ChemPubSoc Europe

DOI: 10.1002/chem.201600721

CHEMISTRY A European Journal Full Paper

Synthetic Biology |*Hot Paper*|

Syntheses of 5′-Nucleoside Monophosphate Derivatives with Unique Aminal, Hemiaminal, and Hemithioaminal Functionalities: A New Class of 5′-Peptidyl Nucleotides

Swarup De, Elisabetta Groaz, Lia Margamuljana, and Piet Herdewijn*[a]

Abstract: A number of synthetically useful transformations have been developed to generate novel 5′-peptidyl nucleoside monophosphate analogues that incorporate sensitive phosphoaminal, -hemiaminal or -hemithioaminal functionalities. The strategies adopted entailed the coupling between dipeptides, which enclose a reactive Cα-functionalized glycine residue and phosphate or phosphorothioate moieties. These developments led to potentially powerful and general methodologies for the preparation of α-phosphorylated pseudopeptides as well as nucleoside monophosphate mimics. The resulting conjugates are of interest for a variety of important applications, which range from drug development to synthetic biology, as pronucleotides or artificial building blocks for the enzymatic synthesis of xenobiotic information systems. The potential of all dipeptide-TMP conjugates as pyrophosphate mimics in the DNA polymerization reaction was tested, and the influence of the nature of the linker was evaluated by in vitro chain elongation assay in the presence of wild-type microbial DNA polymerases.

Introduction

Amino acids are a common, non-toxic structural motif in numerous prodrug compounds, some of which have shown interesting biological activities.[1] For example, peptidomimetic conjugates were designed to increase the solubility of polar nucleoside analogues (i.e., acyclovir, ganciclovir, didanosine), which enhances their cellular permeability and bioavailability following oral administration, by targeting intestinal oligopeptide transporters.[2] The phosphoramidate family of nucleotide prodrugs contains an amino acid ester linked to the 5′-phosphate group through a P–N bond, which is a key characteristic element.[3] This species is sufficiently stable during transport, but the nucleotide adduct is readily cleaved by intracellular enzymes with phosphoramidase activity.[4] As a result, this process releases the promoiety as a biodegradable intermediate, which then re-enters the cellular metabolism and concomitantly leads to the delivery of the parent active nucleoside monophosphate.

We have previously reported that analogous amidate monoester derivatives that comprise amino acids[5] (i.e., L-Asp, L-His) as well as small peptides[6] represent synthetic alternatives to natural nucleoside triphosphates and act as direct substrates in the DNA-polymerase-catalyzed synthesis of DNA. The recognition of the exceptional role that is displayed by this class of compounds is relevant to both the fields of medicinal chemistry and synthetic biology, which allows for the design of new drugs that are able to bypass the multistep anabolic cascade, which is required for nucleoside activation,[7] or unnaturally activated building blocks for the biosynthesis of artificial genetic polymers (XNAs).[8]

A surprising discovery was made by Richert and co-workers, who recently showed that it is possible to form peptidyl chains that are linked to RNA through a 5′-phosphoramidic bond from the spontaneous oligomerization between ribonucleoside phosphates and amino acids under plausible prebiotic conditions and in the absence of cellular machinery.[9] The resulting peptidyl RNA, which is produced alongside simple genetic material, thus emerged as an early protein precursor in primitive metabolic pathways.

However, from a synthetic perspective, the *N*-terminus of an amino acid/peptide chain is the most common site for conjugation to the 5′-phosphate moiety of NMPs.[4a,6,10] Aminoacyl adenylates, which can be considered as molecular isomers of amino acid phosphoramidate nucleosides, are important activated intermediates in protein biosynthesis.[11] The connecting phosphoric–carboxylic mixed anhydride bond makes those molecules too unstable to be synthetically useful.[12] Naturally occurring[13] and relatively less reactive peptidyl nucleotide analogues that contain active carboxyl groups were synthesized.[14]

We wished to explore whether the reaction between nucleoside monophosphates and the alpha carbon of a suitable functionalized amino acid residue would allow us to produce novel, synthetically and biologically useful molecules. To this end, we designed a range of L-Ala-Gly nucleotide conjugates

[a] *Dr. S. De, Dr. E. Groaz, L. Margamuljana, Prof. Dr. P. Herdewijn*
KU Leuven
Rega Institute for Medical Research, Medicinal Chemistry
Minderbroedersstraat 10, 3000 Leuven (Belgium)
E-mail: Piet.Herdewijn@rega.kuleuven.be

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/chem.201600721.

© 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

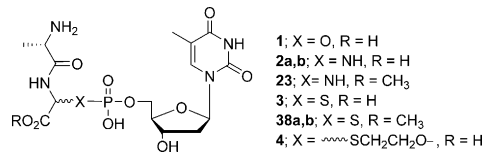


Figure 1. General design of molecular targets: Cα-X-peptide–nucleotide conjugates (X = O, N, S).

that were connected through a phosphodiester or a mixed amido/thioate-phosphoester bond, as illustrated in Figure 1.

The formation of α-heterosubstituted peptides from electrophilic precursors typically requires the presence of a good leaving group at the α-carbon center of a glycine residue, which is commonly in the form of an alkoxide,[15] a carboxylate,[16] or a halide group.[17] Although α-substituted glycines can rapidly decompose by elimination of the α-group, their stabilization can be accomplished by masking the lone pair of the glycyl nitrogen through N-acylation (i.e., formation of a peptide bond).[16a,18] Small heteroatom-centered (i.e., N, O, and S) alkyl and aryl nucleophiles were shown to be suitable reaction partners.[16c,19] Most notably, glycine-containing peptides that were modified at the α-carbon with a toxophoric agent[16a] or a phenyl derivative[20] were found not only to be accepted and actively transported by peptide permeases into bacterial cells but also to undergo hydrolysis in the presence of intracellular peptidases, which thus liberates the attached cargo.

To the best of our knowledge, the synthetic utility of this transformation has not been demonstrated for highly charged compounds. It was anticipated that efficient N–Cα–X–P bond formation involved overcoming challenges that were associated not only with the inherent chemical instability of the parent peptides but also with the increased nucleophilicity of the unit that originates from the phosphorus-containing nucleophile in the final products.

Results and Discussion

Synthetic routes need to be cautiously designed and executed by using suitable orthogonal protection/deprotection methods as well as reaction conditions that are compatible with the sensitive phosphodiester and mixed phosphoester units of newly designed conjugates. Our retrosynthetic analysis for the proposed target compounds is shown in Figure 2. It became apparent that our target conjugates could be obtained by nucleophilic substitution either at the phosphorus group itself (Route 1) or at the α-position of glycine (Route 2).

Among the various options that are available for the formation of a phosphodiester linkage, we selected the approach that is based on the use of a dialkylphosphoramidite intermediate followed by in situ oxidation in our initial efforts to access phosphohemiaminal diester target **1** (Scheme 1). α-Hydroxy glycine derivative **5** was obtained as a diastereoisomeric mixture in good yield by coupling *z*-L-Ala–NH_2 with benzyl glyoxalate in accordance with an improved protocol,[16a] which comprised the addition of a catalytic amount of *p*-TSA to increase the reaction rate. As all attempts to generate bis(alkyl)-*N,N*-diisopropyl phosphoramidite derivative **6** met with failure, we turned to an alternative strategy that was based on the conversion of 3′-*O*-benzyl thymidine **8** into the corresponding pro-moiety **9**[21] followed by phosphoramidite coupling with the α-OH of glycine **5** and in situ oxidation. Although thymidine phosphoramidite **9** could be successfully isolated, the subsequent coupling with glycine **5** in the presence of 1-*H* tet-

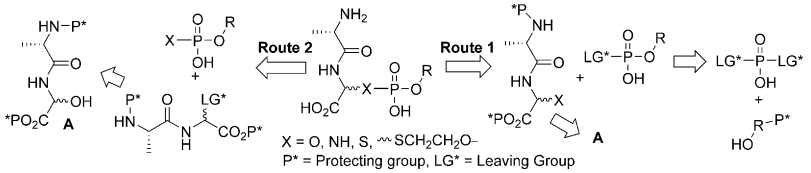
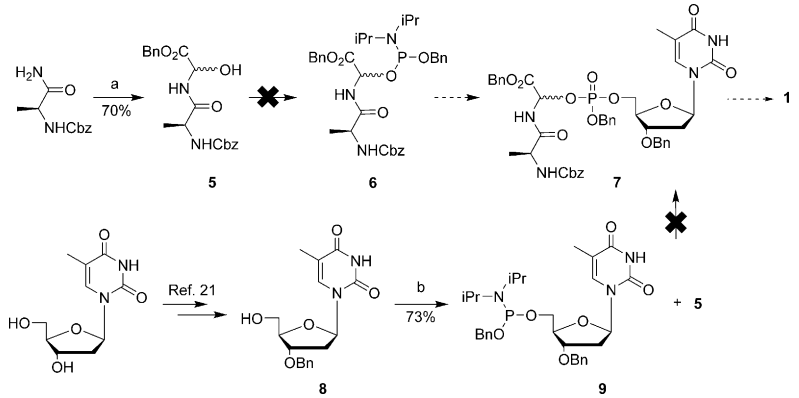


Figure 2. Retrosynthetic analysis of 5′-peptidyl nucleoside monophosphate targets.

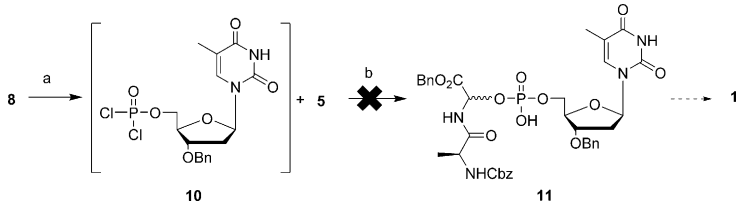


Scheme 1. Attempted routes towards the synthesis of the phosphoester of (L-alanyl-D,L-2-aminoglycine)-TMP **1**. Reagents and conditions: (a) Benzyl glyoxalate, *p*-TSA, CH_2Cl_2, 168 h; (b) Benzyloxy-bis(*N,N*-diisopropylphosphoramino) phosphine, 0.45 M 1*H*-tetrazole in CH_3CN, CH_2Cl_2, 0 °C to RT, 4 h.

© 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

razole and oxidation with H_2O_2 did not produce the desired compound **7**, but 3′-*O*-benzyl-(benzyloxy)-TMP was formed instead. Despite screening of various coupling agents with different pK_a values, which included 4,5-dicyanoimidazole, 2-ethylthio tetrazole, and 2-benzylthio tetrazole, none were found to be beneficial to the outcome of the above reaction.

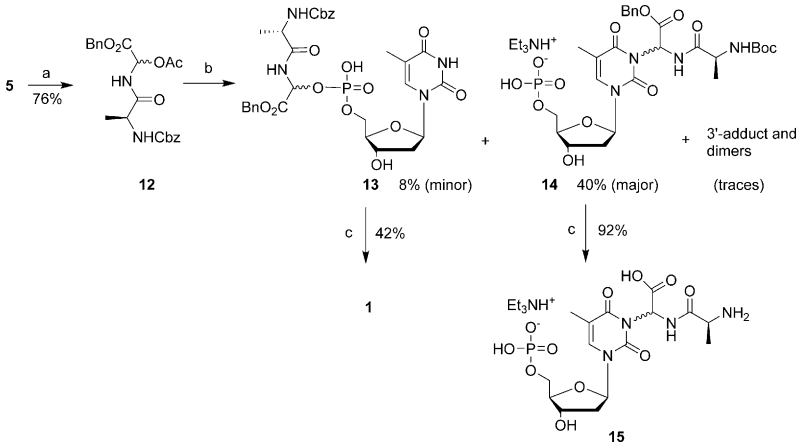
Therefore, we proceeded to investigate a method that was based on a more reactive chlorophosphate intermediate, as depicted in Scheme 2. 3′-*O*-benzyl thymidine **8** was converted into its activated form, 5′-*O*-dichlorophospho-3′-*O*-benzyl-thymidine **10**, by using $POCl_3$ in trimethyl phosphate, which was then treated in situ with glycine **5** in the presence of *N*-methyl imidazole. Surprisingly, the reaction did not afford the desired product, but it resulted in the formation of 3-*O*-benzyl-TMP along with other uncharacterized byproducts. A number of bases with different pK_a's, such as Et_3N, DIPEA, and pyridine, were employed but yielded the same results. This may be due to the low nucleophilicity of the α-OH group, a steric hindrance at the α-position, the instability of the mono-chlorophosphate intermediate during quenching of the reaction, or a base-promoted[22] elimination reaction at the α-position of the desired compound.



Scheme 2. Route towards the synthesis of the phosphoester of (L-alanyl-D,L-2-aminoglycine)-TMP **1**. Reagents and conditions: (a) $POCl_3$, trimethyl phosphate, −20 °C to RT, 8 h; (b) Compound **5**, base, CH_2Cl_2, 0 °C to RT, 16 h.

Owing to the perceived difficulties in achieving the synthesis of compound **1**, we next switched to Route 2 of the proposed retrosynthetic analysis (Figure 2). Accordingly, the α-hydroxy glycine derivative **5** was activated towards nucleophilic substitution at the α-position by conversion to its α-acetoxy derivative **12**, which was then reacted with the triethylammonium salt of TMP (Scheme 3). Although at first a single product was thought to have formed on the basis of mass spectroscopic analysis of the crude residue, further RP-HPLC purification revealed a mixture of products with the same mass. After in-depth two-dimensional NMR analysis of the separated products, it could be concluded that the desired phosphodiester derivative **13** was formed as a minor product (8% yield) along with thymidylate N^3-peptide adduct **14** as the major product (40% yield). Various dimeric products and a 3′-TMP-peptide adduct were also produced in trace amounts in the course of the reaction. This result can be rationalized on the basis of the hard soft acid base principle (HSAB): a soft center, such as an α-carbon, prefers a soft donor; in the present case, the electrons will preferentially be donated from the N^3-atom rather than from the oxygen atom of the phosphate group. Despite numerous attempts to improve the reaction yields, longer times and higher temperatures only led to the formation of additional undesired side products. As the polarity similarities between the 3′-TMP-peptide adduct and compound **13** did not allow their separation by preparative HPLC, the two regioisomers were separated in the subsequent step. Finally, removal of all protecting groups of compound **13** by reductive hydrogenation with 10% Pd/C Degussa-type followed by preparative HPLC purification yielded the desired conjugate **1** in moderate yield (42%). During the conversion of compound **13** into target **1**, we observed the formation of TMP as an impurity, whereas when the same reaction conditions were applied to compound **14**, we isolated the desired product **15** in excellent yield (92%).

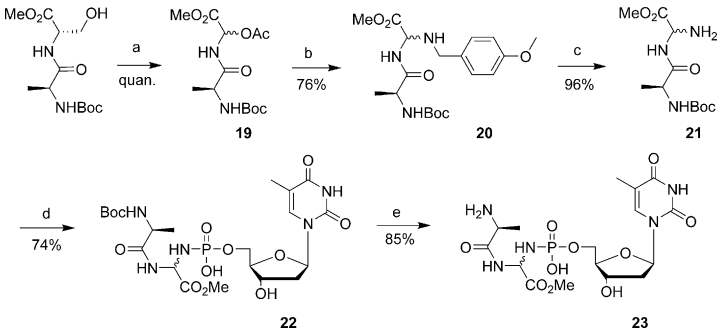
In principle, the synthesis of phosphoaminal analogue **2** could be accessed through either Route 1 or 2, as shown in the retrosynthetic analysis (Figure 2). Above, we reported the competition for nucleophilic attack in Route 1 between the N^3-atom of thymine and the thymidine-5′-*O*-phosphoramidate ni-



Scheme 3. Synthesis of the phosphoester of (L-alanyl-D,L-2-aminoglycine)-TMP **1**. Reagents and conditions: (a) Ac_2O, pyridine, 0 °C, 8 h; (b) TMP triethylammonium salt, Et_3N, DMF, 40 °C, 24 h; (c) 10% Pd/C (Degussa-type), EtOH/H_2O 5:1, RT, 4 h.

