/MS analysis and Lia Margamuljana for helping with the biological test. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement n° ERC-2012-ADG_20120216/320683.

Supplementary data

Supplementary data associated with this article can be found in the online version, at http://dx.doi.org/10.1016/j.tet.2014.09.096.

References and notes

1. Benner, S. A.; Sismour, A. M. *Nat. Rev. Genet.* **2005**, *6*, 533–543.
2. Wu, Y.; Fa, M.; Tae, E. L.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2002**, *124*, 14626–14630.
3. Malyshev, D. A.; Dhami, K.; Lavergne, T.; Chen, T.; Dai, N.; Foster, J. M.; Correa, I. R.; Romesberg, F. E. *Nature* **2014**, *509*, 385–388.
4. De, S.; Groaz, E.; Herdewijn, P. *Eur. J. Org. Chem.* **2014**, *2014*, 2322–2348.
5. Pezo, V.; Liu, F. W.; Abramov, M.; Froeyen, M.; Herdewijn, P.; Marlière, P. *Angew. Chem., Int. Ed.* **2013**, *52*, 8139–8143.
6. Prakash, O.; Eisenberg, M. A. *J. Bacteriol.* **1974**, *120*, 785–791.
7. Piffeteau, A.; Zamboni, M.; Gaudry, M. *Biochim. Biophys. Acta* **1982**, *688*, 29–36.
8. Davis, B. D. *Microbiol. Rev.* **1987**, *51*, 341–350.
9. Magnet, S.; Blanchard, J. S. *Chem. Rev.* **2005**, *105*, 477–498.
10. Taber, H. W.; Mueller, J. P.; Miller, P. F.; Arrow, A. S. *Microbiol. Rev.* **1987**, *51*, 439–457.
11. Becker, B.; Cooper, M. A. *ACS Chem. Biol.* **2013**, *8*, 105–115.
12. Semenyuk, A.; Földesi, A.; Johansson, T.; Estmer-Nilsson, C.; Blomgren, P.; Brännvall, M.; Kirsebom, L. A.; Kwiatkowski, M. *J. Am. Chem. Soc.* **2006**, *128*, 12356–12357.
13. Sekine, M.; Iimura, S.; Nakanishi, T. *Tetrahedron Lett.* **1991**, *32*, 395–398.
14. Manoharan, M.; Lu, Y.; Casper, M. D.; Just, G. *Org. Lett.* **2000**, *2*, 243–246.
15. Luedtke, N. W.; Carmichael, P.; Tor, Y. *J. Am. Chem. Soc.* **2003**, *125*, 12374–12375.
16. Charles, I.; Xue, L.; Arya, D. P. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1259–1262.
17. Michael, K.; Wang, H.; Tor, Y. *Bioorg. Med. Chem.* **1999**, *7*, 1361–1371.
18. Hirao, I.; Itoh, K.; Nishino, S.; Araki, Y.; Ishido, Y. *Carbohydr. Res.* **1986**, *152*, 159–172.
19. Suzuki, M.; Sekido, T.; Matsuoka, S.-i.; Takagi, K. *Biomacromolecules* **2011**, *12*, 1449–1459.
20. Liu, Q.; Luedtke, N. W.; Tor, Y. *Tetrahedron Lett.* **2001**, *42*, 1445–1447.
21. Schepdael, V. A.; Busson, R.; Claes, P. *Bull. Soc. Chim. Belg.* **1992**, *101*, 709–718.
22. Sainlos, M.; Belmont, P.; Vigneron, J.-P.; Lehn, P.; Lehn, J.-M. *Eur. J. Org. Chem.* **2003**, *2003*, 2764–2774.
23. Blount, K. F.; Zhao, F.; Hermann, T.; Tor, Y. *J. Am. Chem. Soc.* **2005**, *127*, 9818–9829.