© 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

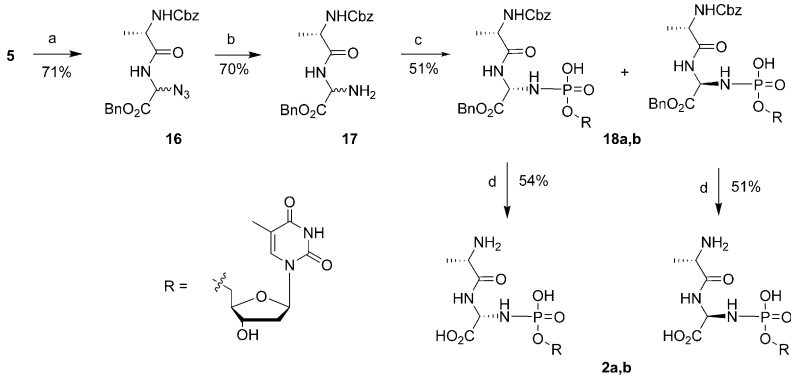
trogen atom; therefore, the more selective approach, Route 1, was employed. The synthesis of α-amino glycine derivative **17** from its corresponding hydroxyl derivative **5** was performed via an azide intermediate (Scheme 4). Azido compound **16** was obtained in good yield from glycine derivative **5** by using the DPPA/DBU method,[23] and it was then converted into its amino form **17** by classical reduction in the presence of PPh_3. The ensuing phosphoramidate coupling was carried out under standard DCC coupling conditions[24] to afford a diastereoisomeric mixture of compounds **18a** and **18b** in good yield. Our initial efforts to separate isomers **18a,b** after the final debenzylation step turned out to be problematic; however, the protected diastereoisomers could be successfully separated by preparative RP-HPLC to yield pure **18a** and **18b** in an approximate 1:1 ratio. Finally, catalytic hydrogenation effected the cleavage of the two protecting groups, which gave rise to the corresponding isomers **2a** and **2b** in moderate yields.



Scheme 5. Synthesis of the phosphoaminal of (L-alanyl-D,L-2-aminoglycine methyl ester)-TMP **23**. Reagents and conditions: (a) $Pb(OAc)_4$, dry EtOAc, 4 Å MS, reflux, 2 h; (b) 4-methoxybenzylamine, DIPEA, dry DMF, 40 °C, 24 h; (c) 10% Pd/C, EtOH, RT, 5 h (d) TMP triethylammonium salt, DCC, Et_3N, *t*BuOH/H_2O 4:1, 85 °C, 2.5 h; (e) TFA, thioanisole, CH_2Cl_2, 0 °C to RT, 4 h.

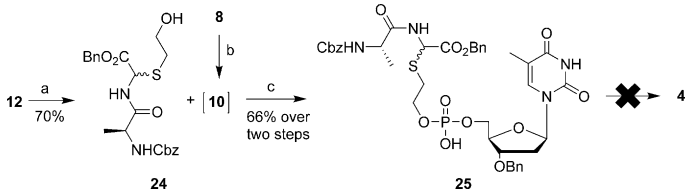
As it is not unreasonable to speculate that the free α-carboxyl group of glycine might play a role in the formation of TMP as a side product during the isolation of nucleotides **1**, **2a** and **2b**, we were interested to synthesize phosphoaminal ester analogue **23**, in which the carboxyl group of glycine is protected as a methyl ester. A suitable choice of orthogonal protecting groups was crucial for the synthesis of compound **23** (Scheme 5), and the main synthetic strategy resembled that which was employed for nucleotides **2ab**. Unlike benzyl-protected α-acetoxy glycine derivative **12**, which was synthesized in a convergent manner, the synthesis of α-acetoxy derivative **19** followed a more practical degradative approach, which started from *N*-Boc-L-Ala-L-Ser-OMe in the presence of $Pb(OAc)_4$,[22] because the related protecting groups were well tolerated during the reaction. The acetoxy derivative **19** was then converted to its protected amino derivative **20** by nucleophilic substitution in the presence of 4-methoxybenzyl amine. Subsequent deprotection by Pd/C-catalyzed hydrogenation yielded the desired amine **21** in excellent yield (96%). Phosphoaminal ester **22** was obtained in good yield by using the standard coupling protocol, and successive removal of the Boc protecting group with TFA in the presence of the radical scavenger thioanisole produced the desired methyl ester derivative **23**; in this case, no TMP formation was observed. As for previously described compounds **18a,b**, a late stage effort was made to separate the diastereoisomers of **22** and **23** by preparative RP-HPLC, but it was not found to be effective in this case.

Although both proposed retrosynthetic routes are feasible for the synthesis of thioethyl phosphate derivative **4**, it was anticipated that the instability of the thioethyl-TMP synthon could be problematic. As described in the literature, this unit can be readily synthesized,[25] but when it is formed it immediately decomposes through nucleophilic attack of the naked thiol group to give TMP and thiirane.[26] According to Route 2, thioethanol analogue **24** was formed by selective S-alkylation of the corresponding acetoxy compound **12** with 2-mercaptoethanol in good yield (70%), as shown in Scheme 6. The phosphoramidite approach could in theory be utilized for the synthesis of phosphoester **25**; however, selective oxidation of P^{III} to P^{V} in the presence of a sulfur atom posed a real challenge.[27] Therefore, the $POCl_3$ method was adopted instead, which involved the formation of intermediate **10** from 3′-*O*-benzyl thymidine and its in situ treatment with thioethanol analogue **24** in the presence of *N*-methyl imidazole. Quenching



Scheme 4. Synthesis of the phosphoaminals of (L-alanyl-D,L-2-aminoglycine)-TMP **2a** and **2b**. Reagents and conditions: (a) DPPA, DBU, dry toluene, 0 °C to RT, 16 h; (b) PPh_3, THF, reflux, 2 h; (c) TMP triethylammonium salt, Et_3N, DCC, *t*BuOH/H_2O 4:1, 85 °C, 2.5 h; (d) 10% Pd/C (Degussa-type), EtOH/H_2O 5:1, RT, 4 h.

© 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



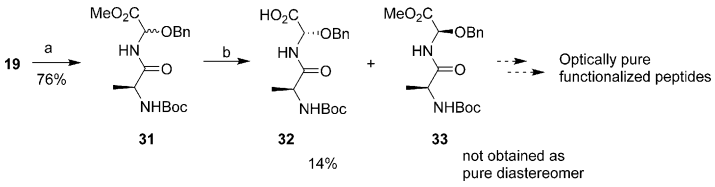
Scheme 6. Attempted route towards the synthesis of the phosphoester of (L-alanyl-D,L-2-mercaptoethanolglycine)-TMP **4**. Reagents and conditions: (a) 2-mercaptoethanol, Et_3N, dry DMF, 0 °C to RT,16 h; (b) $POCl_3$, trimethyl phosphate, −20 °C to RT, 8 h (c) *N*-methyl imidazole, CH_2Cl_2, 0 °C to RT,16 h.

followed by preparative RP-HPLC purification furnished phosphoester **25** as a diastereoisomeric mixture in 66% yield over two steps. For the conversion of protected derivative **25** into nucleotide **4**, we attempted to remove the Cbz and benzyl groups by hydrogenation with a range of palladium compounds, which included 10% Pd/C, 10% Pd/C (Degussa), 20% $Pd(OH)_2$/C, Pd-black, and $Pd(OAc)_2$. Here, the use of an excess of catalyst was necessary because of catalyst poisoning that can occur in the presence of sulfur atoms. However, all catalysts failed to provide nucleotide **4**; harsh reaction conditions led to undesired over-reduced side products, whereas mild conditions resulted in the recovery of unreacted starting material.

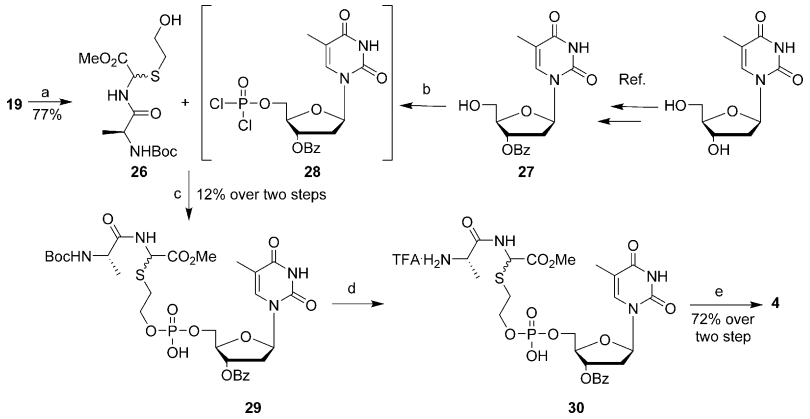
Therefore, an alternative protecting group strategy for the synthesis of target **4** was considered, as illustrated in Scheme 7. Thioethanol methyl ester **26** was isolated in good yield (77%) from its acetoxy derivative **19**. In analogy to the synthesis of compound **25**, the corresponding phosphodiester product **29** was obtained in low yield (12%) over two steps. It was found that the low conversion rate of nucleoside **27** into dichlorophosphoester intermediate **28** was the main cause for the formation of complex byproducts in the next base-promoted step. Other methods were explored to improve the yield of intermediate **28**, which included the use of $POCl_3$ in combination with different bases and solvents,[28] but none of them led to improved conversions, and the formation of dimers was detected even under controlled reaction conditions. Finally, the removal of the Boc group was achieved by using the TFA/thioanisole system followed by the simultaneous hydrolysis of the glycine ester and the 3′-benzoyl group with LiOH, which furnished nucleotide **4** in good yield (72%) over two steps.

As separation of diastereoisomers by preparative RP-HPLC represents a limitation with regards to the quantities of products that can be purified, we tried to perform this step at the later stages of the synthetic pathway for all the conjugates. However, effective separation could only be achieved in some cases. Therefore, it was thought that the enzymatic separation of the starting diastereoisomeric peptides could help to obtain larger quantities of pure final compounds. Accordingly, the diastereoisomeric mixture of *N*-Boc-L-Val-D,L-benzyloxy-Gly-OMe was successfully resolved by enzyme-catalyzed ester hydrolysis by using subtilisin Carlsberg (Scheme 8).[22] Conversely, the application of similar conditions with *N*-Boc-L-Ala-D,L-benzyloxy-Gly-OMe resulted in a poor separation, which may be due to a low substrate specificity.

For our last targets, phosphohemithioaminal ester derivatives **38a** and **38b**, the initially followed retrosynthetic Route 2 was quickly replaced by the approach described in Scheme 9 as it was not possible to isolate the *N*-Boc-L-alanyl-D,L-2-thioglycine methyl ester synthon in a satisfactorily pure form, which is likely due to the instability of the free thiol group. The synthesis of thymidine-5′-*O*-phosphorothioate **36** is reported in the literature by using a solid-phase methodology;[29] despite being effective for small quantities, this protocol was not reproducible upon scale-up.

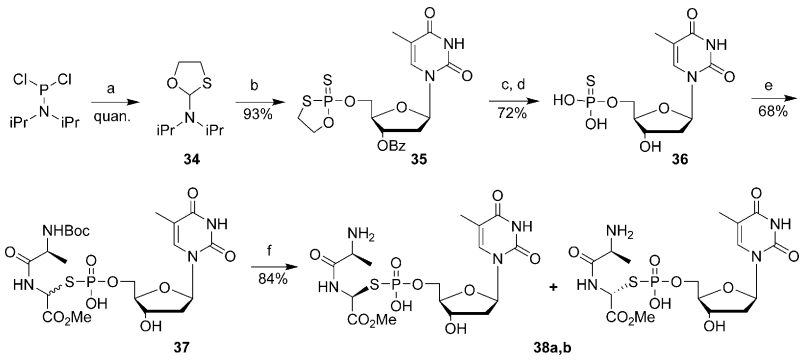


Scheme 8. Route towards the enzymatic separation of diastereoisomeric peptides. Reagents and conditions: (a) DABCO, dry THF, BnOH, −78 °C to RT, 24 h; (b) subtilisin Carlsberg, 1 M NaOH, DMF/H_2O 1:1, pH 7.2, 55 °C.



Scheme 7. Synthesis of the phosphoester of (L-alanyl-D,L-2-mercaptoethanolglycine)-TMP **4**. Reagents and conditions: (a) 2-mercaptoethanol, Et_3N, dry DMF, 0 °C to RT,16 h; (b) $POCl_3$, trimethyl phosphate, −20 °C to RT, 8 h (c) N-methyl imidazole, CH_2Cl_2, 0 °C to RT,16 h; (d) TFA, CH_2Cl_2, thioanisole, 0 °C to RT, 4 h; (e) LiOH, MeOH/H_2O 1:1, 0 °C to RT, 3 h.

© 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 9. Synthesis of the phosphohemithioaminals of (L-alanyl-D,L-2-aminoglycine methyl ester)]-TMP **38a** and **38b**. Reagents and conditions: (a) 2-mercapto ethanol, DIPEA, diethyl ether, −20 °C to RT, 2.5 h, (b) (i) 3′-*O*-benzoyl thymidine, *S*-ethylthio tetrazole, CH_2Cl_2, 0 °C to RT, 3 h; (ii) S_8, RT, overnight; (c) 3-hydroxy propionitrile, DBU, RT, 8 h; (d) 25% NH_3, 55 °C, sealed tube, 16 h; (e) Compound **19**, Et_3N, DMF, 37 °C, 4 h; (f) TFA, thioanisole, CH_2Cl_2, 0 °C to RT, 3 h.

Thus, monothiophosphorylation with $PSCl_3$[30] was next explored to access compound **36** in one step from thymidine, but the reaction rates were found to be sluggish and the process was non-selective in the presence of an excess of reagent. However, the protection of 3′-OH of thymidine and the use of an improved oxathiaphospholane approach that was adapted from the literature could overcome this issue.[31] As depicted in Scheme 9, compound **34** was obtained from *N,N*-diisopropyl phosphoramidous dichloride and 2-mercapto ethanol in the presence of DIPEA, and without further purification, it was reacted with 3′-*O*-benzoyl thymidine and *S*-ethylthio tetrazole. In situ sulfurization with elemental S_8 furnished oxathiaphospholane **35** in excellent yield (93%) over three steps. Opening of the oxathiaphospholane ring of compound **35** with 3-hydroxypropionitrile and simultaneous removal of 2-cyanoethyl and 3′-benzoyl groups by using 25% aqueous ammonia in a sealed tube produced compound **36** in good yield. Selective S-alkylation over N^3-alkylation of compound **36** with fragment **19** gave compound **37** in only 4 h as a diastereoisomeric mixture. Ultimately, Boc cleavage and preparative RP-HPLC purification allowed the isolation of diastereoisomers **38a** and **38b**.

When the two diastereoisomers were separately subjected to ester hydrolysis with LiOH, we isolated TMP as the major product. The in cellulo mode of breakdown of α-substituted glycine-based peptide conjugates after peptidase cleavage was postulated by Gilverg et al.[16a] In this context, the formation of TMP as a side product during the synthesis of compounds **1–3** could be explained by the anchimeric assistance of the α-carboxyl group of glycine towards the phosphorus atom, which led to the formation of a 5-membered-ring intermediate. This phenomenon has previously been noticed at different rates in the presence of phosphate, phosphoramidate, and phosphorothiate moieties.[32] By masking the free α-carboxyl group of glycine as a methyl ester, it was possible to resolve the *meta*-stability of the corresponding compounds.

Chain elongation experiments

In our laboratory, we are currently investigating the non-canonical mimics of nucleoside triphosphates as valuable targets for a range of studies into the substrate scope of DNA polymerases in directed enzyme evolution as well as in medicinal-chemistry-focused programs. As dipeptides are comparable in size to the pyrophosphate group,[6] which is realeased during DNA biosynthesis, it was envisioned that L-Ala-Gly could also act as an alternative leaving group in the enzyme-catalyzed DNA polymerization reaction. In this study, the activities of different thermophilic and mesophilic microbial DNA polymerases were compared, which included Therminator, Vent (*exo*-), and the Klenow fragment (*exo*-) of *E. coli* DNA Polymerase I. To investigate the strand elongation capacity of all the final analogues that exhibit α-substituted glycines connected through various phosphoheteroamic linkers (**1**, **2a**, **2b**, **23**, **4**, **38a** and **38b**), we carried out a polyacrylamide gel-based template-dependent incorporation assay of multiple nucleotides by using a tagged 5′-radiolabelled γ-^{33}P or a fluorescent DNA primer (**P1**) that was annealed to templates **T1**–**3** (Table 1).

Amongst all the compounds screened as potential triphosphate mimics, phosphoaminal ester analogues **2a** and **2b** exhibited interesting incorporation efficiencies with all tested polymerases (Figure 3 A). In contrast, compound **1** showed incorporation with only the Therminator polymerase and with a poor level of efficiency. On the basis of these observations, it can be deduced that the free carboxyl group along with an appropriate chain length are both crucial factors for incorporation. Vent (*exo*-) showed poorer efficiency than Therminator and the Klenow fragment; it only produced 4% (P+4) strand formation for **2a** and 11% (P+3) for **2b** at 1 mM concentration after 60 min. Therminator gave 76% yield of the full length product (P+5) and up to 4% (P+4) strand formation for **2a** and **2b**, respectively. The Klenow fragment showed the best incorporation efficiency, which led to 27% formation of the full length product (P+5) and up to 11% yield of the (P+4) strand at 1 mM concentration after only 15 min. Notably, compound

Table 1. Overview of primer and templates sequences used in the incorporation experiments.

Template T1:	3′ G T C C T T T G T C G A T A C T G A A A A A 5′
Primer P1:	5′/5Cy5/ C A G G A A A C A G C T A T G A C 3′
Template T3:	3′ G T C C T T T G T C G A T A C T G A T T T T 5′

© 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

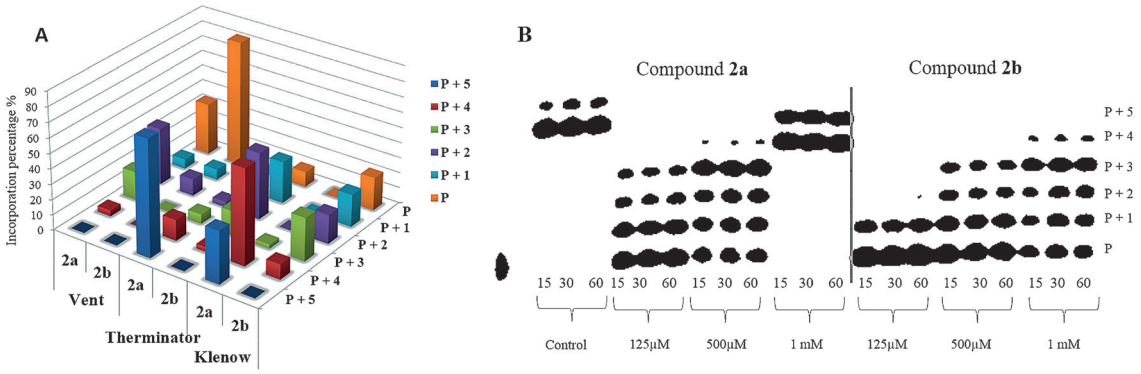


Figure 3. (A) Chain elongation efficiencies of compounds **2a** and **2b** by using Vent (*exo*-) (0.01 U μL^{-1}), Therminator (0.01 U μL^{-1}) and Klenow fragment (0.05 U μL^{-1}) polymerases at 1 mM after 60 min. (B) Chain elongation profile of compounds **2a** and **2b** as substrates into the P1:T1 complex by Klenow polymerase (0.05 U μL^{-1}).

2a exhibited better substrate properties than its diastereoisomer **2b**; this was probably because of different binding efficacy in the enzyme pocket, which might be approached with the aid of a modeling study. As the best incorporation results were obtained with compound **2a** and the Klenow fragment (*exo*-) polymerase, we compared the kinetic parameters of **2a** with its natural substrate counterpart TTP based on the Single Complete Hot model.

Steady-state kinetic analysis (Table 2) indicated a comparable V_{max} value, a K_M that was 991-fold higher, and a final V_{max}/K_M ratio for compound **2a** that was 994-fold lower than the natural substrate. This result falls within the range of kinetic values that have been previously reported for other leaving groups in our laboratory.[24]

Conclusion

A number of useful strategies were developed and effectively refined to generate different phosphodiester and mixed (N/O, S/O) phosphoester linker moieties in the preparation of a novel class of 5′-modified conjugates of TMP with the dipeptide L-Ala-Gly. The considerable challenges posed by the sensitive nature of the derived phosphoaminal, -hemiaminal or -hemithioaminal functionalities demanded a systematic elaboration of specifically tailored synthetic technologies. The introduction of an amino group at the α-position of a glycine residue successfully activated the corresponding dipeptide for attack at the phosphorus atom of TMP, which led to the requisite phosphoaminal ester conjugates. 5′-Peptidyl nucleotides that contain either an O–Cα–O–P or a S–Cα–O–P bond were otherwise assembled by a nucleophilic substitution reaction of the α-acetoxy group of glycine. Overall, we feel that the extensive synthetic optimization that was needed during the execution of this work could develop into new general reaction procedures for the synthesis of α-phosphorylated pseudopeptides and nucleoside monophosphate mimics.

In a template-based multiple incorporation assay, analogues **2a** and **2b** featuring a phosphoaminal ester linker and the Klenow fragment (exonuclease-free) of DNA polymerase I were found to be the best substrate and the most effective polymerase, respectively, and resulted in the best strand elongation rates. We found that a free carboxyl group at the C-terminal of the dipeptidic chain was necessary for binding in the active site of the enzyme, but it was also accountable for the destabilization of the conjugates and led to hydrolytic cleavage through anchimeric assistance. The further exploration of such dipeptides as leaving groups for DNA polymerization can be a valuable tool for synthetic biology, nucleotide delivery, polymerase chain reaction (PCR), and as potential prodrugs for antiviral therapy.

Table 2. Steady-state kinetics of single nucleotide incorporation into P1:T3 by Klenow fragment (0.005 U μL^{-1}).

Substrate	V_{max} [nM min^{-1}]	K_M [μM]	V_{max}/K_M [$\times 10^{-3}$ min^{-1}]
TTP	30.59 ± 0.74	0.088 ± 0.006	347.61
Compound **2a**	30.48 ± 1.04	87.21 ± 9.21	0.3495

Experimental Section

General Information

For all reactions, analytical grade solvents were used. TMP-disodium salt was transformed into its triethylammonium salt by treatment with Amberlite IRA-120B ion exchange resin (H^+ form) and then with triethylamine. All moisture-sensitive reactions were carried out in 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 TLC. Compounds were visualized with UV light ($\lambda = 254$ nm). Products were purified by flash chromatography on ICN silica gel 63–200 μm, 60 Å. All final compounds were purified by preparative RP-HPLC (Xbridge™ Prep C18 5 μm OBD 19×150 mm column) by using an eluent gradient of H_2O/CH_3CN with 50 mmol TEAB as eluent buffer. ^1H, ^{13}C, and ^{31}P NMR spectra were recorded on Bruker Avance 300, 500, or 600 MHz spectrometers. Final compounds were characterized by using 2D NMR (H-COSY, HSQC, HMBC, TOCSY, and NOESY) techniques. For the sake of clarity, NMR signals of protons and carbons for sugar and base moieties are indicated with and without a prime, respectively. Chemical shifts were referenced to residual solvent signals at $\delta_{H/C} = 7.26/77.00$ ($CDCl_3$), 3.31/49.10 (CD_3OD), 1.94/118.7 (CD_3CN), and 2.50/39.50 ([D_6]DMSO) relative to TMS as an internal standard wherever applied. Coupling constants are expressed in hertz (Hz) and were di-

© 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

rectly obtained from the spectra. Splitting patterns are designated as s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), br (broad), and apparent (app). High-resolution mass spectra (HRMS) were obtained on a quadruple orthogonal acceleration time-of-flight mass spectrometer (Synapt G2 HDMS, Waters, Milford, MA). Samples were infused at 3 μL min^{-1} and spectra were obtained in positive (or negative) ionization mode with a resolution of 15000 (FWHM) by using leucine enkephalin as lock mass.

N-Carbobenzyloxy-L-alanyl-D,L-2-hydroxyglycine benzyl ester (5)

Benzyl glyoxalate (1.85 g, 11.25 mmol) and a catalytic amount of *p*-toluene sulfonic acid (0.03 g, 0.18 mmol) were added to a stirred suspension of *N*-carbobenzoxy-L-alanine amide (2.0 g, 9.0 mmol) in CH_2Cl_2 (40 mL). The reaction mixture was stirred at room temperature for 7 days and monitored by TLC. Upon completion, the reaction mixture was concentrated in vacuo and the resulting crude residue was purified by column chromatography on silica gel (EtOAc/hexane, 1:4–2:3–3:2, v/v) to give compound **5** (2.44 g, 70%) as a white solid. ^{1}H NMR (300 MHz, [D$_6$]DMSO) δ = 8.79 and 8.77 (2×d, J = 8.5 Hz, 1H, N*H*–Gly), 7.44 (d, J = 7.9 Hz, 1H, N*H*–Ala), 7.37–7.33 (m, 10H, 2×OBn), 6.74–6.65 (m, 1H, α*H*–Gly), 5.53 (br s, 1H, αO*H*–Gly), 5.16–5.14 (m, 2H, C*H*$_2$–CO$_2$Bn), 5.02 (s, 2H, C*H*$_2$–Cbz), 4.17–4.07 (m, 1H, α*H*–Ala), 1.19 ppm (app t, J = 7.7 Hz, 3H, C*H*$_3$–Ala); ^{13}C NMR (75 MHz, [D$_6$]DMSO) δ = 172.5 and 172.5 (CO-Ala), 169.6 and 169.6 (CO-Gly), 155.6 (NHCONH), 137.0 and 137.0 (1C of OCH$_2$Ph), 135.7 and 135.7 (1C of OCH$_2$Ph), 128.4 and 128.4 (Ar-C), 128.2 (Ar-C), 128.0 (Ar-C), 127.9 (Ar-C), 127.8 and 127.8 (Ar-C), 127.7 and 127.7 (Ar-C), 71.4 and 71.4 (αC-Gly), 66.1 and 66.1 (CH$_2$-CO$_2$Bn), 65.4 and 65.4 (CH$_2$-Cbz), 49.9 and 49.8 (αC-Ala), 18.1 and 18.0 ppm (CH$_3$-Ala); HRMS (ESI): *m/z* calcd for $C_{20}H_{22}N_2O_6$: 385.1405 [*M*−H]$^-$; found: 385.1402.

3′-*O*-Benzyl-5′-*O*-(*O*-benzyl-*N*,*N*-diisopropyl phosphoramidite)thymidine (9)

A solution of benzyloxy-bis(*N*,*N*-diisopropylamino)phosphine (1.68 mL, 0.92 mmol, 0.549 M in anhydrous CH_2Cl_2) was added to a stirred solution of compound **8** (300 mg, 0.903 mmol) in anhydrous CH_2Cl_2 (18 mL) followed by a solution of 1H-tetrazole (2.04 mL, 0.92 mmol, 0.45 M in anhydrous AcCN) at 0 °C. The reaction mixture was stirred at room temperature for 4 h, and it was then quenched with saturated aq. NaHCO$_3$ and extracted with CH_2Cl_2. The organic phase was dried over Na_2SO_4 and concentrated in vacuo. The crude residue was purified by flash chromatography on silica gel (Et$_3$N/EtOAc/hexane, 0.2:20:79.8–0.2:40:59.8–0.2:80:19.8). The fractions that contained the phosphoramidite were combined and concentrated to afford compound **9** (497 mg, 72% over two steps) as a white solid. ^{1}H NMR (300 MHz, CDCl$_3$) δ = 7.67 and 7.56 (2×s, 1H, H-6), 7.36–7.24 (m, 10H, 2×OBn), 6.43–6.34 (m, 1H, H-1′), 4.80–4.63 (m, 2H, C*H*$_2$P–OBn), 4.58–4.34 (m, 2H, H-5′ and H-5″), 4.27 and 4.23 (2×s, 2H, C*H*$_2$–3′-OBn), 3.79–3.78 (m, 1H, H-3′), 3,64–3.60 (m, 3H, H-4′ and 2×*H*-*i*Pr), 2.54–2.38 (m, 1H, H-2′), 2.01–1.95 (m, 1H, H-2″), 1.86 and 1.80 (2×s, 3H, C*H*$_3$–Thy), 1.23–1.15 ppm (m, 12H, C*H*$_3$, 4×*i*Pr); ^{31}P NMR (121 MHz, CDCl$_3$) δ = 148.6 and 148.1 ppm; HRMS (ESI$^+$): *m/z* calcd for $C_{30}H_{40}N_3O_6P$: 570.2727 [*M*+H]$^+$; found: 570.2732.

N-Carbobenzyloxy-L-alanyl-D,L-2-acetoxyglycine benzyl ester (12)

Pyridine (21.6 mL) was added dropwise to a stirred suspension of compound **5** (2.4 g, 6.21 mmol) in acetic anhydride (30 mL) at 0 °C. The reaction mixture was stirred at the same temperature for 8 h and monitored by TLC. Upon completion, the mixture was concentrated in vacuo (bath temp. ≈15 °C). The residue was diluted with EtOAc (150 mL) and washed with H_2O (2×50 mL), 1N HCl (2×50 mL), 5% NaHCO$_3$ (2×50 mL), and brine (100 mL). The organic phase was dried over Na_2SO_4, filtered, and concentrated in vacuo. The resulting crude residue was triturated with cold Et$_2$O to give compound **12** (2.02 g, 76%) as a white solid. ^{1}H NMR (300 MHz, CDCl$_3$) δ = 7.59, 7.54 (2×d, J = 8.5 Hz, 1H, N*H*-Gly), 7.35–7.26 (m, 10H, 2×OBn), 6.43, 6.39 (2×d, J = 8.9 Hz, 1H, α*H*-Gly), 5.31 (app t, 1H, N*H*-Ala), 5.20 (s, 2H, C*H*$_2$-CO$_2$Bn), 5.09 (s, 2H, C*H*$_2$-Cbz), 4.35–4.30 (m, 1H, α*H*-Ala), 2.05 (s, 3H, C*H*$_3$-OAc), 1.38 and 1.36 ppm (2×d, J = 7.0 Hz, 3H, C*H*$_3$-Ala); ^{13}C NMR (75 MHz, CDCl$_3$) δ = 172.4 and 172.4 (CO-Ala), 170.3 and 170.3 (CO-OAc), 166.4 (CO-Gly), 156.0 (OCONH), 136.1 (1C of OCH$_2$Ph), 134.8 (1C of OCH$_2$Ph), 128.8 (Ar-C), 128.7 (Ar-C), 128.4 (Ar-C), 128.2 (Ar-C), 72.3 and 72.3 (αC-Gly), 68.3 (CH$_2$-CO$_2$Bn), 67.4 and 67.4 (CH$_2$-Cbz), 50.5 (αC-Ala), 20.7 (CH$_3$-OAc), 18.5 and 18.2 ppm (CH$_3$-Ala); HRMS (ESI$^+$): *m/z* calcd for $C_{22}H_{24}N_2O_7$: 451.1476 [*M*+Na]$^+$; found: 451.1474.

5′-*O*-(*N*-Carbobenzyloxy-L-alanyl-D,L-2-aminoglycine benzyl ester)-TMP (13) and 3-*N*-(*N*-carbobenzyloxy-L-alanyl-D,L-2-aminoglycine benzyl ester)-TMP triethylammonium salt (14)

Et$_3$N (0.45 mL, 3.234 mmol) was added to a stirred solution of TMP-triethylammonium salt (344 mg, 0.809 mmol) and compound **12** (476 mg, 1.496 mmol) in dry DMF (6 mL). The reaction mixture was heated at 40 °C for 24 h. It was allowed to cool to room temperature, the volatiles were removed in vacuo, and the resulting residue was dissolved in water and lyophilized. The crude product was purified by preparative RP-HPLC (98% H_2O+2% CH_3CN and 98% CH_3CN+2% H_2O). The collected fractions were freeze-dried repeatedly until the mass remained constant to afford compound **13** (49.6 mg, 8%) as a minor product and compound **14** (270 mg, 40%) as the major product, both as white solids. Data for compound **13**: ^{1}H NMR (600 MHz, CD$_3$OD) δ = 7.77 and 7.76 (2×s, 1H, H-6), 7.36–7.26 (m, 10H, 2×OBn), 6.32–6.29 (m, 1H, H-1′), 6.06 and 6.01 (2×d, J = 10.0 Hz, 1H, α*H*-Gly), 5.22–5.15 (m, 2H, C*H*$_2$-CO$_2$Bn), 5.09–5.03 (m, 2H, C*H*$_2$-Cbz), 4.49–4.46 (m, 1H, H-3′), 4.19–4.16 (m, 1H, α*H*-Ala), 4.07–4.02 (m, 2H, H-5′ and H-5″), 3.99–3.97 (m, 1H, H-4′), 2.22–2.14 (m, 2H, H-2′ and H-2″), 1.90 and 1.89 (2×s, 3H, C*H*$_3$-Thy), 1.30–1.27 ppm (m, 3H, C*H*$_3$-Ala, merged with Et$_3$N); ^{13}C NMR (125 MHz, CD$_3$OD) δ = 175.1 and 175.0 (CO-Ala), 169.2 (CO-Gly), 166.6 and 166.5 (C-4), 158.2 and 158.1 (OCONH), 152.5 (C-2), 138.1 (C-6 and 1C OBn), 136.9 (1C OBn), 129.6 (Ar-C), 129.5 (Ar-C), 129.4 (Ar-C), 129.3 (Ar-C), 129.0 (Ar-C), 128.9 (Ar-C), 112.0 (C-5), 87.5 and 87.4 (d, $^3J_{C,P}$ = 9.4 Hz, C-4′), 86.2 and 86.1 (C-1′), 75.3 (d, $^2J_{C,P}$ = 12.2 Hz, αC-Gly), 72.9 and 72.8 (C-3′), 68.5 (CH$_2$-CO$_2$Bn), 67.7 (CH$_2$-Cbz), 66.6 and 66.5 (d, $^2J_{C,P}$ = 5.4 Hz, C-5′), 51.8 (αC-Ala), 40.9 (C-2′), 18.4 and 18.2 (CH$_3$-Ala), 12.6 ppm (CH$_3$-Thy); ^{31}P NMR (202 MHz, CD$_3$OD) δ = −2.1 and −2.2 ppm; HRMS (ESI): *m/z* calcd for $C_{30}H_{35}N_4O_{13}P$: 689.1865 [*M*−H]$^-$; found: 689.1879.

Data for compound **14**: ^{1}H NMR (500 MHz, CD$_3$OD) δ = 7.93 (s, 1H, H-6), 7.37–7.29 (m, 11H, α*H*-Gly and 2×OBn), 6.34–6.30 (m, 1H, H-1′), 5.28–5.08 (m, 2H, C*H*$_2$-CO$_2$Bn), 5.06–5.04 (m, 2H, C*H*$_2$-Cbz), 4.50 (br s, 1H, H-3′), 4.26–4.19 (m, 1H, α*H*-Ala), 4.08–4.04 (m, 3H, H-4′, H-5′ and H-5″), 2.26–2.19 (m, 2H, H-2′ and H-2″), 1.95 (s, 3H, C*H*$_3$-Thy), 1.34 and 1.32 ppm (2×d, J = 7.1 Hz, 3H, C*H*$_3$-Ala, merged with Et$_3$N); ^{13}C NMR (125 MHz, CD$_3$OD) δ = 175.4 and 175.3 (CO-Ala), 168.3 and 168.2 (CO-Gly), 164.3 (C-4), 158.3 and 158.2 (OCONH), 151.7 (C-2), 138.1 and 138.0 (1C OBn), 137.5 and 137.4 (C-6), 136.7 (1C OBn), 129.6 (Ar C), 129.5 (Ar C), 129.4 (Ar C), 129.2 (Ar C), 129.1 (Ar C), 129.0 (Ar C), 128.9 (Ar C), 128.8 (Ar C), 111.3 and 111.2 (C-5),

© 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

87.9 and 87.8 (d, $^{3}J_{C,P}=8.8$ Hz, C-4′), 87.1 and 87.0 (C-1′), 72.7 and 72.6 (C-3′), 68.9 (CH_2-CO_2Bn), 67.8 and 67.7 (CH_2-Cbz), 65.9 (d, $^{2}J_{C,P}=4.8$ Hz, C-5′), 58.2 and 58.1 (αC-Gly), 52.0 and 51.6 (αC-Ala), 41.1 and 41.0 (C-2′), 18.0 and 17.9 (CH_3-Ala), 13.0 ppm (CH_3-Thy); ^{31}P NMR (202 MHz, CD_3OD) $\delta=0.7$ ppm; HRMS (ESI): *m/z* calcd for $C_{30}H_{35}N_4O_{13}P$: 689.1865 $[M-H]^-$; found: 689.1862.

5′-*O*-(L-Alanyl-D,L-2-aminoglycine)-TMP (1)

10% Pd/C, Degussa-type (6 mg, 0.2 equiv w/w) was added to a stirred solution of compound **13** (30 mg, 0.0434 mmol, 1 equiv) in EtOH/H_2O (5:1, 5 mL), and evacuation was carried out with hydrogen atmosphere replacements (3×). The reaction mixture was stirred at room temperature for 4 h under an atmospheric pressure of hydrogen. Upon completion, the catalyst was removed by filtration through a cellulose filter (0.45 μm), and the ethanol was removed in vacuo (bath temp. ≈10 °C). The residue was lyophilized to obtain a crude product that was purified by preparative RP-HPLC (98% H_2O+2% CH_3CN and 98% CH_3CN+2% H_2O). The collected fractions were freeze-dried repeatedly until the mass remained constant to afford compound **1** as a white solid (8.5 mg, 42%). ^{1}H NMR (300 MHz, D_2O) $\delta=7.73$ and 7.70 (2×s, 1H, H-6), 6.37–6.30 (m, 1H, H-1′), 5.73 (d, $J=9.0$ Hz, 1H, α*H*-Gly), 4.60–4.54 (m, 1H, H-3′), 4.18–4.02 (m, 4H, H-4′, α*H*-Ala, H-5′ and H-5″), 2.37–2.33 (m, 2H, H-2′ and H-2″), 1.91 (s, 3H, CH_3-Thy), 1.53 (d, $J=7.1$ Hz, 3H, CH_3-Ala); ^{13}C NMR (125 MHz, D_2O) $\delta=172.8$ (CO-Gly), 169.9 and 169.8 (CO-Ala), 166.4 (C-4), 151.5 (C-2), 137.1 and 137.0 (C-6), 111.5 and 111.5 (C-5), 85.1 and 85.0 (d, $^{3}J_{C,P}=9.2$ Hz, C-4′), 84.7 and 84.6 (C-1′), 74.8 (C-3′), 70.9 and 70.8 (αC-Gly), 64.9 and 64.8 (d, $^{2}J_{C,P}=5.2$ Hz, C-5′), 48.8 and 48.7 (αC-Ala), 38.4 and 38.3 (C-2′), 15.9 and 15.8 (CH_3-Ala), 11.4 ppm (CH_3-Thy); ^{31}P NMR (121 MHz, D_2O) $\delta=-2.0$ ppm; HRMS (ESI): *m/z* calcd for $C_{15}H_{23}N_4O_{11}P$: 465.1028 $[M-H]^-$; found: 465.1029.

3-*N*-(L-Alanyl-D,L-2-aminoglycine)-TMP triethylammonium salt (15)

Following a similar procedure as the one used for the synthesis of compound **1**, compound **15** was obtained as a white solid (179.1 mg, 92%) starting from **14** (260 mg, 0.343 mmol), 10% Pd/C, Degussa-type (52 mg, 0.2 eq w/w) in a 5:1 EtOH:H_2O mixture (15 mL). ^{1}H NMR (500 MHz, D_2O) $\delta=7.76$ and 7.74 (2×s, 1H, H-6), 6.90 and 6.86 (2×s, 1H, α*H*-Gly), 6.32 (app t, $J=6.8$ Hz, 1H, H-1′), 4.55–4.52 (m, 1H, H-3′), 4.21–4.16 (m, 1H, α*H*-Ala), 4.15–4.14 (m, 1H, H-4′), 4.11–4.01 (m, 2H, H-5′ and H-5″), 2.36–2.33 (m, 2H, H-2′ and H-2″), 1.91 (s, 3H, CH_3-Thy), 1.58 and 1.41 ppm (2×d, $J=7.1$ Hz, 3H, CH_3-Ala); ^{13}C NMR (150 MHz, D_2O) $\delta=171.0$ and 170.9 (CO-Gly), 170.5 and 170.4 (CO-Ala), 164.1 (C-4), 150.7 (C-2), 135.9 (C-6), 110.6 (C-5), 85.6 (C-1′), 85.3 (d, $^{3}J_{C,P}=9.3$ Hz, C-4′), 70.7 (C-3′), 64.4 (C-5′), 58.3 and 58.2 (αC-Gly), 48.7 and 48.6 (αC-Ala), 38.6 (C-2′), 16.1 and 16.0 (CH_3-Ala), 11.9 ppm (CH_3-Thy); ^{31}P NMR (202 MHz, D_2O) $\delta=0.04$ ppm; HRMS (ESI): *m/z* calcd for $C_{15}H_{23}N_4O_{11}P$: 465.1028 $[M-H]^-$; found: 465.1025.

N-Carbobenzyloxy-L-alanyl-D,L-2-azidoglycine benzyl ester (16)

DPPA (1.56 mL, 7.26 mmol) was added dropwise to a stirred suspension of compound **5** (2.0 g, 5.18 mmol) and DBU (1.08 mL, 7.26 mmol) in dry toluene (50 mL) at 0 °C. The reaction mixture was slowly warmed to room temperature and left stirring for 16 h. The resulting mixture was concentrated in vacuo (bath temp. ≈15 °C). The residue was diluted with EtOAc (150 mL) and washed with brine (50 mL) and 5% HCl (2×50 mL). The organic phase was dried over Na_2SO_4, filtered, and concentrated in vacuo, and the resultant crude residue was purified by column chromatography on silica gel (EtOAc/hexane, 1:9–1:4–2:3, v/v) to give compound **16** (1.51 g, 71%) as a colorless foam. ^{1}H NMR (300 MHz, $CDCl_3$) $\delta=$ 7.81–7.77 (2×d, $J=7.1$ Hz, 1H, N*H*-Gly), 7.34–7.30 (m, 10H, 2× OBn), 5.79 and 5.78 (2×d, $J=8.2$ Hz, 1H, α*H*-Gly), 5.65 (d, $J=$ 6.3 Hz, 1H, N*H*-Ala), 5.21 and 5.20 (2×s, 2H, CH_2-CO_2Bn), 5.08 and 5.07 (2×s, 2H, CH_2-Cbz), 4.43–4.38 (m, 1H, α*H*-Ala), 1.36 ppm (app t, $J=6.3$ Hz, 3H, CH_3-Ala); ^{13}C NMR (75 MHz, $CDCl_3$) $\delta=173.3$ (CO-Ala), 166.4 (CO-Gly), 156.1 (OCONH), 136.2 (1C of OCH_2Ph), 134.4 (1C of OCH_2Ph), 128.9 (Ar-C), 128.7 (Ar-C), 128.6 (Ar-C), 128.5 (Ar-C), 128.3 (Ar-C), 128.1 (Ar-C), 68.5 (αC-Gly), 67.3 (CH_2-CO_2Bn), 64.7 and 64.6 (CH_2-Cbz), 50.4 (αC-Ala), 18.6, 18.3 ppm (CH_3-Ala); HRMS (ESI^+): *m/z* calcd for $C_{20}H_{21}N_5O_5$: 434.1435 $[M+Na]^+$; found: 434.1432.

N-Carbobenzyloxy-L-alanyl-D,L-2-aminoglycine benzyl ester (17)

Triphenylphosphine (0.92 g, 3.50 mmol) was added to a stirred solution of compound **16** (1.2 g, 2.92 mmol) and water (0.26 mL, 14.6 mmol) in THF (15 mL). The reaction mixture was heated at reflux for 2 h, cooled to room temperature, and concentrated in vacuo. Water (100 mL) was added to the residue, and the solution was extracted with CH_2Cl_2 (3×100 mL). The organic phase was dried over Na_2SO_4, filtered, and concentrated in vacuo. The resultant crude residue was purified by column chromatography on silica gel (MeOH/CH_2Cl_2, 1:99–1:49–1:19, v/v) to give compound **17** (0.79 g, 70%) as a colorless foam and triphenylphosphine oxide as a minor impurity. ^{1}H NMR (300 MHz, $CDCl_3$) $\delta=7.77$ and 7.73 (2×d, $J=8.3$ Hz, 1H, N*H*-Gly), 7.31–7.29 (m, 10H, 2×OBn), 5.74 (app t, $J=$ 9.0 Hz, 1H, N*H*-Ala), 5.17–5.07 (m, 3H, α*H*-Gly and CH_2-CO_2Bn), 5.06 and 5.04 (2×s, 2H, CH_2-Cbz), 4.30–4.20 (m, 1H, α*H*-Ala), 2.84 (br s, 2H, αNH_2-Gly), 1.31 and 1.30 ppm (2×d, $J=7.0$ Hz, 3H, CH_3-Ala); ^{13}C NMR (75 MHz, $CDCl_3$) $\delta=172.7$ (CO-Ala), 170.5 (CO-Gly), 156.0 (OCONH), 136.3 (1C of OCH_2Ph), 135.1 (1C of OCH_2Ph), 128.6 (Ar-C), 128.5 (Ar-C), 128.3 (Ar-C), 128.2 (Ar-C), 128.1 (Ar-C), 67.6 (αC-Gly), 67.0 (CH_2-CO_2Bn), 60.3 and 60.2 (CH_2-Cbz), 50.5 (αC-Ala), 18.5 ppm (CH_3-Ala); HRMS (ESI^+): *m/z* calcd for $C_{20}H_{23}N_3O_5$: 408.1530 $[M+Na]^+$; found: 408.1523.

Thymidine-5′-*O*-(*N*-carbobenzyloxy-L-alanyl-2-aminoglycine benzyl ester) phosphoramidates (18a) and (18b)

Et_3N (0.1 mL, 0.737 mmol) was added to a stirring suspension of TMP-triehtylammonium salt (400 mg, 0.737 mmol) and compound **17** (640 mg, 1.659 mmol) in *t*BuOH/H_2O (4:1, 15 mL) to facilitate dissolution, followed by DCC (532 mg, 2.579 mmol). The reaction mixture was heated at 85 °C for 2.5 h, and the reaction progress was monitored by TLC (CH_2Cl_2/MeOH/H_2O, 17:7:1). Upon completion, the mixture was cooled to room temperature and concentrated in vacuo. The resulting residue was resuspended in water and lyophilized to give a crude material, which was purified by column chromatography on silica gel (CH_2Cl_2/MeOH/H_2O 19:1:0–9:1:0–17:7:1) to provide the desired nucleoside phosphoramidates as salts. Semi-preparative RP-HPLC (98% H_2O+2% CH_3CN and 98% CH_3CN+2% H_2O) was employed for further separation to obtain the pure diastereoisomers in a ≈1:1 ratio as white powders (131 mg of compound **18a**, 129 mg of compound **18b**, overall yield **18a+18b**: 260 mg, 51%). The isolated products were freeze-dried repeatedly until the mass remained constant. Data for compound **18a**: ^{1}H NMR (500 MHz, CD_3CN+D_2O, 50 °C) $\delta=7.55$ (s, 1H, H-6), 7.32–7.28 (m, 10H, 2×OBn), 6.14 (app t, $J=6.8$ Hz, 1H, H-1′), 5.21 and 5.16 (2×d, $J=10.9$ Hz, 1H, α*H*-Gly), 5.14–4.99 (m, 4H, CH_2-CO_2Bn and CH_2-Cbz), 4.38–4.36 (m, 1H, H-3′), 4.11–4.06 (m, 1H,

© 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

αH-Ala), 3.93–3.91 (m, 1H, H-4′), 3.87–3.82 (m, 2H, H-5′ and H-5″), 2.17–2.13 (m, 2H, H-2′ and H-2″), 1.80 and 1.79 (2×s, 3H, CH_3-Thy), 1.20 ppm (app t, $J=7.0$ Hz, 3H, CH_3-Ala); ^{13}C NMR (150 MHz, CD_3CN+D_2O, 10 °C) $\delta=174.6$ (CO-Ala), 171.3 (CO-Gly), 166.3 (C-4), 157.5 (OCONH), 151.5 (C-2), 137.9 (C-6), 137.5 (1C of OCH_2Ph), 136.2 (1C of OCH_2Ph), 129.5 (Ar-C), 129.4 (Ar-C), 129.2 (Ar-C), 129.0 (Ar-C), 128.7 (Ar-C), 111.9 (C-5), 86.4 (d, $^3J_{C,P}=8.9$ Hz, C-4′), 84.8 (C-1′), 71.7 (C-3′), 68.3 (CH_2-CO_2Bn), 67.4 (CH_2-Cbz), 64.9 (d, $^2J_{C,P}=3.9$ Hz, C-5′), 61.1 (d, $^2J_{C,P}=9.8$ Hz, αC-Gly), 51.1 (αC-Ala), 39.8 (C-2′), 18.1 (CH_3-Ala), 12.6 ppm (CH_3-Thy); ^{31}P NMR (202 MHz, CD_3CN+D_2O, 50 °C) $\delta=4.4$ ppm; $[\alpha]_{589}^{20}=-0.147°$ ($c=1$ in CH_3OH); HRMS (ESI): *m/z* calcd for $C_{30}H_{36}N_5O_{12}P$: 688.2025 $[M-H]^-$; found: 688.2040.

Data for compound **18b**: ^1H NMR (300 MHz, CD_3CN+D_2O, 25 °C) $\delta=7.54$ (s, 1H, H-6), 7.32–7.27 (m, 10H, 2×OBn), 6.14 (app t, $J=6.8$ Hz, 1H, H-1′), 5.21 and 5.16 (2×d, $J=10.9$ Hz, 1H, αH-Gly), 5.11–4.93 (m, 4H, CH_2-CO_2Bn and CH_2-Cbz), 4.37–4.32 (m, 1H, H-3′), 4.11–4.01 (m, 1H, αH-Ala), 3.95–3.90 (m, 1H, H-4′), 3.89–3.85 (m, 2H, H-5′ and H-5″), 2.17–2.11 (m, 2H, H-2′ and H-2″), 1.79 and 1.77 (2×s, 3H, CH_3-Thy), 1.19 ppm (app t, $J=7.0$ Hz, 3H, CH_3-Ala); ^{13}C NMR (150 MHz, CD_3CN+D_2O, 10 °C) $\delta=174.7$ (CO-Ala), 171.3 (CO-Gly), 166.3 (C-4), 157.5 (OCONH), 152.0 (C-2), 137.9 (C-6), 137.5 (1C of OCH_2Ph), 136.3 (1C of OCH_2Ph), 129.5 (Ar-C), 129.4 (Ar-C), 129.0 (Ar-C), 128.7 (Ar-C), 112.0 (C-5), 86.3 (d, $^3J_{C,P}=8.7$ Hz, C-4′), 85.5 (C-1′), 71.6 (C-3′), 68.4 (CH_2-CO_2Bn), 67.4 (CH_2-Cbz), 64.9 (d, $^2J_{C,P}=5.1$ Hz, C-5′), 61.1 (d, $^2J_{C,P}=9.1$ Hz, αC-Gly), 51.1 (αC-Ala), 39.7 (C-2′), 18.1 (CH_3-Ala), 12.6 ppm (CH_3-Thy); ^{31}P NMR (121 MHz, CD_3CN+D_2O, 25 °C) $\delta=3.2$ ppm; $[\alpha]_{589}^{20}=-0.052°$ ($c=1$ in CH_3OH); HRMS (ESI): *m/z* calcd for $C_{30}H_{36}N_5O_{12}P$: 688.2025 $[M-H]^-$; found: 688.2019.

Thymidine-5′-*O*-(L-alanyl-2-aminoglycine) phosphoramidate (2a)

Following a similar procedure as the one used for the synthesis of compound **1**, compound **2a** was obtained as a white solid (45.5 mg, 54%) starting from **18a** (125 mg, 0.136 mmol, 1 equiv), 10% Pd/C, Degussa-type (25 mg, 0.2 eq w/w) in EtOH/H_2O (5:1, 10 mL). ^1H NMR (300 MHz, D_2O) $\delta=7.63$ (d, $J=1.1$ Hz, 1H, H-6), 6.14 (app t, $J=6.8$ Hz, 1H, H-1′), 5.23 (d, $J=11.3$ Hz, 1H, αH-Gly), 4.50–4.45 (m, 1H, H-3′), 4.06–4.02 (m, 1H, H-4′), 4.00–3.88 (m, 3H, αH-Ala, H-5′ and H-5″), 2.31–2.26 (m, 2H, H-2′ and H-2″), 1.82 (d, $J=1.1$ Hz, 3H, CH_3-Thy), 1.44 ppm (d, $J=7.1$ Hz, 3H, CH_3-Ala); ^{13}C NMR (150 MHz, D_2O) $\delta=172.0$ (d, $^3J_{C,P}=9.1$ Hz, CO-Gly), 169.9 (CO-Ala), 166.3 (C-4), 151.5 (C-2), 137.3 (C-6), 111.5 (C-5), 85.2 (d, $^3J_{C,P}=9.0$ Hz, C-4′), 84.8 (C-1′), 70.8 (C-3′), 64.2 (d, $^2J_{C,P}=4.6$ Hz, C-5′), 61.1 (d, $^2J_{C,P}=6.5$ Hz, αC-Gly), 48.7 (αC-Ala), 38.4 (C-2′), 16.0 (CH_3-Ala), 11.5 ppm (CH_3-Thy); ^{31}P NMR (121 MHz, D_2O) $\delta=4.4$ ppm; HRMS (ESI): *m/z* calcd for $C_{15}H_{24}N_5O_{10}P$: 464.1188 $[M-H]^-$; found: 464.1185.

Thymidine-5′-*O*-(L-alanyl-2-aminoglycine) phosphoramidate (2b)

Following a similar procedure as the one used for the synthesis of **1**, compound **2b** was obtained as a white solid (43.0 mg, 51%) starting from compound **18b** (125 mg, 0.136 mmol, 1 equiv) and 10% Pd/C, Degussa-type (25 mg, 0.2 eq w/w) in EtOH/H_2O (5:1, 10 mL). ^1H NMR (300 MHz, D_2O) $\delta=7.68$ (s, 1H, H-6), 6.25 (app t, $J=6.9$ Hz, 1H, H-1′), 5.24 and 5.19 (2×d, $J=11.0$ Hz, 1H, αH-Gly), 4.52–4.47 (m, 1H, H-3′), 4.11–4.07 (m, 1H, H-4′), 4.06–3.90 (m, 3H, αH-Ala, H-5′ and H-5″), 2.34–2.29 (m, 2H, H-2′ and H-2″), 1.86 (s, 3H, CH_3-Thy), 1.47 and 1.46 (2×d, $J=7.0$ Hz, 3H, CH_3-Ala); ^{13}C NMR (150 MHz, D_2O) $\delta=172.2$ (d, $^3J_{C,P}=9.2$ Hz, CO-Gly), 169.9 (CO-Ala), 166.4 (C-4), 151.6 (C-2), 137.3 (C-6), 111.5 (C-5), 85.3 (d, $^3J_{C,P}=8.4$ Hz, C-4′), 84.9 (C-1′), 71.0 (C-3′), 64.3 (d, $^2J_{C,P}=4.7$ Hz, C-5′), 61.1 (d, $^2J_{C,P}=9.1$ Hz, αC-Gly), 48.7 (αC-Ala), 38.5 (C-2′), 16.0 (CH_3-Ala), 11.5 ppm (CH_3-Thy); ^{31}P NMR (121 MHz, D_2O) $\delta=4.2$ ppm; HRMS (ESI): *m/z* calcd for $C_{15}H_{24}N_5O_{10}P$: 464.1188 $[M-H]^-$; found: 464.1187.

N-Boc-L-Alanyl-D,L-2-acetoxyglycine methyl ester (19)

Molecular sieves (4 Å, 5 g) and $Pb(OAc)_4$ (17.89 g, 40.30 mmol) were added to a stirring solution of Boc-L-Ala-L-Ser-OMe (3.9 g, 13.43 mmol) in dry ethyl acetate (220 mL)under an inert atmosphere. The reaction mixture was heated at reflux for 2 h and cooled to room temperature. The solid was removed by filtration through a pad of Celite, and the organic phase was stirred with 10% aq. citric acid until it became nearly colourless. The organic phase was separated, washed with 10% aq. citric acid, water, and brine, dried over Na_2SO_4, filtered, and concentrated in vacuo to give a crude diastereoisomeric mixture of compound **19** (4.27 g, quant.) as a colourless foam, which was used in the next step without any further purification. ^1H NMR (300 MHz, $CDCl_3$) $\delta=7.70–7.52$ (m, 1H, NH-Gly), 6.40 and 6.38 (2×d, $J=9.1$ Hz, 1H, αH-Gly), 5.31 (app t, $J=6.9$ Hz, 1H, NH-Ala), 4.27–4.13 (m, 1H, αH-Ala), 3.81 and 3.80 (2×s, 3H, OCH_3-Gly), 2.11 (s, 3H, CH_3-OAc), 1.45 (s, 9H, *t*Bu), 1.38 and 1.37 ppm (2×d, $J=7.1$ Hz, 3H, CH_3-Ala); ^{13}C NMR (75 MHz, $CDCl_3$) $\delta=172.8$ (CO-Ala), 170.3 and 170.2 (CO-Gly), 167.2 and 167.1 (CO-OAc), 155.6 (OCONH), 88.7 (1C-*t*Bu), 72.3 and 72.2 (αC-Gly), 53.4 and 53.3 (OCH_3-Gly), 50.3 and 50.1 (αC-Ala), 28.4 (*t*Bu), 20.7 (CH_3-OAc), 17.8 and 17.6 ppm (CH_3-Ala); HRMS (ESI^+): *m/z* calcd for $C_{13}H_{22}N_2O_7$: 341.1319 $[M+Na]^+$; found: 341.1318.

N-Boc-L-Alanyl-D,L-2-*p*-anisylaminoglycine methyl ester (20)

DIPEA (3.01 mL, 17.28 mmol) was added to a stirring solution of compound **19** (2.2 g, 6.91 mmol) and 4-methoxybenzylamine (1.0 mL, 7.60 mmol) in dry DMF (20 mL) at room temperature. The reaction mixture was heated at 40 °C for 24 h and then concentrated in vacuo. Water (200 mL) was added to the residue and the organics were extracted with ethyl acetate (3×120 mL). The combined organic phases were washed with brine, dried over Na_2SO_4, filtered, and concentrated in vacuo. The resultant crude material was purified by column chromatography on silica gel (EtOAc/hexane, 2:3–3:2–4:1) to give a diastereoisomeric mixture of compound **20** (2.07 g, 76%) as a colorless foam. ^1H NMR (300 MHz, $CDCl_3$) $\delta=7.22$ (d, $J=8.5$ Hz, 1H, *o*H-PMB), 7.04–6.93 (m, 1H, NH-Gly), 6.84 (d, $J=8.5$ Hz, 1H, *m*H-PMB), 5.27 (d, $J=7.8$ Hz, 1H, αH-Gly), 5.08–4.99 (m, 1H, NH-Ala), 4.20–4.18 (m, 1H, αH-Ala), 3.78 (s, 3H, OCH_3-PMB), 3.73–3.70 (m, 5H, OCH_3-Gly and CH_2-PMB), 2.37 (br s, 1H, NH-PMB), 1.46 and 1.45 (2×s, 9H, *t*Bu), 1.37 ppm (app t, $J=7.0$ Hz, 3H, CH_3-Ala); ^{13}C NMR (75 MHz, $CDCl_3$) $\delta=173.1$ (CO-Ala), 170.6 and 170.5 (CO-Gly), 159.0 (Ar-C), 155.6 (OCONH), 131.3 and 131.2 (1C of PMB), 129.7 (Ar-C), 129.3 (Ar-C), 114.3 (Ar-C), 113.9 (Ar-C), 80.4 (1C-*t*Bu), 64.4 and 64.3 (αC-Gly), 55.4 (OCH_3-PMB), 52.9 (OCH_3-Gly), 50.3 (αC-Ala), 48.5 (CH_2-PMB), 28.4 (*t*Bu), 18.3 and 18.1 ppm (CH_3-Ala); HRMS (ESI): *m/z* calcd for $C_{19}H_{29}N_3O_6$: 394.1983 $[M-H]^-$; found: 394.1973.

N-Boc-L-Alanyl-D,L-2-aminoglycine methyl ester (21)

10% Pd/C (0.18 g, 0.1 equiv w/w) was added to a stirring solution of compound **20** (1.8 g, 2.44 mmol) in EtOH (80 mL), and the mixture was hydrogenated at atmospheric pressure by using a balloon filled with H_2 for 5 h at room temperature. The catalyst was re-

© 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

moved by filtration through a pad of Celite, and the filtrate was concentrated in vacuo. The resulting crude residue was purified by column chromatography on silica gel ($MeOH/CH_2Cl_2$, 1:99–1:49–1:24) to give a diastereoisomeric mixture of compound **21** (1.2 g, 96%) as a colorless foam. ^1H NMR (300 MHz, $CDCl_3$) $\delta = 7.41$ (br s, 1H, N*H*-Gly), 5.28–5.24 (m, 2H, α*H*-Gly and N*H*-Ala), 4.20–4.18 (m, 1H, α*H*-Ala), 3.79 (s, 3H, OCH_3-Gly), 2.24 (br s, 1H, αNH_2-Gly), 1.45 (s, 9H, *t*Bu), 1.36 ppm (d, $J = 7.1$ Hz, 3H, CH_3-Ala); ^{13}C NMR (75 MHz, $CDCl_3$) $\delta = 173.1$ and 173.0 (CO-Ala), 171.2 and 171.1 (CO-Gly), 155.5 (OCONH), 80.3 (1C-*t*Bu), 60.2 and 60.1 (αC-Gly), 53.0 (OCH_3-Gly), 50.1 (αC-Ala), 28.3 (*t*Bu), 18.3 and 18.2 ppm (CH_3-Ala); HRMS (ESI^+): *m/z* calcd for $C_{19}H_{29}N_3O_6$: 298.1377 $[M+Na]^+$; found: 298.1390.

Thymidine-5′-*O*-(*N*-Boc-L-alanyl-D,L-2-aminoglycine methyl ester) phosphoramidate triethylammonium salt (22)

Following a similar procedure as the one used for the synthesis of compounds **18a,b**, the triethylammonium salt of compound **22** was obtained as a diastereoisomeric mixture (538 mg, 74%, white solid) starting from TMP-triehtylammonium salt (560 mg, 1.07 mmol), compound **21** (647 mg, 2.35 mmol), triethylamine (163 μL, 1.17 mmol), and DCC (881 mg, 4.27 mmol) in *t*BuOH/H_2O (4:1, 15 mL). ^1H NMR (500 MHz, CD_3OD) $\delta = 7.68$ (d, $J = 1.1$ Hz, 1H, H-6), 6.35–6.31 (m, 1H, H-1′), 5.30 and 5.22 (2×d, $J = 11.1$ Hz, 1H, α*H*-Gly) 4.48–4.45 (m, 1H, H-3′), 4.10–4.05 (m, 1H, α*H*-Ala), 4.02–4.00 (m, 1H, H-4′), 3.99–3.95 (m, 2H, H-5′ and H-5″), 3.71 and 3.69 (2×s, 3H, OCH_3-Gly), 2.29–2.17 (m, 2H, H-2′ and H-2″), 1.94 (d, $J = 1.1$ Hz, 3H, CH_3-Thy), 1.44 and 1.43 (2×s, 9H, *t*Bu), 1.29 ppm (app t, $J = 7.1$ Hz, 3H, CH_3-Ala, merged with Et_3N); ^{13}C NMR (125 MHz, CD_3OD) $\delta = 175.2$ (CO-Ala), 172.1 and 172.0 (CO-Gly), 166.5 (C-4), 157.5 (OCONH), 152.5 (C-2), 138.3 (C-6), 111.9 (C-5), 87.6 (app t, $^3J_{C,P} = 9.5$ Hz, C-4′), 86.2 and 86.0 (C-1′), 80.6 (1C-*t*Bu), 73.0 and 72.9 (C-3′), 65.7 (app t, $^2J_{C,P} = 6.9$ Hz, C-5′), 61.7 (αC-Gly), 53.1 and 53.0 (OCH_3-Gly), 51.3 (αC-Ala), 40.8 (C-2′), 28.7 (*t*Bu), 18.4 (CH_3-Ala), 12.7 ppm (CH_3-Thy); ^{31}P NMR (202 MHz, CD_3OD) $\delta = 3.4$ ppm; HRMS (ESI): *m/z* calcd for $C_{21}H_{34}N_5O_{12}P$: 578.1869 $[M-H]^-$; found: 578.1873.

Thymidine-5′-(L-alanyl-D,L-2-aminoglycine methyl ester) phosphoramidate TFA salt (23)

TFA (1.5 mL) was added to a stirring solution of compound **22** (220 mg, 0.323 mmol) and thioanisole (42 μL, 0.356 mmol) in CH_2Cl_2 (4.5 mL) at 0 °C. The reaction mixture was stirred at 0 °C for 1 h and then at room temperature for 3 h. After reaction completion, the volatiles were removed in vacuo and coevaporated three times with toluene to remove residual TFA. The residue was redissolved in H_2O and lyophilized (2×) to obtain a crude residue, which was purified by preparative RP-HPLC (98% H_2O+2% CH_3CN and 98% CH_3CN+2% H_2O). The collected fractions were freeze-dried repeatedly until the mass remained constant to obtain a diastereoisomeric mixture of compound **23** (158 mg, 85%, white solid). ^1H NMR (600 MHz, D_2O) $\delta = 7.72$ (s, 1H, H-6), 6.29 (app t, $J = 7.0$ Hz, 1H, H-1′), 5.64 (d, $J = 9.5$ Hz, 1H, α*H*-Gly), 4.54–4.52 (m, 1H, H-3′), 4.18–4.15 (m, 1H, α*H*-Ala), 4.14–4.12 (m, 1H, H-4′), 4.07–4.03 (m, 2H, H-5′ and H-5″), 3.84 (s, 3H, OCH_3-Gly), 2.32–2.30 (m, 2H, H-2′ and H-2″), 1.86 (s, 3H, CH_3-Thy), 1.53 and 1.52 ppm (2×d, $J = 7.2$ Hz, 3H, CH_3-Ala); ^{13}C NMR (150 MHz, D_2O) $\delta = 171.6$ and 171.4 (CO-Ala), 166.3 (CO-Gly), 166.2 (C-4), 151.5 (C-2), 137.1 (C-6), 111.4 (C-5), 85.2 (app t, $^3J_{C,P} = 9.0$ Hz, C-4′), 84.8 (C-1′), 70.9 (C-3′), 64.7 (d, $^2J_{C,P} = 4.5$ Hz, C-5′), 56.8 (d, $^2J_{C,P} = 7.5$ Hz, αC-Gly), 54.1 (OCH_3-Gly), 48.5 (αC-Ala), 38.6 (C-2′), 15.9 and 15.8 (CH_3-Ala), 11.4 ppm (CH_3-Thy); ^{31}P NMR (202 MHz, D_2O) $\delta = -0.1$ ppm; HRMS (ESI): *m/z* calcd for $C_{16}H_{26}N_5O_{10}P$: 478.1344 $[M-H]^-$; found: 478.1346.

N-Carbobenzyloxy-L-alanyl-D,L-2-mercaptoethanolglycine benzyl ester (24)

Et_3N (0.55 mL, 3.92 mmol) was added dropwise to a stirring solution of compound **12** (1.4 g, 3.27 mmol) and 2-mercaptoethanol (0.25 mL, 3.59 mmol) in dry DMF (12 mL) at 0 °C. The reaction mixture was slowly warmed to room temperature and left stirring for 16 h. The resulting mixture was concentrated in vacuo. Water (100 mL) was added to the residue and the mixture was extracted with ethyl acetate (3×100 mL). The combined organic phases were washed with brine, dried over Na_2SO_4, filtered, and concentrated. The resultant crude material was purified by column chromatography on silica gel (EtOAc/hexane, 1:4–2:3–3:2) to give a diastereoisomeric mixture of compound **24** (1.2 g, 70%) as a colorless foam. ^1H NMR (300 MHz, $[D_6]DMSO$) $\delta = 8.83$ and 8.79 (2×d, $J = 8.2$ Hz, 1H, N*H*-Gly), 7.47 (d, $J = 7.3$ Hz, 1H, N*H*-Ala), 7.39–7.32 (m, 10H, 2×OBn), 5.46 and 5.44 (2×d, $J = 10.0$ Hz, 1H, α*H*-Gly), 5.19–5.16 (m, 2H, CH_2-CO_2Bn), 5.00 (s, 2H, CH_2-Cbz), 4.96–4.93 (m, 1H, O*H*), 4.19–4.12 (m, 1H, α*H*-Ala), 3.54–3.48 (m, 2H, OCH_2CH_2S), 2.72–2.65 (m, 2H, OCH_2CH_2S), 1.19 ppm (app t, $J = 7.8$ Hz, 3H, CH_3-Ala); ^{13}C NMR (75 MHz, $[D_6]DMSO$) $\delta = 172.7$ and 172.6 (CO-Ala), 168.5 and 168.4 (CO-Gly), 155.7 and 155.6 (OCONH), 137.0 (1C of OCH_2Ph), 135.6 and 135.5 (1C of OCH_2Ph), 128.5 (Ar-C), 128.4 (Ar-C), 128.2 (Ar-C), 127.9 (Ar-C), 127.7 (Ar-C), 66.8 and 66.7 (CH_2-CO_2Bn), 65.5 (CH_2-Cbz), 60.4 and 60.3 (OCH_2CH_2S), 53.2 and 53.1 (αC-Gly), 49.9 and 49.8 (αC-Ala), 32.7 and 32.6 (OCH_2CH_2S), 18.1 and 18.0 ppm (CH_3-Ala); HRMS (ESI): *m/z* calcd for $C_{22}H_{26}N_2O_6S$: 445.1439 $[M-H]^-$; found: 445.1443.

3′-*O*-Benzyl-5′-*O*-(*N*-carbobenzyloxy-L-alanyl-D,L-2-mercaptoethanolglycine benzyl ester) TMP triethylammonium salt (25)

$POCl_3$ (169 μL, 3.92 mmol) was added dropwise to a stirring solution of compound **8** (0.5 g, 1.50 mmol) in dry trimethyl phosphate (5 mL) at −20 °C. The reaction mixture was slowly warmed to room temperature and was left stirring for 8 h until reaction completion (monitored by TLC). The solution was cooled to 0 °C, and a solution of compound **24** (0.672 g, 1.50 mmol) and *N*-methyl imidazole (0.36 mL, 4.51 mmol) in dry CH_2Cl_2 (14 mL) was added. The resulting mixture was stirred at room temperature for 16 h. It was then quenched with water at 0 °C, the volatiles were removed in vacuo, water was added to the residue, and the mixture was lyophilized (3×). The resultant crude residue was purified by column chromatography on silica gel ($MeOH/CH_2Cl_2$, 1:99–3:97–8:92) and by preparative RP-HPLC (25 mmol TEAB in 98% H_2O+2% CH_3CN and 25 mmol TEAB in 98% CH_3CN+2% H_2O). The collected fractions were freeze-dried repeatedly until the mass remained constant to obtain a diastereoisomeric mixture of compound **25** (935 mg, 66%) as a white solid. ^1H NMR (300 MHz, CD_3OD) $\delta = 7.76$ (s, 1H, H-6), 7.35–7.25 (m, 15H, 3×OBn), 6.45 (dd, $J = 8.7$, 5.8 Hz, 1H, H-1′), 5.51 (d, $J = 13.4$ Hz, 1H, α*H*-Gly), 5.24–5.13 (m, 2H, CH_2-CO_2Bn), 5.06 (s, 2H, CH_2-Cbz), 4.57–4.55 (m, 2H, CH_2-3′OBn), 4.34–4.31 (m, 1H, α*H*-Ala), 4.25–4.18 (m, 1H, H-3′ and H-4′), 4.07–4.01 (m, 2H, H-5′ and H-5″), 4.00–3.90 (m, 2H, OCH_2CH_2S), 2.93–2.78 (m, 2H, OCH_2CH_2S), 2.40–2.17 (m, 2H, H-2′ and H-2″), 1.91 (s, 3H, CH_3-Thy), 1.34 and 1.32 ppm (2×d, $J = 7.2$ Hz, 3H, CH_3-Ala); ^{13}C NMR (75 MHz, CD_3OD) $\delta = 175.2$ (CO-Ala), 169.7 and 169.6 (CO-Gly), 166.4 and 166.3 (C-4), 158.1 (OCONH), 152.5 and 152.4 (C-2), 139.4 (1C Ar-C), 138.2 (1C Ar-C), 137.9 (C-6), 136.8 (1C Ar-C), 129.6 (Ar-C), 129.5 (Ar-C), 129.4 (Ar-C), 129.3 (Ar-C), 129.2 (Ar-C), 129.0 (Ar-C), 128.9 (Ar-C),

© 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

128.8 (Ar-C), 128.7 (Ar-C), 112.1 and 112.0 (C-5), 86.2 and 86.1 (C-1′), 85.1 (d, $^3J_{C,P}=8.7$ Hz, C-4′), 81.1 and 81.0 (C-3′), 72.2 and 72.1 (CH_2-3′OBn), 68.6 and 68.5 (CH_2-CO_2Bn), 67.6 (CH_2-Cbz), 66.9 (d, $^2J_{C,P}=5.5$ Hz, C-5′), 65.8 and 65.7 (d, $^2J_{C,P}=5.4$ Hz, OCH_2CH_2S), 54.8 and 54.7 (αC-Gly), 52.0 and 51.9 (αC-Ala), 38.1 (C-2′), 32.1 and 32.0 (OCH_2CH_2S), 18.3 and 18.2 (CH_3-Ala), 12.7 ppm (s, CH_3-Thy); ^{31}P NMR (121 MHz, CD_3OD) $\delta=-0.5$ ppm; HRMS (ESI): *m/z* calcd for $C_{39}H_{45}N_4O_{13}PS$: 839.2368 $[M-H]^-$; found: 839.2376.

N-Boc-L-Alanyl-D,L-2-mercaptoethanolglycine methyl ester (26)

Following a similar procedure as the one used for the synthesis of compound **24**, a diastereoisomeric mixture of compound **26** was obtained as a colorless foam (3.5 g, 77%) starting from compound **19** (4.28 g, 13.43 mmol), 2-mercaptoethanol (1.04 mL, 14.77 mmol), and Et_3N (2.24 mL, 16.12 mmol) in dry DMF (12 mL). ^{1}H NMR (300 MHz, $CDCl_3$) $\delta=7.63$ (br s, 1H, N*H*-Gly), 5.56–5.50 (m, 1H, α*H*-Gly), 5.10 and 5.06 (2×d, $J=7.0$ Hz, 1H, N*H*-Ala), 4.28–4.13 (m, 1H, α*H*-Ala), 3.92–3.77 (m, 5H, OCH_2CH_2S and OCH_3-Gly), 2.90–2.85 (m, 2H, OCH_2CH_2S), 1.46 and 1.45 (2×s, *t*Bu), 1.38 and 1.37 ppm (2×d, $J=7.0$ Hz, 3H, CH_3-Ala); ^{13}C NMR (75 MHz, $CDCl_3$) $\delta=172.5$ (CO-Ala), 169.8 and 169.6 (CO-Gly), 155.6 (OCONH), 80.7 (1C-*t*Bu), 62.1 (OCH_2CH_2S), 53.2 (OCH_3-Gly), 52.9 (αC-Gly), 50.4 (αC-Ala), 34.2 and 34.1 (OCH_2CH_2S), 28.4 (*t*Bu), 18.0 and 17.9 ppm (CH_3-Ala); HRMS (ESI$^+$): *m/z* calcd for $C_{13}H_{24}N_2O_6S$: 359.1247 $[M+Na]^+$; found: 359.1240.

3′-*O*-Benzoyl-5′-*O*-(*N*-Boc-L-alanyl-D,L-2-mercaptoethanolglycine methyl ester)TMP (29)

Following a similar procedure as the one used for the synthesis of compound **25**, a diastereoisomeric mixture of compound **29** was obtained as a white solid (101.5 mg, 12%) starting from **27** (394 mg, 1.138 mmol), $POCl_3$ (128 μL, 1.366 mmol) in dry trimethyl phosphate (4 mL) and compound **26** (383 mg, 1.138 mmol), *N*-methyl imidazole (272 μL, 1.366 mmol) in dry CH_2Cl_2 (10 mL). The resultant crude residue was purified by column chromatography on silica gel (MeOH/CH_2Cl_2, 1:99–3:97–8:92) and preparative RP-HPLC (98% H_2O+2% CH_3CN and 98% CH_3CN+2% H_2O). The collected fractions were freeze-dried repeatedly until the mass remained constant. ^{1}H NMR (300 MHz, CD_3OD) $\delta=8.06$ (d, $J=7.5$ Hz, 2H, *oH*-OBz), 7.78 (s, 1H, H-6), 7.64 (t, $J=7.4$ Hz, 1H, *pH*-OBz), 7.64 (t, $J=7.6$ Hz, 2H, *mH*-OBz), 6.45 (app t, $J=7.5$ Hz, 1H, H-1′), 5.67–5.62 (m, 1H, H-3′), 5.59–5.56 (m, 1H, α*H*-Gly), 4.67–4.37 (m, 1H, H-4′), 4.35–4.23 (m, 2H, H-5′ and H-5″), 4.18–4.04 (m, 3H, OCH_2CH_2S and α*H*-Ala), 3.74 (s, 3H, OCH_3-Gly), 3.02–2.92 (m, 2H, OCH_2CH_2S), 2.58–2.50 (m, 2H, H-2′ and H-2″), 1.95 (s, 3H, CH_3-Thy), 1.42 (s, 9H, *t*Bu), 1.30 and 1.29 ppm (2×d, $J=7.0$ Hz, 3H, CH_3-Ala); ^{13}C NMR (150 MHz, CD_3OD) $\delta=175.4$ and 175.3 (CO-Ala), 170.4 and 170.3 (CO-Gly), 167.3 (CO-OBz), 166.3 and 166.2 (C-4), 157.6 (OCONH), 152.5 and 152.4 (C-2), 137.8 (C-6), 134.6 and 134.5 (1C Ar-C), 130.9 (Ar-C), 130.8 (Ar-C), 130.7 (Ar-C), 130.6 (Ar-C), 129.7 (Ar-C), 129.6 (Ar-C), 112.4 and 112.3 (C-5), 86.3 and 86.1 (C-1′), 85.0 (d, $^3J_{C,P}=8.5$ Hz, C-4′), 80.6 (1C-*t*Bu), 77.5 and 77.2 (C-3′), 67.0 (d, $^2J_{C,P}=3.0$ Hz, C-5′), 66.1 (d, $^2J_{C,P}=6.0$ Hz, OCH_2CH_2S), 54.7 and 54.6 (αC-Gly), 53.5 (OCH_3-Gly), 51.5 (αC-Ala), 38.2 and 38.1 (C-2′), 32.0 (OCH_2CH_2S), 28.7 (*t*Bu), 18.2 and 18.1 (CH_3-Ala), 12.9 and 12.8 ppm (CH_3-Thy); ^{31}P NMR (202 MHz, D_2O) $\delta=-0.5$ ppm; HRMS (ESI): *m/z* calcd for $C_{30}H_{41}N_4O_{14}PS$: 743.2005 $[M-H]^-$; found: 743.2014.

5′-*O*-(L-Alanyl-D,L-2-mercaptoethanolglycine) TMP (4)

Following a similar procedure as the one used for the synthesis of compound **23**, the TFA salt of the crude diastereoisomeric amine **30** was obtained as a sticky mass (≈100 mg, quant.) starting from compound **29** (100 mg, 0.134 mmol), thioanisole (16 μL, 0.134 mmol), and TFA (0.5 mL) in CH_2Cl_2 (1.5 mL). HRMS (ESI): calcd for $C_{25}H_{33}N_4O_{12}PS$ [M-H]$^-$ calcd.: 643.1480, found: 643.1472. The obtained residue **30** was treated, without further purification, with LiOH (19.4 mg, 0.809 mmol) in MeOH/H_2O (1:1, 2 mL) at 0 °C, and the solution was stirred at room temperature for 3 h. After reaction completion, the mixture was neutralized with 5% acetic acid and concentrated in vacuo. The crude residue was purified by preparative RP-HPLC (98% H_2O+2% CH_3CN and 98% CH_3CN+2% H_2O). The collected fractions were freeze-dried repeatedly until the mass remained constant to afford a diastereoisomeric mixture of compound **4** (51 mg, 72% over two steps) as a white solid. ^{1}H NMR (500 MHz, D_2O) $\delta=7.60$ (s, 1H, H-6), 6.39 (app t, $J=6.9$ Hz, 1H, H-1′), 5.22 (d, $J=20.5$ Hz, 1H, α*H*-Gly), 4.56–4.53 (m, 1H, H-3′), 4.12–4.10 (m, 1H, H-4′), 4.07–4.03 (m, 2H, H-5′ and H-5″), 4.00–3.94 (m, 2H, OCH_2CH_2S), 3.51–3.46 (m, 1H, α*H*-Ala), 2.83–2.79 (m, 2H, OCH_2CH_2S), 2.31–2.26 (m, 2H, H-2′ and H-2″), 1.85 (s, 3H, CH_3-Thy), 1.21 and 1.19 (2×d, $J=7.0$ Hz, 3H, CH_3-Ala); ^{13}C NMR (125 MHz, D_2O) $\delta=177.0$ (CO-Gly), 173.0 (CO-Ala), 165.5 (C-4), 156.7 (C-2), 135.9 and 135.8 (C-6), 111.3 (C-5), 84.4 (d, $^3J_{C,P}=8.4$ Hz, C-4′), 84.3 (C-1′), 70.5 and 70.4 (C-3′), 64.5 (d, $^2J_{C,P}=5.1$ Hz, C-5′), 64.1 (d, $^2J_{C,P}=5.5$ Hz, OCH_2CH_2S), 55.6 (αC-Gly), 49.2 (αC-Ala), 38.2 (C-2′), 29.6 (d, $^3J_{C,P}=7.8$ Hz, OCH_2CH_2S), 19.1 (CH_3-Ala), 12.0 and 11.9 ppm (CH_3-Thy); ^{31}P NMR (202 MHz, D_2O) $\delta=-0.3$ ppm; HRMS (ESI): *m/z* calcd for $C_{17}H_{27}N_4O_{11}PS$: 525.1062 $[M-H]^-$; found: 525.1058.

N-Boc-L-Alanyl-D,L-2-*O*-benzylglycine methyl ester (31)

A solution of DABCO (1.35 g, 12.0 mmol) in dry THF (15 mL) was added to a stirring solution of compound **19** (1.59 g, 5.0 mmol) in dry THF (100 mL) at −78 °C. After stirring at the same temperature for 10 min, benzyl alcohol (0.62 mL, 6 mmol) was added and the solution was stirred for an additional 6 h at −78 °C. The reaction mixture was slowly warmed to room temperature and the stirring was continued for 24 h. The reaction was cooled to 0 °C, quenched with 10% aq. citric acid (100 mL), and the volatiles were removed in vacuo. The residue was extracted with ethyl acetate (3×150 mL) and the combined organic phases were washed with 10% aq. citric acid, saturated aq. $NaHCO_3$, water, and brine, dried over Na_2SO_4, filtered, and concentrated in vacuo. The resulting crude product was purified by column chromatography on silica gel (acetone/petroleum ether, 1:99–1:19–1:9) to give a diastereoisomeric mixture of **31** (0.82 g, 45%) as a colorless oil. ^{1}H NMR (300 MHz, [D_6]acetone) $\delta=8.01$ (d, $J=8.6$ Hz, N*H*-Gly), 7.39–7.28 (m, Ar-H OBn), 6.26 (br s, 1H, N*H*-Ala), 5.69 and 5.67 (2×d, $J=9.3$ Hz, 1H, α*H*-Gly), 4.71–4.56 (m, 2H, CH_2-OBn), 4.24–4.19 (m, 1H, α*H*-Ala), 3.74 and 3.73 (2×s, 3H, OCH_3-Gly), 1.42 and 1.41 (2×s, 9H, *t*Bu), 1.36 and 1.35 ppm (2×d, $J=7.2$ Hz, 3H, CH_3-Ala); ^{13}C NMR (75 MHz, [D_6]acetone) $\delta=174.5$ (CO-Ala), 168.9 and 168.8 (CO-Gly), 156.4 (OCONH), 138.7 and 138.6 (1C of OBn), 129.1 (Ar-C), 129.0 (Ar-C), 128.9 (Ar-C), 128.8 (Ar-C), 128.5 (Ar-C), 79.5 (1C-*t*Bu), 77.7 and 77.6 (αC-Gly), 70.7 and 70.6 (CH_2-OBn), 52.8 (OCH_3-Gly), 51.2 (αC-Ala), 28.5 (*t*Bu), 18.2 and 18.0 ppm (CH_3-Ala); HRMS (ESI$^+$): *m/z* calcd for $C_{18}H_{26}N_2O_6$: 367.1863 $[M+H]^+$; found: 367.1860.

N-Boc-L-Alanyl-L-2-*O*-benzylglycine (32)

subtilisin Carlsberg (20 mg) was added to a stirring solution of the diastereoisomeric mixture **31** (400 mg, 1.092 mmol) in DMF/H_2O

© 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

(1:1, 16 mL) at 55 °C. NaOH (0.61 mL, 1 M) was added dropwise to maintain the pH of the reaction at 7.2 (monitored by a pH-meter). Upon completion, the volatiles were removed in vacuo. The residue was dissolved in 5% $NaHCO_3$ and extracted with ethyl acetate (3×50 mL), and the combined organic phases were washed with 5% $NaHCO_3$, 10% aq. citric acid, water, and brine, dried over Na_2SO_4, filtered, and concentrated in vacuo to give the diastereoisomeric mixture (320 mg, recovered). The aqueous phase was acidified with 10% citric acid and extracted with ethyl acetate (3× 50 mL). The combined organic phases were washed with 10% aq. citric acid, water, and brine, dried over Na_2SO_4, filtered, and concentrated in vacuo to get the pure diastereoisomer (L,L) **32** (54 mg, 14%) as a white solid. ^{1}H NMR (300 MHz, CD_3CN) δ=7.56 (d, J=8.6 Hz, N*H*-Gly), 7.36–7.30 (m, Ar-H OBn), 5.67 (br s, 1H, N*H*-Ala), 5.55–5.50 (m, 1H, α*H*-Gly), 4.67–4.53 (m, 2H, CH_2-OBn), 4.10–4.02 (m, 1H, α*H*-Ala), 1.41 (s, 9H, *t*Bu), 1.28 ppm (d, J=7.2 Hz, 3H, CH_3-Ala); ^{13}C NMR (75 MHz, CD_3CN) δ=175.1 (CO-Gly), 169.2 (CO-Ala), 156.6 (OCONH), 138.4 (1C of OBn), 129.3 (Ar-C), 129.1 (Ar-C), 128.8 (Ar-C), 80.2 (1C-*t*Bu), 77.6 (αC-Gly), 70.8 (CH_2-OBn), 51.4 (αC-Ala), 28.5 (*t*Bu), 17.9 ppm (CH_3-Ala); HRMS (ESI$^+$): *m/z* calcd for $C_{17}H_{24}N_2O_6$: 351.1561 $[M+H]^+$, found: 351.1563.

2-*N*,*N*-Diisopropylamino-1,3,2-oxathiaphospholane (34)

DIPEA (664 μL, 3.811 mmol) was slowly added to a stirring solution of *N*,*N*-diisopropyl phosphoramidous dichloride (293 μL, 1.588 mmol) and 2-mercapto ethanol (112 μL, 1.588 mmol) in dry diethyl ether (8 mL) at −20 °C. The reaction mixture was then stirred at the same temperature for 30 min and slowly warmed to room temperature over 2 h. The resultant white precipitate was isolated by filtration under an inert atmosphere, and the filtrate was concentrated (bath temp. ≈15 °C) and dried in vacuo to give crude compound **34** (301 mg, quant.) as a colorless oil. The product was checked for purity by using ^{31}P NMR and was immediately used in the next step without any further purification.

3′-*O*-Benzoyl-5′-(2-thio-1,3,2-oxathiaphospholane)-thymidine (35)

A solution of *S*-ethylthiotetrazole (273 mg, 2.094 mmol) in dry CH_2Cl_2 (5 mL) was added slowly to a stirring suspension of crude **34** (301 mg, 1.588 mmol) and compound **27** (500 mg, 1.444 mmol) in dry CH_2Cl_2 (18 mL) at 0 °C, and the reaction mixture was left stirring at room temperature for 3 h. When the solution became homogeneous, dry elemental sulfur (153 mg) was added and it was left stirring overnight. The mixture was filtered, and the filtrate was concentrated in vacuo. The resulting crude residue was purified by column chromatography on silica gel ($MeOH/CHCl_3$, 0:100–0.5:95.5–1:99) to give product **35** (650 mg, 93%) as a pale yellow foam. ^{1}H NMR (300 MHz, $CDCl_3$) δ=9.09 (br s, 1H, N*H*-Thy), 8.07 (d, J=1.0 Hz, 1H, H-6), 8.05–7.46 (m, 5H, OBz), 6.53–6.46 (m, 1H, H-1′), 4.58–4.54 (m, 1H, H-3′), 4.64–4.44 (m, 4H, H-5′, H-5″ and OCH_2CH_2S), 4.42–4.39 (m, 1H, H-4′), 3.61–3.52 (m, 2H, OCH_2CH_2S), 2.67–2.60 (m, 1H, H-2′), 2.36–2.23 (m, 1H, H-2″), 2.01 ppm (d, J=1.0 Hz, 3H, CH_3); ^{13}C NMR (75 MHz, $CDCl_3$) δ=166.3 (CO-OBz), 164.0 (C-4), 150.6 (C-2), 135.4 and 135.3 (C-6), 133.9 (1C-OBz), 129.9 (Ar-C), 128.7 (Ar-C), 112.1 and 112.0 (C-5), 85.1 and 85.0 (C-1′), 83.2 (d, $^3J_{C,P}$=8.7 Hz, C-4′), 75.5 and 75.4 (C-3′), 70.9 (d, $^3J_{C,P}$=11.3 Hz, OCH_2CH_2S), 68.1 (d, $^2J_{C,P}$=6.6 Hz, C-5′), 37.7 and 37.6 (C-2′), 37.1 (d, $^3J_{C,P}$=16.3 Hz, OCH_2CH_2S), 12.7 and 12.6 ppm (CH_3-Thy); ^{31}P NMR (121 MHz, $CDCl_3$) δ=104.3 and 104.1 ppm; HRMS (ESI$^+$): *m/z* calcd for $C_{19}H_{21}N_2O_6PS_2$: 507.0420 $[M+Na]^+$; found: 507.0421.

Thymidine-5′-*O*-phosphorothioate triethylammonium salt (36)

DBU (0.22 mL, 1.476 mmol) was added dropwise to a stirring solution of compound **35** (650 mg, 1.342 mmol) and 3-hydroxypropionitrile (0.46 mL, 6.708 mmol) in dry CH_2Cl_2 (18 mL), and the reaction mixture was stirred at room temperature for 8 h. It was concentrated in vacuo and the crude residue was purified by flash chromatography on silica gel ($MeOH/CHCl_3$, 1:49–1:9–1:4). The fractions that contained the phosphorothioate monoester [^{31}P NMR (121 MHz, CD_3OD) δ=56.6 ppm; HRMS (ESI): *m/z* calcd for $C_{20}H_{22}N_3O_8PS$: 494.0792 $[M-H]^-$; found: 494.0789] were combined and concentrated in vacuo. The resulting residue was treated with 25% NH_3 in a sealed tube at 55 °C for 16 h. Ammonia was removed in vacuo, and the residue was diluted with water and lyophilized. The crude product was purified by preparative RP-HPLC (50 mmol TEAB in 98% H_2O+2% CH_3CN and 50 mmol TEAB in 50% CH_3CN+50% H_2O). The collected fractions were freeze-dried repeatedly until the mass remained constant to afford compound **36** (497 mg, 72% over two steps) as a white solid. ^{1}H NMR (300 MHz, D_2O) δ=7.80 (d, J=1.0 Hz, 1H, H-6), 6.34 (app t, J=7.0 Hz, 1H, H-1′), 4.60–4.56 (m, 1H, H-3′), 4.20–4.18 (m, 1H, H-4′), 4.12–4.08 (m, 2H, H-5′ and H-5″), 2.36–2.32 (m, 2H, H-2′ and H-2″), 1.92 ppm (d, J=1.0 Hz, 3H, CH_3); ^{13}C NMR (75 MHz, D_2O) δ=166.2 (C-4), 151.4 (C-2), 137.1 (C-6), 111.4 (C-5), 85.2 (d, $^3J_{C,P}$=9.4 Hz, C-4′), 84.7 (C-1′), 71.0 (C-3′), 64.5 (d, $^2J_{C,P}$=5.1 Hz, C-5′), 38.4 (C-2′), 11.3 ppm (CH_3-Thy); ^{31}P NMR (121 MHz, D_2O) δ=50.2 ppm; HRMS (ESI): *m/z* calcd for $C_{10}H_{15}N_2O_7PS$: 337.0265 $[M-H]^-$; found: 337.0262.

Thymidine-5′-O-(*N*-Boc-L-alanyl-D,L-2-aminoglycine methyl ester)-phosphorothioate (37)

Et_3N (0.45 mL, 3.234 mmol) was added to a stirred solution of compound **19** (344 mg, 0.809 mmol) and compound **36** (309 mg, 0.970 mmol) in dry DMF (6 mL). The reaction mixture was heated at 37 °C for 4 h. After cooling to RT, DMF was removed in vacuo, and the resulting residue was dissolved in water and lyophilized. The crude product was purified by preparative RP-HPLC (98% H_2O+2% CH_3CN and 98% CH_3CN+2% H_2O). The collected fractions were freeze-dried repeatedly until the mass remained constant to afford a diastereoisomeric mixture of compound **37** (329 mg, 68%) as a white solid. ^{1}H NMR (500 MHz, CD_3OD) δ=7.74 (s, 1H, H-6), 6.35 (q, J=6.9 Hz, 1H, H-1′), 5.76–5.69 (m, 1H, α*H*-Gly), 4.50–4.46 (m, 1H, H-3′), 4.16–4.08 (m, 4H, H-4′, α*H*-Ala, H-5′ and H-5″), 3.79 and 3.78 (2×s, 3H, OCH_3-Gly), 2.29–2.25 (m, 2H, H-2′ and H-2″), 1.96 and 1.94 (2×s, 3H, CH_3-Thy), 1.47 and 1.46 (2×s, 9H, *t*Bu), 1.35 and 1.34 ppm (2d, J=7.1 Hz, 3H, CH_3-Ala; merged with Et_3N); ^{13}C NMR (125 MHz, CD_3OD) δ=175.4 and 175.3 (CO-Ala), 170.8 and 170.7 (CO-Gly), 166.4 (C-4), 157.6 (OCONH), 152.4 (C-2), 138.0 (C-6), 112.0 and 111.9 (C-5), 87.0 (d, $^3J_{C,P}$=7.6 Hz, C-4′), 86.2 (C-1′), 80.8 and 80.8 (1C-*t*Bu), 72.7 and 72.6 (C-3′), 67.0 (C-5′), 55.0 and 54.9 (αC-Gly), 53.6 (OCH_3-Gly), 51.7 and 51.6 (αC-Ala), 40.8 and 40.7 (C-2′), 28.7 (*t*Bu), 18.2 and 18.0 (CH_3-Ala), 12.7 and 12.6 ppm (CH_3-Thy); ^{31}P NMR (202 MHz, CD_3OD) δ=13.1 ppm; HRMS (ESI): *m/z* calcd for $C_{21}H_{33}N_4O_{12}PS$: 595.1480 $[M-H]^-$, found: 595.1486.

Thymidine-5′-*O*-(L-alanyl-(−)-2-aminoglycine methyl ester)-phosphorothioate (38a) and thymidine-5′-*O*-(L-alanyl-(+)-2-aminoglycine methyl ester)-phosphorothioate (38b)

A similar synthetic protocol as the one used for the synthesis of compound **23** was employed for the synthesis of **38a,b** starting from compound **37** (160 mg, 0.2293 mmol), thioanisole (33 μL, 0.356 mmol), and TFA (1.5 mL) in CH_2Cl_2 (4.5 mL). Semi-preparative

© 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

RP-HPLC (98% H_2O+2% CH_3CN and 98% CH_3CN+2% H_2O) was employed for further separation to obtain the TFA salts of pure diastereoisomers as a ≈1:1 ratio (isomer **38a**: 59.6 mg; isomer **38b**: 60 mg) as white powders (overall yield **38a+38b**: 119.6 mg, 84%). The isolated products were freeze-dried repeatedly until the mass remained constant. Data for compound **38a**: ^{1}H NMR (300 MHz, D_2O) δ=7.71 (d, J=1.0 Hz, 1H, H-6), 6.35 (app t, J=7.0 Hz, 1H, H-1′), 5.60 (d, J=15.7 Hz, 1H, α*H*-Gly) 4.61–4.56 (m, 1H, H-3′), 4.20–4.18 (m, 1H, H-4′), 4.16–4.09 (m, 3H, α*H*-Ala, H-5′ and H-5″), 3.81 (s, 3H, O*CH*$_3$-Gly), 2.42–2.37 (m, 2H, H-2′ and H-2″), 1.94 (d, J=1.0 Hz, 3H, *CH*$_3$-Thy), 1.55 ppm (d, J=7.1 Hz, 3H, *CH*$_3$-Ala); ^{13}C NMR (150 MHz, D_2O) δ=170.0 (d, $^3J_{C,P}$=6.0 Hz, CO-Gly), 169.7 (CO-Ala), 166.3 (C-4), 151.5 (C-2), 137.1 (C-6), 111.5 (C-5), 84.8 (d, $^3J_{C,P}$=9.2 Hz, C-4′), 84.7 (C-1′), 70.8 (C-3′), 65.3 (d, $^2J_{C,P}$=5.4 Hz, C-5′), 53.6 (αC-Gly and OCH$_3$-Gly), 48.7 (αC-Ala), 38.4 (C-2′), 15.9 (CH$_3$-Ala), 11.5 ppm (CH$_3$-Thy); ^{31}P NMR (202 MHz, D_2O) δ=15.1 ppm; $[\alpha]_{589}^{20}=-0.518^\circ$ (c=1 in CH_3OH); HRMS (ESI): *m/z* calcd for $C_{16}H_{25}N_4O_{10}PS$: 495.0956 $[M-H]^-$; found: 495.0956. Data for compound **38b**: ^{1}H NMR (500 MHz, D_2O) δ=7.69 (s, 1H, H-6), 6.36 (app t, J=7.0 Hz, 1H, H-1′), 5.64 (d, J=14.5 Hz, 1H, α*H*-Gly) 4.57–4.54 (m, 1H, H-3′), 4.19–4.16 (m, 1H, H-4′), 4.15–4.04 (m, 3H, α*H*-Ala, H-5′ and H-5″), 3.75 (s, 3H, O*CH*$_3$-Gly), 2.42–2.35 (m, 2H, H-2′ and H-2″), 1.88 (s, 3H, *CH*$_3$-Thy), 1.53 ppm (d, J=7.1 Hz, 3H, *CH*$_3$-Ala); ^{13}C NMR (125 MHz, D_2O) δ=169.8 (d, $^3J_{C,P}$=6.8 Hz, CO-Gly), 169.7 (CO-Ala), 166.2 (C-4), 151.5 (C-2), 137.3 (C-6), 111.4 (C-5), 84.9 (d, $^3J_{C,P}$=8.9 Hz, C-4′), 84.8 (C-1′), 70.6 (C-3′), 65.4 (d, $^2J_{C,P}$=5.3 Hz, C-5′), 53.6 (OCH$_3$-Gly), 53.3 (d, $^2J_{C,P}$=3.0 Hz, αC-Gly), 48.8 (αC-Ala), 38.1 (C-2′), 16.0 (CH$_3$-Ala), 11.5 ppm (CH$_3$-Thy); ^{31}P NMR (202 MHz, D_2O) δ=15.1 ppm; $[\alpha]_{589}^{20}=+0.730^\circ$ (c=1 in CH_3OH); HRMS (ESI): *m/z* calcd for $C_{16}H_{25}N_4O_{10}PS$: 495.0956 $[M-H]^-$; found: 495.0959.

Acknowledgements

We thank FWO (Vlaanderen) (G.078014N and G.08816N) for financial support. 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 no ERC-2012-ADG 20120216/320683. We thank Chantal Biernaux for final typesetting.

Keywords: phosphoaminals · nucleosides · peptides · polymerase chain reaction · synthetic biology

[1] B. S. Vig, K. M. Huttunen, K. Laine, J. Rautio, *Adv. Drug Delivery Rev.* **2013**, *65*, 1370–1385.
[2] M. Lalanne, K. Andrieux, P. Couvreur, *Curr. Med. Chem.* **2009**, *16*, 1391–1399.
[3] D. P. Drontle, C. R. Wagner, *Mini-Rev. Med. Chem.* **2004**, *4*, 409–419.
[4] a) T. W. Abraham, T. I. Kalman, E. J. McIntee, C. R. Wagner, *J. Med. Chem.* **1996**, *39*, 4569–4575; b) J. Kim, T.-f. Chou, G. W. Griesgraber, C. R. Wagner, *Mol. Pharm.* **2004**, *1*, 102–111; c) T. F. Chou, P. Bieganowski, K. Shilinski, J. L. Cheng, C. Brenner, C. R. Wagner, *J. Biol. Chem.* **2005**, *280*, 15356–15361; d) T.-F. Chou, J. Baraniak, R. Kaczmarek, X. Zhou, J. Cheng, B. Ghosh, C. R. Wagner, *Mol. Pharm.* **2007**, *4*, 208–217; e) J. Cheng, X. Zhou, T.-F. Chou, B. Ghosh, B. Liu, C. R. Wagner, *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6379–6381.
[5] a) O. Adelfinskaya, P. Herdewijn, *Angew. Chem. Int. Ed.* **2007**, *46*, 4356–4358; *Angew. Chem.* **2007**, *119*, 4434–4436; b) O. Adelfinskaya, M. Terrazas, M. Froeyen, P. Marliere, K. Nauwelaerts, P. Herdewijn, *Nucleic Acids Res.* **2007**, *35*, 5060–5072.
[6] X.-P. Song, C. Bouillon, E. Lescrinier, P. Herdewijn, *Chem. Biodiversity* **2012**, *9*, 2685–2700.
[7] E. Groaz, P. Herdewijn, *Curr. Med. Chem.* **2015**, *22*, 3980–3990.
[8] P. Herdewijn, P. Marliere, *FEBS Lett.* **2012**, *586*, 2049–2056.
[9] M. Jauker, H. Griesser, C. Richert, *Angew. Chem. Int. Ed.* **2015**, *54*, 14564–14569.
[10] J. G. Zhu, H. Fu, Y. Y. Jiang, Y. F. Zhao, *J. Org. Chem.* **2006**, *71*, 1722–1724.
[11] a) M. B. Hoagland, E. B. Keller, P. C. Zamecnik, *J. Biol. Chem.* **1956**, *218*, 345–358; b) P. Schimmel, *Annu. Rev. Biochem.* **1987**, *56*, 125–158.
[12] a) N. Wickramasinghe, J. C. Lacey, *Origins Life Evol. Biospheres* **1992**, *22*, 361–368; b) R. Kluger, R. W. Loo, V. Mazza, *J. Am. Chem. Soc.* **1997**, *119*, 12089–12094; c) T. Moriguchi, T. Yanagi, T. Wada, M. Sekine, *J. Chem. Soc. Perkin Trans. 1* **1999**, 1859–1865; d) J. P. Biron, L. L. Parkes, R. Pascal, J. D. Sutherland, *Angew. Chem. Int. Ed.* **2005**, *44*, 6731–6734; *Angew. Chem.* **2005**, *117*, 6889–6892.
[13] a) G. Harris, J. W. Davies, R. Parsons, *Nature* **1958**, *182*, 1565–1567; b) G. Harris, J. W. Davies, *Nature* **1959**, *184*, 788–789.
[14] G. Harris, I. C. Macwilliam, *J. Chem. Soc.* **1961**, 2053–2057.
[15] U. Zoller, D. Benishai, *Tetrahedron* **1975**, *31*, 863–866.
[16] a) W. D. Kingsbury, J. C. Boehm, R. J. Mehta, S. F. Grappel, C. Gilvarg, *J. Med. Chem.* **1984**, *27*, 1447–1451; b) G. Apitz, W. Steglich, *Tetrahedron Lett.* **1991**, *32*, 3163–3166; c) G. Apitz, M. Jager, S. Jaroch, M. Kratzel, L. Schaffeler, W. Steglich, *Tetrahedron* **1993**, *49*, 8223–8232.
[17] a) T. Bretschneider, W. Miltz, P. Munster, W. Steglich, *Tetrahedron* **1988**, *44*, 5403–5414; b) J. T. Repine, J. S. Kaltenbronn, A. M. Doherty, J. M. Hamby, R. J. Himmelsbach, B. E. Kornberg, M. D. Taylor, E. A. Lunney, C. Humblet, S. T. Rapundalo, B. L. Batley, M. J. Ryan, C. A. Painchaud, *J. Med. Chem.* **1992**, *35*, 1032–1042.
[18] H. Horikawa, T. Iwasaki, K. Matsumoto, M. Miyoshi, *Tetrahedron Lett.* **1976**, *17*, 191–194.
[19] a) J. L. Kraus, G. Attardo, *Eur. J. Med. Chem.* **1992**, *27*, 19–26; b) V. Niddam, M. Camplo, D. LeNguyen, J. C. Chermann, J. L. Kraus, *Bioorg. Med. Chem. Lett.* **1996**, *6*, 609–614; c) C. Paulitz, W. Steglich, *J. Org. Chem.* **1997**, *62*, 8474–8478; d) L. Yaouancq, M. Anissimova, M. A. Badet-Denisot, B. Badet, *Eur. J. Org. Chem.* **2002**, 3573–3579.
[20] a) W. D. Kingsbury, J. C. Boehm, D. Perry, C. Gilvarg, *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 4573–4576; b) W. D. Kingsbury, J. C. Boehm, *Intl J. Peptide Protein Res.* **2009**, *27*, 659–665.
[21] C. V. Miduturu, S. K. Silverman, *J. Org. Chem.* **2006**, *71*, 5774–5777.
[22] M. Bogenstätter, W. Steglich, *Tetrahedron* **1997**, *53*, 7267–7274.
[23] F. Felluga, W. Baratta, L. Fanfoni, G. Pitacco, P. Rigo, F. Benedetti, *J. Org. Chem.* **2009**, *74*, 3547–3550.
[24] S. De, E. Groaz, L. Margamuljana, M. Abramov, P. Marliere, P. Herdewijn, *Org. Biomol. Chem.* **2015**, *13*, 3950–3962.
[25] C. Ausín, A. Grajkowski, J. Cieslak, S. L. Beaucage, *Org. Lett.* **2005**, *7*, 4201–4204.
[26] S. Iyer, A. C. Hengge, *J. Org. Chem.* **2008**, *73*, 4819–4829.
[27] a) S. De, E. Groaz, P. Herdewijn, *Eur. J. Org. Chem.* **2014**, 2322–2348; b) S. De, E. Groaz, M. Maiti, V. Pezo, P. Marliere, P. Herdewijn, *Tetrahedron* **2014**, *70*, 8843–8851.
[28] a) R. L. Alexander, S. L. Morris-Natschke, K. S. Ishaq, R. A. Fleming, G. L. Kucera, *J. Med. Chem.* **2003**, *46*, 4205–4208; b) G. L. Kucera, C. L. Goff, N. Iyer, S. Morris-Natschke, K. S. Ishaq, S. D. Wyrick, R. A. Fleming, L. S. Kucera, *Antiviral Res.* **2001**, *50*, 129–137; c) N. Schlienger, T. Beltran, C. Perigaud, I. Lefebvre, A. Pompon, A. M. Aubertin, G. Gosselin, J. L. Imbach, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 3003–3006.
[29] Y. Ahmadibeni, K. Parang, *Org. Lett.* **2005**, *7*, 1955–1958.
[30] H. H. Zepik, P. Walde, T. Ishikawa, *Angew. Chem. Int. Ed.* **2008**, *47*, 1323–1325; *Angew. Chem.* **2008**, *120*, 1343–1345.
[31] M. Olesiak, D. Krajewska, E. Wasilewska, D. Korczynski, J. Baraniak, A. Okruszek, W. J. Stec, *Synlett* **2002**, 0967–971.
[32] a) M. Maiti, S. Michielssens, N. Dyubankova, M. Maiti, E. Lescrinier, A. Ceulemans, P. Herdewijn, *Chem. Eur. J.* **2012**, *18*, 857–868; b) M. Trmčić, F. L. Chadbourne, P. M. Brear, P. W. Denny, S. L. Cobb, D. R. W. Hodgson, *Org. Biomol. Chem.* **2013**, *11*, 2660–2675.

Received: February 16, 2016
Published online on May 2, 2016

© 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim