Dipeptides as leaving group in the enzyme catalyzed DNA synthesis

by **Xiao-Ping Song, Camille Bouillon, Eveline Lescrinier and Piet Herdewijn***

Laboratory of Medicinal Chemistry, Rega Institute for Medical Research, KU Leuven,
Minderbroedersstraat 10, 3000 Leuven, Belgium (Phone: Tel. +32 16 337387; Fax: +32
16 337340; E-mail: Piet.Herdewijn@rega.kuleuven.be)

Abstract

Conjugates between 2'-deoxyadenosine monophosphate and dipeptides have been synthesized and tested as substrates for several polymerases. Although the incorporation efficiency is not very high, it demonstrates that some of these dipeptides can be accommodated in the active site of polymerases and function as leaving groups in the enzymatic synthesis of DNA.

Keywords: maximum 5 keywords

Dipeptides; DNA synthesis; polymerases; leaving group; nucleotides

Introduction. - Nucleic acids are synthesized in a cell making use of nucleoside triphosphates as substrate and polymerases as enzymes. Pyrophosphate functions as leaving group in this process. Diversifying this leaving group could serve the goal to develop a new biochemistry for the enzymatic synthesis of nucleic acids. Besides the potential application in drug design (developing of modified nucleoside triphosphate mimics as direct substrates for polymerases), these molecules can also be applied in synthetic biology (developing of an XNA information system) [1], and as tools to better understand the chemical origin of nucleic acid replication in life.

The constructions of an XNA system could rely on a systematic diversification of the sugar moiety, the base moiety and the leaving group as well as enzyme evolution.

Generally, native polymerases have limited ability to recognize modified nucleotide building blocks. However, some engineered polymerase mutants are able to efficiently incorporate unnatural substrates. For example, Pinheiro et al reported that some engineered polymerases are able to generate XNA sequences (HNA, CeNA, ANA, FANA, TNA and LNA) up to 72 nt, while still retaining canonical base-pairing selectivity [2]. Previously, we have demonstrated that some natural amino acids (e.g. L-aspartic acid) or unnatural amino acids (e.g. L- β -imidazole lactic acid, iminodiacetate (IDA), iminodipropionate (IDP) and 3-phosphono-L-alanine) are able to act as pyrophosphate mimics in enzymatic DNA polymerization [3-7]. We also observed that d4TMP adducts of 3-phosphono-L-alanine (3-phosphono-L-alanine-d4TMP) fails to deliver d4TMP into CEM/TK⁻ cells, suggesting the metabolic instability of this nucleotide adduct [8].

In the present study, we have evaluated if dipeptides can function as leaving group in the DNA polymerization reaction. Dipeptides could have an additional advantage as they might contribute to the cell penetration ability of the DNA precursors bearing an alternative leaving group. Dipeptides might facilitate the membrane delivery of these nucleotide adducts with the help of membrane transporters [9]. Dipeptides consisting of two amino acid molecules joined by a peptide bond, are metabolically degradable by hydrolase enzymes *in vivo* [10]. Therefore, the dipeptide that would be generated in the cell after nucleotide incorporation can be degraded into single amino acids, making the polymerization reaction irreversible. Dipeptides bearing two additional carboxylic acid residues could retain the chelating properties necessary for binding the nucleotide substrates in the active site of polymerases. Compared to a single amino acid, a dipeptide is a larger molecule, allowing us to probe the geometry of the enzyme active site pocket.

Result and Discussion. - *Synthesis of compound 1a-1k.* The 11 compounds synthesized are depicted in Table 1. The synthetic routes used for obtaining compound **1a-1k** are outlined in Scheme 1, 2, 3 and 4.

➔ Insert Table 1 here

Scheme 1 shows the synthesis of the phosphoramidate nucleoside dipeptide analogues **1a-1e**, obtained in four steps. First the protected dipeptides **2a-2e** were obtained in good yield starting from commercially available amino acids, using the conventional coupling reagent DCC/HoSu in CH₂Cl₂. The Boc protecting group was removed by treating **2a-2e** with trifluoro acetic acid giving the corresponding amines **3a-3e** in excellent yields. The

synthesis of compounds **1a-1e** was carried out according to the procedure described by Wagner and colleagues [11], starting from deoxyadenosine monophosphate and amines **3a-3e** using DCC as coupling reagent. Finally, deprotection of the carboxylic acid ester function with sodium hydroxide in methanol/water solution (for compound **1a**, **1b** and **1e**) or by hydrogenation with 10 % Pd/C (for compound **1c** and **1d**) provided the target compounds **1a-1e**.

➔ insert Scheme 1 here

The synthesis of compound **1f-1h** was accomplished as described in Scheme 2. N-alkylation of benzyl glycine ethyl ester with *tert*-butyl bromoacetate in the presence of diisopropylethylamine (DIPEA) gave compound **4**, which was further treated with TFA to remove selectively the *tert*-butyl group giving the corresponding acid **5** as TFA salt. Then the obtained compound **5** was coupled to commercially available amino acids, using the coupling reagent DCC/HoSu or EDCi/HOBt to afford compounds **6a-6c** in good yields. Compounds **7a-7c** were obtained after removal of the benzyl protecting group and the resulting secondary amine was coupled with dAMP in the presence of DCC to provide the protected nucleoside analogues. The deprotection of the carboxylic acid function with sodium hydroxide in methanol-water solution gave compound **1f-1h**.

➔ insert Scheme 2 here

The synthesis of compound **1i** and **1j** is described in Scheme 3. Two Michael addition reactions of benzylamine on methyl acrylate and *tert*-butyl acrylate successively gave compound **8** and **9**. Treatment of the latter with TFA provided compound **10**. The trifluoroacetic salt **10** was coupled with glycine methyl ester or L-aspartic acid dimethyl

ester in the presence of EDCi/HOBt to obtain **11a** and **11b** respectively. The hydrogenolysis reaction gave the free amine **12a** and **12b**. Finally, as described previously, a coupling reaction of **12a-12b** with dAMP, followed by a saponification reaction gave the final compounds **1i** and **1j**.

➔ insert Scheme 3 here

Scheme 4 shows the synthesis of compound **1k** obtained in four steps, starting with a coupling reaction between diethyl aminomalonate and carbobenzyloxy glycine. Then compound **13** was engaged in a hydrogenation reaction in the presence of hydrochloric acid to avoid intermolecular and intramolecular side reactions. The chlorhydrate salt **14** was coupled with dAMP and deprotection of the carboxylic acid ester function provides compound **1k**, after a decarboxylation reaction.

➔ insert Scheme 4 here

Single nucleotide incorporation experiments. Polymerases play a central role during the enzyme catalyzed DNA synthesis. They promote DNA synthesis by catalyzing the nucleophilic attack of the 3'-terminal hydroxyl group of a primer to the α -phosphate of the incoming nucleoside triphosphate, leading to formation of a 5'-3' phosphodiester bond [12]. In order to investigate the substrate property of compound **1a-1k**, polymerases from different families were tested (HIV-1 RT, Taq, Vent (exo⁻) polymerase and Terminator DNA polymerase). Primer P₁ and template T₁ were used as duplex (Table 2), deoxyadenosine triphosphate (dATP) was used as positive control. Incorporation efficiency was analyzed by the polyacrylamide gel-based single nucleotide incorporation assay [13-14].

➔ Insert Table 2 here

Single nucleotide incorporation by HIV-1 RT. HIV-1 RT is an error-prone polymerase, therefore has high mutation rate [15]. Its flexibility and high tolerance towards chemical modified nucleotide substrates made it the primary choice as enzyme catalyst for our pyrophosphate mimic screening. The single nucleotide incorporation results of compound **1a-1k** by HIV-1 RT are shown in Table 3. Unfortunately, all 11 compounds are not good substrates for HIV-1 RT. Most of the compounds only showed 10 %-35 % yield of primer extension. Compound **1c**, **1g** and **1h** are not able to extend the primer after 2 hours reaction time. The best single incorporation result was obtained by using compound **1a** as substrate, with 33 % yield of P+1 strand. It's interesting to note that compounds **1a**, **1b** and **1c** which all contain an L-aspartic acid molecule in the leaving group showed quite different incorporation results, with 30 %, 0 % and 20 % primer conversion respectively. This is much less than the 90 % conversion obtained when using L-Asp-dAMP as substrate.

➔ Insert Table 3 here

Single nucleotide incorporation by Therminator DNA polymerase. Therminator DNA polymerase is a 9°N variant, recognized for its ability to incorporate a broad spectrum of modified nucleotides [16-17]. As showed in Figure 1, compound **1a-1k** showed good substrate properties using Therminator as catalyst. Except for compound **1d**, all the tested analogues are accepted as substrate by Therminator DNA polymerase quite efficiently, some moderate to high yield of primer conversion are obtained. However, using Therminator polymerase as catalyst higher ratio of misincorporation was observed.

For example, using compound **1g** as substrate in the single nucleotide incorporation assay over 90 % primer conversion was obtained with Terminator polymerase (Figure 2), but only 22 % of the primer was converted to the P+1 level, 49 % converted to the P+2 level, and 19 % primer converted to the P+3 and P+4 level.

➔ insert Figure 1 here

➔ insert Figure 2 here

Single nucleotide incorporation by other DNA polymerases. Furthermore, the substrate properties of compound **1a-1k** were investigated by Taq and Vent (exo⁻) polymerase, as showed in Figure 3. The incorporation efficiency is significantly less appealing than using Terminator DNA polymerase as catalyst. When using Taq as polymerase, only compound **1d** (30 % primer conversion) and **1g** (27 % primer conversion) give primer extension. Most of the analogues are poor substrates for Taq polymerase. In contrast, Vent (exo⁻) polymerase showed slightly better catalytic ability. Compound **1f** with iminodiacetic acid joint by L-histidine as leaving group demonstrates 75 % primer extension at a concentration of 500 μM (Figure 4). Compound **1g** containing an iminodiacetic acid and a glycine in the leaving group gives 60 % P+1 product formation. On the contrary, compound **1h** composed of iminodiacetic acid and L-aspartic acid in the leaving group was poorer substrate than compound **1f** and **1g**, and only 15 % yield of P+1 product was obtained. The different incorporation results obtained for compound **1f-1g** indicated that it's still challenging to balance binding properties, and potential substrate properties for polymerases of modified nucleoside triphosphates.

➔ insert Figure 3 here

➔ insert Figure 4 here

It is worth to note that compound **1k** is the smallest sized dipeptide among the tested analogues. It contains two glycine molecules and only one carboxylic acid group in the leaving group, and still can be accepted as substrate by all four enzymes tested. It showed better substrate properties than some of the analogues with two or three carboxylic acid groups, especially when using Terminator polymerase as catalyst (Figure 1). Enzyme catalyzed nucleotide incorporation is a very complicated procedure. A lot of factors might be involved in this process. Based on the results obtained for compound **1k** and for other analogues that have been studied here, it is clear that the size of the dipeptide leaving group is a very important factor influencing the incorporation reactions. Co-crystallization experiments and molecular modeling to analyze the possible interactions between the substrate and the enzyme active site will be very helpful for the future rational design of new pyrophosphate mimics.

Conclusion. - A series of phosphoramidate analogues bearing dipeptides as leaving group were synthesized and tested as substrates by HIV-1 RT, Taq, Vent (exo⁻) and Terminator DNA polymerase. Terminator polymerase showed the best catalytic efficiency between all the four used enzymes, with moderate to high yield of primer conversions. Higher level of misincorporation was also observed with Terminator polymerase. Compound **1k** bearing only one carboxylic acid residue in the leaving group is still well accepted by the four polymerases. This suggests that the spacial size of the dipeptide leaving group might play an important role in the recognition by the enzyme active site. Although these dipeptide leaving groups show inferior incorporation

efficiency when compare with, for example, iminodipropionate, the obtained results are encouraging. We will further explore the potential of using such protected or unprotected dipeptide-dNMP analogues as potential prodrugs for modified nucleotides. Another interesting question is, if such conjugates would be able to freely penetrate the membrane of vesicles and/or function as substrate in a non-enzymatic nucleic acid polymerization reaction.

Experimental Part

General Methods. For all reactions, chemicals of analytical or synthetic grade were obtained from commercial sources and were used without purification. Technical solvents were obtained from Brenntag (Deerlijk, Belgium). Flash silica chromatography was performed on Davisil® silica gel 60, 0.040–0.063 mm (Grace Davison). Analytical thin Layer Chromatography was performed on Alugram® silica gel UV254 mesh 60, 0.20 mm (Macherey-Nagel). NMR Spectra were recorded on a Bruker Avance™ II 300 MHz or 500 MHz NMR spectrometer. Chemical shifts are expressed as δ units (part per million) down field from TMS (tetramethyl silane) for ^1H and ^{13}C . ^{31}P NMR chemical shifts are referenced to an external 85 % H_3PO_4 standard ($\delta = 0.00$ ppm). Exact mass spectra were obtained with a quadrupol/orthogonal-acceleration time-of-flight tandem mass spectrometer (Q-ToF 2, Micromass) equipped with a standard electrospray ionization (ESI) source.

Dibenzyl 2-({4-(benzyloxy)-2-[(tert-butoxycarbonyl)amino]-4-oxobutanoyl}amino) succinate (2a), general procedure (I). To a solution of Boc-L-aspartic acid benzyl ester

(646 mg, 2.0 mmol) and L-aspartic acid dibenzyl ester (971 mg, 2.0 mmol) in CH₂Cl₂ (15 mL), N-methylmorpholine (205 μ L, 2.0 mmol), DCC (619 mg, 3.0 mmol) and N-hydroxysuccinimide (232 mg, 2.0 mmol) were added. The reaction mixture was stirred for 14h under argon at room temperature, white precipitate was removed by filtration, the resulting crude material was purified on silica chromatography (n-Hexane/EtoAc 5:1, v/v) to provide compound **2a** as a white solid (yield 92 %). ¹H NMR (300 MHz, CDCl₃): δ 7.45-7.30 (m, 15H, *Ph*), 5.14-5.07 (m, 6H, *CH*₂*Ph*), 4.92-4.86 (m, 1H, Asp α -*CH*), 4.60-4.55 (m, 1H, Asp β -*CH*), 3.11-3.03 (m, 2H, Asp α -*CH*₂), 2.93-2.70 (m, 2H, Asp β -*CH*₂), 1.47 (s, 9H, *tBu*); ¹³C NMR (75 MHz, CDCl₃): δ 171.2, 170.3, 170.1, 169.7, 155.2, 135.1, 134.8, 127.9-128.3 (15C), 81.2, 67.2, 66.5 (2C), 50.3, 48.6, 36.0, 35.8, 27.9; HRMS (ESI): *m/z* calcd. for (C₃₄H₃₈N₂O₉Na⁺) 641.2470, found 641.2450.

Benzyl 4-({2-(benzyloxy)-2-oxoethyl}amino)-3-[(tert-butoxycarbonyl)amino]-4-oxobutanoate (2b). The general procedure (I) was applied using glycine benzyl ester (404 mg, 2.0 mmol), Boc-L-aspartic acid benzyl ester (646 mg, 2.0 mmol), N-methylmorpholine (205 μ L, 2.0 mmol), DCC (619 mg, 3.0 mmol) and N-hydroxysuccinimide (232 mg, 2.0 mmol). Purification was carried out according to the method used for compound **2a**, the title compound was obtained as white solid (yield 82 %). ¹H NMR (300 MHz, CDCl₃): δ 7.37-7.35 (m, 10H, *Ph*), 5.19 (s, 2H, *CH*₂-*Ph*), 5.15 (s, 2H, *CH*₂-*Ph*), 4.61-4.60 (m, 1H, NH*CHCH*₂), 4.06 (d, 2H, *J*=5.0 Hz, NH*CH*₂CO₂Bn), 3.12-3.05 (m, 1H, NH*CHCH*_{2a}), 2.80-2.72 (m, 1H, NH*CHCH*_{2b}), 1.47 (s, 9H, *tBu*); ¹³C NMR (75 MHz, CDCl₃): δ 171.7, 171.1, 169.4, 155.6, 135.5, 135.3, 128.4-128.7 (10C), 80.7, 67.3, 66.9, 50.6, 36.2, 34.0, 28.4; HRMS (ESI): *m/z* calcd. for (C₂₅H₃₂N₂O₇⁺) 471.2126, found 471.2112.

Methyl 3-[(tert-butoxycarbonyl)amino]-4-{[1-(1H-imidazol-4-ylmethyl)-2-methoxy-2-oxoethyl] amino}-4-oxobutanoate (2c). The general procedure (I) was applied using L-histidine methyl ester dihydrochloride (484 mg, 2.0 mmol), Boc-L-aspartic acid methyl ester (495 mg, 2.0 mmol), N-methylmorpholine (205 μ L, 2.0 mmol), DCC (619 mg, 3.0 mmol) and N-hydroxysuccinimide (232 mg, 2.0 mmol). Purification was carried out according to the method used for compound **2a**, the title compound was obtained as a colorless oil (yield 58 %). ^1H NMR (300 MHz, CDCl_3): δ 7.72 (s, 1H, $\text{NHCH}=\text{N}$), 6.79 (s, 1H, $\text{NHCH}=\text{C}$), 4.72-4.70 (m, 1H, $\text{NHCHCH}_2\text{CO}_2\text{Me}$), 4.51-4.50 (m, 1H, NHCHCH_2Im), 3.65 (s, 3H, OCH_3), 3.63 (s, 3H, OCH_3), 3.12-3.10 (m, 2H, CH_2Im), 2.88-2.73 (m, 2H, $\text{NHCHCH}_2\text{CO}_2\text{Me}$), 1.39 (s, 9H, *tBu*); ^{13}C NMR (75 MHz, CDCl_3): δ 171.7, 170.9, 170.6, 155.3, 135.0, 131.1, 118.7, 80.1, 52.6, 52.1, 51.7, 50.6, 36.0, 29.3, 27.9; HRMS (ESI): *m/z* calcd. for ($\text{C}_{17}\text{H}_{27}\text{N}_4\text{O}_7^+$) 399.1880, found 399.1873.

Dimethyl 2-({2-[(tert-butoxycarbonyl)amino]-3-{(1H-imidazol-4-yl)propanoyl}amino} succinate (2d). The general procedure (I) was applied using Boc-L-histidine (510 mg, 2.0 mmol), Boc-L-aspartic acid dimethyl ester hydrochloride (395 mg, 2.0 mmol), N-methylmorpholine (205 μ L, 2.0 mmol), DCC (619 mg, 3.0 mmol) and N-hydroxysuccinimide (232 mg, 2.0 mmol). Purification was carried out according to the method used for compound **2a**, the title compound was obtained as a colorless oil (yield 74 %). ^1H NMR (300 MHz, CDCl_3): δ 7.77 (s, 1H, $\text{NHCH}=\text{N}$), 6.79 (s, 1H, $\text{NHCH}=\text{C}$), 4.76-4.74 (m, 1H, NHCHCO_2Me), 4.37-4.35 (m, 1H, NHCHCH_2Im), 3.63 (s, 3H, OCH_3), 3.57 (s, 3H, OCH_3), 2.99-2.82 (m, 2H, CH_2Im), 2.81-2.60 (m, 2H, $\text{NHCHCH}_2\text{CO}_2\text{Me}$), 1.31 (s, 9H, *tBu*); ^{13}C NMR (75 MHz, CDCl_3): δ 174.7, 171.8,

170.9, 155.5, 135.0, 132.1, 118.0, 79.9, 54.3, 52.6, 51.9, 48.5, 35.8, 29.2, 28.1; HRMS (ESI): m/z calcd. for ($C_{17}H_{27}N_4O_7^+$) 399.1880, found 399.1871.

Dibenzyl 2-[(tert-butoxycarbonyl)amino]acetyl}amino]succinate (2e). The general procedure (I) was applied using Boc-L-glycine (350 mg, 2.0 mmol), L-aspartic acid dibenzyl ester (971 mg, 2.0 mmol), N-methylmorpholine (205 μ L, 2.0 mmol), DCC (619 mg, 3.0 mmol) and N-hydroxysuccinimide (232 mg, 2.0 mmol) were added.

Purification was carried out according to the method used for compound **2a**. The title compound was obtained as white solid (yield 95 %). 1H NMR (300 MHz, $CDCl_3$): δ 7.35-7.11 (m, 10H, *Ph*), 5.19 (s, 2H, CH_2Ph), 5.06 (s, 2H, CH_2Ph), 4.95-4.92 (m, 1H, $NHCHCO_2Bn$), 3.82 (s, 2H, $NHCH_2CO$), 3.13-2.82 (m, 2H, $NHCHCH_2CO_2Bn$), 1.46 (s, 9H, *tBu*); ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.3, 169.9, 169.1, 155.6, 135.0, 134.8, 127.9-128.3 (10C), 79.9, 67.3, 66.5, 48.3, 43.8, 36.0, 28.0; HRMS (ESI): m/z calcd. for ($C_{25}H_{31}N_2O_7 Na^+$) 493.1945, found 493.1957.

Dibenzyl 2-[[2-amino-4-(benzyloxy)-4-oxobutanoyl]amino]succinate trifluoroacetate (3a), general procedure (II). Compound **2a** was dissolved in CH_2Cl_2 (10 mL), TFA (5 mL) was added to the solution. Reaction solution was stirred for 24h at room temperature. After removal of solvent and excess of TFA, the desired product **3a** was obtained as TFA salt as colorless oil (yield 98 %). 1H NMR (300 MHz, $CDCl_3$): δ 7.30-7.22 (m, 15H, *Ph*), 5.06-4.99 (m, 6H, $PhCH_2$), 4.56-4.50 (m, 1H, $Asp\alpha-CH$), 4.09-4.05 (m, 1H, $Asp\beta-CH$), 3.01-2.88 (m, 4H, $Asp\alpha-CH_2+Asp\beta-CH_2$); ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.4, 170.4, 169.4, 135.0, 134.7, 134.6, 127.9-128.2 (15C), 67.2, 66.6, 64.0,

49.3, 48.7, 35.2, 34.6; HRMS (ESI): m/z calcd. for ($C_{29}H_{32}N_2O_7^+$) 519.2126, found 519.2131.

Benzyl 3-amino-4-{[2-(benzyloxy)-2-oxoethyl]amino}-4-oxobutanoate trifluoroacetate (3b). The general procedure (II) was applied using compound **2b** (1.4 g, 3 mmol) and TFA (5 mL). The desired product was obtained as TFA salt as colorless oil (yield 98 %). 1H NMR (300 MHz, $CDCl_3$): δ 7.31-7.30 (m, 10H, *Ph*), 5.11-5.09 (m, 4H, *PhCH*₂), 4.66-4.65 (m, 1H, *NH*₂*CHCH*₂), 4.00-3.99 (m, 2H, *NHCH*₂*CO*₂*Bn*), 3.04-3.03 (m, 2H, *NH*₂*CHCH*₂); ^{13}C NMR (75 MHz, $CDCl_3$): δ 170.9, 169.5, 168.6, 134.9, 134.6, 128.5-128.8 (10C), 68.1, 67.8, 50.2, 34.7, 33.2; HRMS (ESI): m/z calcd. for ($C_{20}H_{24}N_2O_5^+$) 371.1601, found 371.1592.

*Methyl 3-amino-4-{[1-(1*H*-imidazol-4-ylmethyl)-2-methoxy-2-oxoethyl]amino}-4-oxobutanoate trifluoroacetate (3c)*. The general procedure (II) was applied using compound **2c** (1.2 g, 3 mmol) and TFA (5 mL). The desired product was obtained as TFA salt as colorless oil (yield 95 %). 1H NMR (300 MHz, $CDCl_3$): δ 7.60 (s, 1H, *NHCH=N*), 6.77 (s, 1H, *NHCH=C*), 4.08-4.03 (m, 1H, *NHCHCH*₂*CO*₂*Me*), 3.87-3.83 (m, 1H, *NHCHCH*₂*Im*), 3.66 (s, 3H, *OCH*₃), 3.61 (s, 3H, *OCH*₃), 3.53-3.30 (m, 2H, *NHCHCH*₂*CO*₂*Me*), 2.56-2.43 (m, 2H, *NHCHCH*₂*Im*); ^{13}C NMR (75 MHz, $CDCl_3$): δ 172.9, 171.6, 169.9, 134.4, 133.2, 117.6, 55.3, 54.3, 52.2, 51.5, 36.6, 28.9; HRMS (ESI): m/z calcd. for ($C_{12}H_{19}N_4O_5^+$) 299.1355, found 299.1346.

*Dimethyl 2-{[2-amino-3-(1*H*-imidazol-4-yl)propanoyl]amino} succinate trifluoroacetate (3d)*. The general procedure (II) was applied using compound **2d** (1.2 g, 3 mmol) and TFA (5 mL). The desired product was obtained as TFA salt as colorless oil (yield 90 %).

^1H NMR (300 MHz, CDCl_3): δ 7.64 (s, 1H, $\text{NHCH}=\text{N}$), 6.90 (s, 1H, $\text{NHCH}=\text{C}$), 4.79-4.76 (m, 1H, $\text{NHCHCH}_2\text{CO}_2\text{Me}$), 4.30-4.26 (m, 1H, NHCHCH_2Im), 3.65 (s, 3H, OCH_3), 3.60 (s, 3H, OCH_3), 3.15-3.06 (m, 2H, $\text{NHCHCH}_2\text{CO}_2\text{Me}$), 2.85-2.80 (m, 2H, NHCHCH_2Im); ^{13}C NMR (75 MHz, CDCl_3): δ 174.9, 172.4, 169.3, 136.5, 134.2, 119.5, 56.3, 53.2, 52.7, 52.6, 38.7, 32.6; HRMS (ESI): m/z calcd. for $(\text{C}_{12}\text{H}_{19}\text{N}_4\text{O}_5^+)$ 299.1355, found 299.1350.

Dibenzyl 2-[(aminoacetyl)amino]succinate trifluoroacetate (3e). The general procedure (II) was applied using compound **2e** (1.5 g, 3 mmol) and TFA (5 mL). The desired product was obtained as TFA salt as colorless oil (yield 90 %). ^1H NMR (300 MHz, CDCl_3): δ 7.28-7.21 (m, 10H, *Ph*), 5.00-4.97 (m, 5H, $\text{PhCH}_2+\text{NHCHCH}_2$), 3.95-3.86 (m, 2H, NHCH_2CO), 3.02-2.85 (m, 2H, NHCHCH_2); ^{13}C NMR (75 MHz, CDCl_3): δ 170.5, 170.2, 134.9, 134.6, 127.9-128.2 (10C), 67.5, 66.7, 48.7, 40.7, 35.5; HRMS (ESI): m/z calcd. for $(\text{C}_{20}\text{H}_{24}\text{N}_2\text{O}_5^+)$ 371.1601, found 371.1609.

O-(2'-deoxyadenosin-5'-yl)-*N*-[1-({1,2-dicarboxyethyl}amino)-1-carbonyl-2-carboxyethyl] phosphoramidate (**1a**), general procedure (III). 2'-Deoxyadenosine-5'-monophosphate (dAMP) (100 mg, 0.30 mmol) and compound **3a** (777 mg, 1.5 mmol) were dissolved in a mixture of 1, 4-dioxane (8 mL) and H_2O (2 mL) in the presence of one drop of triethylamine. A solution of DCC (433 mg, 2.10 mmol) in 1, 4-dioxane (2 mL) was added to the reaction mixture, reaction was heated at 85 °C for 3 hours under argon atmosphere. After completion, the reaction mixture was cooled down, solvent was removed under reduced pressure. The resulting residue was purified by chromatography on silica gel ($\text{CHCl}_3/\text{MeOH}/\text{H}_2\text{O}$ 5:2:0.5, v/v/v) to provide compound *O*-

(2'-deoxyadenosin-5'-yl)-*N*-[1-({[3-(benzyloxy)-1-[(benzyloxy)carbonyl]-3-oxopropyl]amino}carbonyl)-3-(benzyloxy)-3-oxopropyl] phosphoramidate as a white solid (yield 43 %). ¹H NMR (300 MHz, D₂O): δ 8.47 (s, 1H, H₈), 8.17 (s, 1H, H₂), 7.29-7.24 (m, 15H, *Ph*), 6.47 (t, 1H, *J* = 6.5 Hz, H_{1'}), 5.04-4.99 (m, 6H, CH₂Ph), 4.64-4.60 (m, 1H, H_{3'}), 4.12-4.10 (m, 1H, H_{4'}), 4.02-3.95 (m, 2H, H_{5'}), 3.47-3.44 (m, 2H, Aspα-CH + Aspβ-CH), 2.87-2.75 (m, 5H, H_{2'a} + Aspα-CH₂ + Aspβ-CH₂), 2.50-2.45 (m, 1H, H_{2'b}); ¹³C NMR (125 MHz, D₂O): 173.8, 171.2, 170.2, 170.0, 158.1, 155.5, 152.1, 148.7, 139.3, 135.6, 135.5, 135.2, 127.4-127.8 (15C), 118.6, 86.2, 83.6, 71.2, 66.6, 66.0, 65.6, 64.0, 52.0, 48.6, 39.6, 37.2, 35.2; ³¹P NMR (121 MHz, D₂O) δ 5.70; HRMS (ESI): *m/z* calcd. for (C₃₉H₄₀N₇O₁₂P⁺) 830.2556, found 830.2591. This intermediate was dissolved in methanol (5 mL) and reduced under H₂ atmosphere for 24 hours, in the presence of 10 % Pd/C (20 mg). The reaction mixture was filtered on Celite, and concentrated under reduced, the resulting crude material was purified by chromatography on silica gel (gradient: CHCl₃/MeOH/H₂O 5:2:0.5, v/v/v; 5:2:1, v/v/v; 5:3:0.5, v/v/v) to provide **1a** as a white solid (64 mg, yield 85 %). ¹H NMR (600 MHz, D₂O): δ 8.45 (s, 1H, H₈), 8.22 (s, 1H, H₂), 6.49 (t, 1H, *J* = 6.6 Hz, H_{1'}), 4.68-4.64 (m, 1H, H_{3'}), 4.34-4.31 (m, 1H, H_{4'}), 4.24-4.22 (m, 1H, Aspα-CH), 3.99-3.90 (m, 2H, H_{5'}), 3.81-3.77 (m, 1H, Aspβ-CH), 2.83-2.78 (m, 1H, H_{2'a}), 2.63-2.44 (m, 5H, H_{2'b} + Aspα-CH₂ + Aspβ-CH₂); ¹³C NMR (150 MHz, D₂O): δ 178.5, 178.3, 178.1, 175.7, 155.5, 152.6, 148.7, 139.9, 118.6, 86.1, 83.6, 71.4, 64.2, 53.1, 52.7, 40.3, 39.4, 38.9; ³¹P NMR (121 MHz, D₂O) δ 6.28; HRMS (ESI): *m/z* calcd. for (C₁₈H₂₂N₇O₁₂P⁺) 560.1148, found 560.1154.

O-(2'-deoxyadenosin-5'-yl)-*N*-[1-({1-carboxymethyl}amino)-1-carbonyl-2-carboxyethyl] phosphoramidate (**1b**). The general procedure (III) was applied using 2'-

deoxyadenosine-5'-monophosphate (dAMP) (100 mg, 0.30 mmol), compound **3b** (555 mg, 1.5 mmol) and DCC (433 mg, 2.10 mmol). After purification the desired product was obtained as white solid (yield 76 %). ¹H NMR (600 MHz, D₂O): δ 8.33 (s, 1H, H₈), 8.09 (s, 1H, H₂), 7.33-7.12 (m, 10H, *Ph*), 6.35 (t, 1H, *J* = 7.2 Hz, H_{1'}), 5.08 (s, 2H, *PhCH*₂), 4.88 (s, 2H, *PhCH*₂), 4.64-4.63 (m, 1H, H_{3'}), 4.18-4.17 (m, 2H, H_{4'} + *NHCHCH*₂CO₂Bn), 3.95-3.94 (m, 2H, H_{5'}), 3.88 (d, 2H, *J* = 2.4 Hz, *NHCH*₂CO₂Bn), 2.67-2.64 (m, 3H, H_{2'a} + *NHCHCH*₂CO₂Bn), 2.63-2.61 (m, 1H, H_{2'b}); ¹³C NMR (150 MHz, D₂O): δ 175.3, 172.2, 170.9, 154.8, 151.8, 148.3, 139.7, 135.1, 135.1, 127.5-128.7 (10C), 118.5, 85.8, 83.7, 70.9, 67.4, 63.9, 52.1, 46.7, 41.4, 39.1, 37.9; ³¹P NMR (121 MHz, D₂O) δ 7.81; HRMS (ESI): *m/z* calcd. for (C₃₀H₂₉N₇O₉P)⁻ 682.2032, found 682.2038. The intermediate was dissolved in methanol (5 mL) and reduced under H₂ atmosphere for 24 hours, in the presence of 10 % Pd/C (22 mg). The reaction mixture was filtered on Celite, and concentrated under reduced pressure, the resulting crude material was purified by chromatography on silica gel (gradient: CHCl₃/MeOH/H₂O 5:2:0.5, v/v/v; 5:2:1, v/v/v; 5:3:0.5, v/v/v) to provide **1b** as a white solid (yield 88 %). ¹H NMR (600 MHz, D₂O): δ 8.45 (s, 1H, H₈), 8.24 (s, 1H, H₂), 6.49 (t, 1H, *J* = 8.4 Hz, H_{1'}), 4.70-4.68 (m, 1H, H_{3'}), 4.27-4.24 (m, 1H, H_{4'}), 4.02-3.91 (m, 2H, H_{5'}), 3.80-3.75 (m, 1H, *NHCHCH*₂CO₂H), 2.85-2.79 (m, 1H, H_{2'a}), 2.61-2.45 (m, 3H, H_{2'b} + *NHCHCH*₂CO₂H); ¹³C NMR (150 MHz, D₂O): δ 178.6, 176.4, 175.9, 155.6, 152.8, 148.8, 139.9, 118.7, 86.1, 83.8, 71.5, 64.2, 53.2, 43.4, 40.6, 39.0; ³¹P NMR (121 MHz, D₂O) δ 6.21; HRMS (ESI): *m/z* calcd. for (C₁₆H₂₁N₇O₁₀P)⁻ 502.1093, found 502.1074.

O-(2'-deoxyadenosin-5'-yl)-*N*-[1-({1-carboxy-2-(1*H*-imidazol-4-yl)-ethyl}amino)-1-carbonyl-2-carboxyethyl] phosphoramidate (**1c**). The general procedure (III) was

applied using 2'-deoxyadenosine-5'-monophosphate (dAMP) (100 mg, 0.30 mmol), compound **3c** (447 mg, 1.5 mmol) and DCC (433 mg, 2.10 mmol). After purification the desired product was obtained as white solid. The resulting crude product was treated with 0.4 M sodium hydroxide solution in MeOH/H₂O (v/v, 1:4) (5 mL) for 3 hours, solvent was removed under reduced pressure, the resulting crude product was purified by chromatography on silica gel (gradient: CHCl₃/MeOH/H₂O 5:2:0.5, v/v/v; 5:2:1, v/v/v; 5:3:0.5, v/v/v) to provide compound **1c** as a white solid (yield 22 % for two steps). ¹H NMR (600 MHz, D₂O): δ 8.41 (s, 1H, H₈), 8.14 (s, 1H, H₂), 7.51 (s, 1H, NHCH=N), 6.73 (s, 1H, NHCH=C), 6.41 (t, 1H, *J* = 6.7 Hz, H_{1'}), 4.60-4.58 (m, 1H, H_{3'}), 4.31-4.27 (m, 1H, CH-His), 4.17-4.15 (m, 1H, H_{4'}), 3.90-3.82 (m, 2H, H_{5'}), 3.75-3.67 (m, 1H, CH-Asp), 2.97-2.86 (m, 2H, CH₂-His), 2.83-2.68 (m, 1H, H_{2'a}), 2.52-2.42 (m, 3H, CH₂-Asp+H_{2'b}); ¹³C NMR (150 MHz, D₂O): δ 179.3, 172.6, 170.7, 155.2, 153.5, 146.8, 139.4, 136.5, 134.9, 119.2, 118.6, 89.4, 84.3, 72.6, 65.2, 56.1, 43.2, 40.9, 39.8, 32.4; ³¹P NMR (121 MHz, D₂O) δ 6.61; HRMS (ESI): *m/z* calcd. for (C₁₉H₂₅N₉O₁₀P⁺) 582.1467, found 582.1471.

O-(2'-deoxyadenosin-5'-yl)-*N*-[1-(1,2-dicarboxyethyl)amino]-1-carbonyl-2-(1H-imidazol-4-yl)-ethyl] phosphoramidate (**1d**). The general procedure (III) was applied using 2'-deoxyadenosine-5'-monophosphate (dAMP) (100 mg, 0.30 mmol), compound **3d** (447 mg, 1.5 mmol) and DCC (433 mg, 2.10 mmol). After purification the desired product was obtained as white solid, the resulting crude product was then treated with 0.4 M sodium hydroxide solution in MeOH/H₂O (v/v, 1:4) (5 mL) for 3 hours. Solvent was removed under reduced pressure, and the resulting crude product was purified by chromatography on silica gel (gradient: CHCl₃/MeOH/H₂O 5:2:0.5, v/v/v; 5:2:1, v/v/v;

5:3:0.5, v/v/v) to provide compound **1d** as a white solid (yield 30 % for two steps). ¹H NMR (600 MHz, D₂O): δ 8.43 (s, 1H, H₈), 8.13 (s, 1H, H₂), 7.45 (s, 1H, NHCH=N), 6.70 (s, 1H, NHCH=C), 6.41 (t, 1H, J = 7.1 Hz, H_{1'}), 4.58-4.56 (m, 1H, H_{3'}), 4.28-4.26 (m, 1H, NHCHCO₂H), 4.15-4.13 (m, 1H, H_{4'}), 3.77-3.71 (m, 3H, H_{5'} + NHCHCH₂Im), 2.50-2.46 (m, 3H, H_{2a'} + NHCHCH₂CO₂H), 2.61-2.59 (m, 3H, H_{2b'} + NHCHCH₂Im); ¹³C NMR (150 MHz, D₂O): δ 177.4, 176.9, 170.6, 155.3, 152.1, 148.9, 140.3, 139.6, 132.0, 118.8, 116.7, 89.8, 87.4, 70.2, 63.0, 52.2, 51.8, 38.4, 34.9, 32.1; ³¹P NMR (121 MHz, D₂O) δ 6.02; HRMS (ESI): m/z calcd. for (C₂₀H₂₅N₉O₁₀P⁻) 582.1467, found 582.1475.

O-(2'-deoxyadenosin-5'-yl)-N-[2-(1,2-dicarboxyethyl)amino]-2-oxoethyl] phosphoramidate (**1e**). The general procedure (III) was applied using 2'-deoxyadenosine-5'-monophosphate (dAMP) (100 mg, 0.30 mmol), compound **3e** (555 mg, 1.5 mmol) and DCC (433 mg, 2.10 mmol). After purification the desired crude product was dissolved in methanol (10 mL) and reduced under H₂ atmosphere for 24 hours in the presence of 10 % Pd/C (22 mg). The reaction mixture was filtered on Celite, and concentrated under reduced pressure, the resulting crude material was purified by chromatography on silica gel (gradient: CHCl₃/MeOH/H₂O 5:2:0.5, v/v/v; 5:2:1, v/v/v; 5:3:0.5, v/v/v) to provide **1e** as a white solid (yield 49 % for two steps). ¹H NMR (600 MHz, D₂O): δ 8.43 (s, 1H, H₈), 8.22 (s, 1H, H₂), 6.47 (t, 1H, J = 6.6 Hz, H_{1'}), 4.67-4.66 (m, 1H, H_{3'}), 4.33-4.31 (m, 1H, NHCHCO₂H), 4.23-4.22 (m, 1H, H_{4'}), 3.97-3.95 (m, 2H, H_{5'}), 3.42-3.39 (m, 2H, NHCH₂CONH), 2.83-2.78 (m, 1H, H_{2'a}), 2.64-2.53 (m, 3H, H_{2'b} + NHCHCH₂CO₂H); ¹³C NMR (150 MHz, D₂O): δ 178.8, 178.2, 173.4, 155.6, 152.7,

148.8, 139.8, 118.6, 85.9, 83.6, 71.2, 64.2, 52.5, 44.4, 39.5, 38.9; ^{31}P NMR (121 MHz, D_2O) δ 7.81; HRMS (ESI): m/z calcd. for $(\text{C}_{16}\text{H}_{20}\text{N}_7\text{O}_{10}\text{P}^-)$ 502.1093, found 502.1090.

N-Benzyl-*N'*-(2-(*tert*-butoxy)-2-oxyethyl)glycine ethyl ester (**4**). To a solution of *N*-benzyl glycine ethyl ester (2.06 g, 10.7 mmol) and diisopropylethylamine (3.5 mL, 21.3 mmol) in anhydrous acetonitrile (10 mL), *tert*-butyl bromoacetate (1.55 mL, 10.7 mmol) was added dropwise. The reaction mixture was stirred for 30 min at room temperature and refluxed for 4 h, solvent was then removed in *vacuo*. The residue was diluted with EtOAc (10 mL), washed with 10 % NaHCO_3 solution (2 x 10 mL), the organic layer dried over anhydrous Na_2SO_4 , and concentrated under reduce pressure. The resulting crude material was purified by flash chromatography (Hexane/EtOAc 5:1, v/v), to provide compound **4** as a yellow oil (yield 87 %). ^1H NMR (300 MHz, CDCl_3): δ 7.42-7.27 (m, 5H, *Ph*), 4.18 (q, 6H, $J = 7.2$ Hz, OCH_2CH_3), 3.93 (s, 2H, NCH_2Ph), 3.55 (s, 2H, $\text{NCH}_2\text{CO}_2\text{Et}$), 3.45 (s, 2H, $\text{NCH}_2\text{CO}_2t\text{Bu}$), 1.49 (s, 9H, *t*Bu), 1.29 (t, 3H, $J = 7.2$ Hz, OCH_2CH_3); ^{13}C NMR (125 MHz, CDCl_3): δ 171.3, 170.5, 138.4, 129.0, 128.3, 127.2, 81.0, 60.3, 57.7, 55.1, 54.3, 28.2, 14.2; HRMS (ESI): m/z calcd. for $(\text{C}_{17}\text{H}_{26}\text{NO}_4^+)$ 308.3926, found 308.1837.

2-[benzyl-(2-ethoxy-2-oxoethyl)amino]acetic acid trifluoroacetate (**5**). The *tert*-butyl ester **4** (2.84 g, 9.2 mmol) was treated by TFA (6.8 mL, 92.3 mmol) in CH_2Cl_2 (60 mL) for 2 days at room temperature. After removal of solvent and excess of TFA, the oily crude residue was triturated with Et_2O and filtered, affording **5** as a white solid of a TFA salt (yield 78 %). ^1H NMR (300 MHz, MeOD): δ 7.52-7.44 (m, 5H, *Ph*), 4.42 (s, 2H, NCH_2Ph), 4.23 (q, 2H, $J = 7.1$ Hz, OCH_2CH_3), 4.07 (s, 2H, $\text{NCH}_2\text{CO}_2\text{Et}$), 4.02 (s, 2H,

$\text{NCH}_2\text{CO}_2\text{H}$), 1.27 (t, 3H, $J = 7.1$ Hz, OCH_2CH_3); ^{13}C NMR (75 MHz, MeOD): 170.6, 168.5, 139.8, 132.3, 131.0, 130.3, 63.4, 60.3, 54.9, 54.6, 14.3; HRMS (ESI): m/z calcd for $(\text{C}_{13}\text{H}_{18}\text{NO}_4^+)$ 252.1236, found 252.1229.

Methyl 2-[[2-[benzyl(2ethoxy-2-oxoethyl)amino]acetyl]amino]-3-(1H-imidazol-4-yl)propanoate (6a), general procedure (IV). A solution of **5** (TFA salt) (500.0 mg, 1.4 mmol), L-histidine methyl ester dihydrochloride (331.4 mg, 1.4 mmol) and triethylamine (630 μL , 4.5 mmol) in dichloromethane (5 mL) was stirred at room temperature for 30 min. The reaction mixture was cooled to 0 $^\circ\text{C}$, 1-hydroxybenzotriazole (278.5 mg, 2.1 mmol) and N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (393.6 mg, 2.1 mmol) was then added. The reaction solution was stirred for 24 hours at room temperature. After removal of solvent, the obtained residue was diluted with EtOAc (20 mL), washed with saturated NaHCO_3 solution (2×10 mL) and brine (10 mL), dried over anhydrous MgSO_4 , and concentrated *in vacuo*. The resulting crude material was purified by chromatography on silica gel ($\text{CH}_2\text{Cl}_2/\text{MeOH}$ 10:1, v/v) to provide **6a** as a white solid (431.0 mg, yield 75 %). ^1H NMR (300 MHz, CDCl_3): δ 9.77 (broad s, 1H, *NH*), 8.39 (d, 1H, *NH*), 7.49 (s, 1H, $\text{NHCH}=\text{N}$), 7.26-7.18 (m, 5H, *Ph*), 6.78 (s, 1H, $\text{NHCH}=\text{C}$), 4.85-4.78 (m, 1H, NHCHCO_2Me), 4.10 (q, 2H, $J = 7.1$ Hz, OCH_2CH_3), 3.82-3.69 (m, 2H, NCH_2CONH), 3.61 (s, 3H, CO_2CH_3), 3.31 (s, 4H, $\text{NCH}_2\text{Ph} + \text{NCH}_2\text{CO}_2\text{Et}$), 3.14-3.11 (m, 2H, CH_2Im), 1.19 (t, 3H, OCH_2CH_3); ^{13}C NMR (75 MHz, CDCl_3): 171.5, 170.9, 170.8, 137.1, 135.1, 132.9, 128.9, 128.2, 127.4, 116.7, 60.4, 58.6, 57.5, 54.1, 52.0, 51.9, 29.2, 13.9; HRMS (ESI): m/z calcd for $(\text{C}_{20}\text{H}_{19}\text{N}_4\text{O}_5^+)$ 403.1981, found 403.1981.

Methyl ({2-(benzyl[2ethoxy-2-oxoethyl]amino)acetyl}amino)acetate (6b). The general procedure (I) was applied using glycine methyl ester (279 mg, 2.0 mmol), compound **5** (502 mg, 2.0 mmol), N-methylmorpholine (205 μ L, 2.0 mmol), DCC (619 mg, 3.0 mmol) and N-hydroxysuccinimide (232 mg, 2.0 mmol). Purification was carried out according to the method used for compound **2a**. The title compound was obtained as white solid (yield 45 %). ^1H NMR (300 MHz, CDCl_3): δ 7.25-7.40 (m, 5H, *Ph*), 4.17-4.12 (m, 2H, OCH_2CH_3), 3.90 (d, 2H, $J=12.4$ Hz, $\text{CH}_2\text{-Gly}$), 3.72 (s, 3H, OCH_3), 3.32-3.49 (m, 4H, NCH_2CO), 1.29-1.22 (t, 3H, $J=7.4$ Hz, OCH_2CH_3); ^{13}C NMR (75 MHz, CDCl_3): δ 170.5, 169.0, 166.2, 137.4, 129.0, 128.3, 127.5, 60.4, 58.9, 57.4, 53.8, 51.7, 40.6, 14.0; HRMS (ESI): m/z calcd. for ($\text{C}_{16}\text{H}_{24}\text{N}_2\text{O}_5^+$) 323.1601, found 323.1601.

Methyl ({2-(benzyl[2ethoxy-2-oxoethyl]amino)acetyl}amino)acetate (6c). The general procedure (I) was applied using L-aspartic acid dimethyl ester hydrochloride (395 mg, 2.0 mmol), compound **5** (502 mg, 2.0 mmol), N-methylmorpholine (205 μ L, 2.0 mmol), DCC (619 mg, 3.0 mmol) and N-hydroxysuccinimide (232 mg, 2 mmol). Purification was carried out according to the method used for compound **2a**, the title compound was obtained as colorless oil (yield 75 %). ^1H NMR (300 MHz, CDCl_3): δ 7.36-7.27 (m, 5H, *Ph*), 4.91-4.87 (m, 1H, NHCHCO_2Me), 4.20-4.15 (m, 3H, OCH_2CH_3), 3.85-3.77 (m, 2H, CH_2Ph), 3.74 (s, 3H, OCH_3), 3.69 (s, 3H, OCH_3), 3.39-3.35 (m, 4H, NCH_2CON), 3.06-3.02 (m, 1H, $\text{NHCHCH}_2\text{CO}_2\text{Me}$), 2.84-2.79 (m, 1H, $\text{NHCHCH}_2\text{CO}_2\text{Me}$), 1.28 (t, 3H, $J=8.4$ Hz, OCH_2CH_3); ^{13}C NMR (75 MHz, CDCl_3): δ 171.2, 171.1, 170.8 (2C), 137.4, 129.1, 128.6, 127.7, 60.7, 58.8, 57.6, 54.6, 52.7, 52.0, 48.0, 36.2, 14.2; HRMS (ESI): m/z calcd. for ($\text{C}_{19}\text{H}_{27}\text{N}_2\text{O}_7^+$) 395.1818, found 395.1810.

Methyl 2-[(2-[(2-ethoxy-2-oxoethyl)amino]acetyl)amino]acetate (7b), general procedure (V). Compound **6b** (322.0 mg, 1.0 mmol) was dissolved in methanol (5 mL) and reduced under H₂ atmosphere for 24 hours in the presence of 10 % Pd/C (37 mg). The reaction mixture was filtered on Celite, and concentrated under reduced pressure, the desired product **7b** was obtained as colourless oil (280.0 mg). ¹H NMR (300 MHz, CDCl₃): δ 4.23-4.16 (m, 2H, OCH₂CH₃), 4.07 (d, 2H, *J* = 5.7 Hz, NHCH₂CO₂Me), 3.76 (s, 3H, OCH₃), 3.47 (s, 2H, CH₂CO₂Et), 3.42 (s, 2H, NHCH₂CON), 1.28 (d, 3H, *J* = 7.1 Hz, OCH₂CH₃); ¹³C NMR (75 MHz, CDCl₃): δ 171.8, 171.6, 170.3, 61.2, 52.4, 51.9, 50.6, 40.8, 14.2; HRMS (ESI): *m/z* calcd. for (C₉H₁₈N₂O₅⁺) 233.1132, found 233.1138.

Dimethyl 2-[(2-[(2-ethoxy-2-oxoethyl)amino]acetyl)amino]succinate (7c). The general procedure (V) was applied using compound **6c** (788.0 mg, 2.0 mmol) and 10 % Pd/C (37 mg) under H₂ atmosphere. The desired product **7c** (yield 95 %) was obtained as colourless oil. ¹H NMR (300 MHz, CDCl₃): δ 4.93-4.87 (m, 1H, NHCHCO₂Me), 4.23-4.16 (m, 2H, CH₂Ph), 3.77 (s, 3H, OCH₃), 3.70 (s, 3H, OCH₃), 3.42 (s, 2H, CH₂CO₂Et), 3.36 (s, 2H, NHCH₂CON), 3.08-2.82 (m, 2H, NHCHCH₂CO₂Me), 1.28 (t, 3H, *J* = 7.1 Hz, OCH₂CH₃); ¹³C NMR (75 MHz, CDCl₃): δ 172.1, 171.3, 171.2, 171.2, 61.1, 52.8, 52.2, 52.1, 50.7, 48.1, 36.3, 14.1; HRMS (ESI): *m/z* calcd. for (C₁₂H₂₁N₂O₇⁺) 305.1349, found 305.1455.

O-(2'-deoxyadenosin-5'-yl)-N-[carboxymethyl]-N'-[2-[(1-carboxy-2-(1H-imidazol-4-yl)ethyl)amino]-2-oxoethyl] phosphoramidate (1f). The general procedure (V) was applied using compound **6a** (431.0 mg, 1.1 mmol) and 10 % Pd/C (37 mg) under H₂ atmosphere, product **7a** was obtained as colourless oil (358 mg). 2'-Deoxyadenosine-5'-

O-monophosphoric acid (50 mg, 0.15 mmol), compound **7a** (200 mg, 0.75 mmol) and DCC (156 mg, 0.75 mmol) were dissolved in a mixture of tert-butanol (3 mL) and water (1 mL). The reaction was heated at 85 °C for 6 hours under argon atmosphere. After completion, the reaction mixture was cooled down and solvent was removed under reduced pressure. The resulting crude material was purified by chromatography on silica gel (gradient: iPrOH/H₂O/NH₃ 7:0.5:0.5, v/v/v; 7:1:1, v/v/v; 7:1:2, v/v/v), the obtained product treated with 0.4 M sodium hydroxide solution in MeOH/H₂O (v/v, 1:4) (3 mL) for 4 hrs, solvent was removed under reduced pressure, the resulting crude material was purified by HPLC, to give compound **1f** as white solid (yield 22 % for two steps). ¹H NMR (500 MHz, D₂O): δ 8.40 (s, 1H, H₈), 8.11 (s, 1H, H₂), 7.50 (s, 1H, NHCH=C), 6.66 (s, 1H, NHCH=N), 6.41 (t, 1H, *J* = 5.8 Hz, H_{1'}), 4.58-4.56 (m, 1H, H_{3'}), 4.28-4.26 (m, 1H, CONHCH), 4.15-4.14 (m, 1H, H_{4'}), 4.01-3.94 (m, 2H, H_{5'}), 3.70-3.56 (m, 2H, NCH₂CO₂H), 3.41-3.13 (m, 2H, NCH₂CONH), 2.97-2.82 (m, 2H, CHCH₂Im) 2.55-2.48 (m, 2H, H_{2'}); ¹³C (75 MHz, D₂O): δ 178.2, 171.6, 173.1, 155.2, 152.2, 148.4, 139.5, 135.3, 135.2, 118.2, 118.1, 85.8, 83.2, 71.1, 64.3, 54.8, 51.7, 51.1, 38.3, 29.9; ³¹P NMR (121 MHz, D₂O) δ 7.98; HRMS (ESI): *m/z* calcd. for (C₂₀H₂₅N₉O₁₀P⁻) 582.1467, found 582.1469.

O-(2'-deoxyadenosin-5'-yl)-*N*-[carboxymethyl]-*N'*-[2-{(carboxymethyl)amino}-2-oxoethyl] phosphoramidate (**1g**). The general procedure (III) was applied using 2'-deoxyadenosine-5'-monophosphate (dAMP) (100 mg, 0.30 mmol), compound **7b** (348 mg, 1.5 mmol) and DCC (433 mg, 2.10 mmol). After purification the desired product was observed as white solid (yield 33 %). ¹H NMR (600 MHz, D₂O): δ 8.40 (s, 1H, H₈), 8.17 (s, 1H, H₂), 6.42 (t, 1H, *J* = 7.8 Hz, H_{1'}), 4.63-4.62 (m, 1H, H_{3'}), 4.24-4.23 (m, 1H,

H_{4'}), 4.09-3.99 (m, 4H, H_{5'} + OCH₂CH₃), 3.93 (s, 2H, NHCH₂CO₂H), 3.69 (s, 3H, OCH₃), 3.68-3.67 (m, 4H, NCH₂CO₂Et + NCH₂CON), 2.88-2.82 (m, 1H, H_{2a'}), 2.64-2.59 (m, 1H, H_{2b'}), 1.14 (t, 3H, *J* = 9.0 Hz, OCH₂CH₃); ¹³C NMR (150 MHz, D₂O): δ 174.3, 173.4, 172.8, 172.2, 155.5, 152.6, 148.7, 139.8, 118.5, 85.9, 83.5, 71.2, 64.4, 62.0, 52.7, 50.2, 48.8, 38.7, 35.2, 13.2; ³¹P NMR (121 MHz, D₂O) δ 6.89; HRMS (ESI): *m/z* calcd. for (C₁₉H₂₆N₇O₁₀P⁻) 544.1562, found 544.1545. This intermediate was then treated with 0.4 M sodium hydroxide solution in MeOH/H₂O (v/v, 1:4) (5 mL) for 3 hours, solvent was removed under reduced pressure, and the resulting residue was purified by chromatography on silica gel (gradient: iPrOH/H₂O/NH₃ 7:0.5:0.5, v/v/v; 7:1:1, v/v/v; 7:1:2, v/v/v) to provide compound **1g** as a white solid (yield 44 %). ¹H NMR (600 MHz, D₂O): δ 8.49 (s, 1H, H₈), 8.24 (s, 1H, H₂), 6.49 (t, 1H, *J* = 7.8 Hz, H_{1'}), 4.70-4.67 (m, 1H, H_{3'}), 4.24-4.22 (m, 1H, H_{4'}), 4.10-4.02 (m, 2H, H_{5'}), 3.75 (d, 2H, *J* = 9.6 Hz, NCH₂CO₂H), 3.73 (s, 2H, NHCH₂CO₂H), 3.59 (d, 2H, *J* = 13.2 Hz, NCH₂CON), 2.85-2.80 (m, 1H, H_{2a'}), 2.61-2.59 (m, 1H, H_{2b'}); ¹³C NMR (150 MHz, D₂O): δ 178.8, 176.5, 174.2, 155.6, 152.7, 148.8, 134.0, 118.7, 86.1, 83.7, 71.3, 64.3, 52.4, 52.2, 43.1, 39.0; ³¹P NMR (121 MHz, D₂O) δ 8.12; HRMS (ESI): *m/z* calcd. for (C₁₆H₂₁N₇O₁₀P⁻) 502.1088, found 502.1079.

O-(2'-deoxyadenosin-5'-yl)-*N*-[carboxymethyl]-*N'*-[2-{(1,2-dicarboxyethyl)amino}-2-oxoethyl] phosphoramidate (**1h**). The general procedure (III) was applied using 2'-deoxyadenosine-5'-monophosphate (dAMP) (100 mg, 0.30 mmol), compound **7c** (456 mg, 1.5 mmol) and DCC (433 mg, 2.10 mmol). After purification the desired product was obtained as white solid (yield 59 %). ¹H NMR (600 MHz, D₂O): δ 8.40 (s, 1H, H₈), 8.19 (s, 1H, H₂), 6.43 (t, 1H, *J* = 6.6 Hz, H_{1'}), 4.69-4.67 (m, 2H, H_{3'} +

NHCHCO₂Me), 4.19-4.18 (m, 1H, H_{4'}), 4.06-3.97 (m, 4H, H_{5'} + OCH₂CH₃), 3.68 (s, 3H, OCH₃), 3.64 (s, 3H, OCH₃), 2.87-2.82 (m, 3H, NHCH₂CO₂Me + H_{2a'}), 2.60-2.56 (m, 1H, H_{2b'}), 1.14 (t, 3H, *J* = 7.2 Hz, OCH₂CH₃); ¹³C NMR (150 MHz, D₂O): δ 173.4, 173.4, 172.8, 172.2, 155.5, 152.6, 148.7, 139.8, 118.6, 85.9, 83.5, 71.2, 64.4, 62.0, 53.1, 52.5, 52.0, 50.3, 48.8, 38.7, 35.2, 13.2; ³¹P NMR (121 MHz, D₂O) δ 7.14; HRMS (ESI): *m/z* calcd. for (C₁₉H₂₆N₇O₁₀P⁻) 544.1562, found 544.1545. This intermediate was then treated with 0.4 M sodium hydroxide solution in MeOH/H₂O (v/v, 1:4) (5 mL) for 3 hours, solvent was removed under reduced pressure, and the resulting residue was purified by chromatography on silica gel (gradient: iPrOH/H₂O/NH₃ 7:0.5:0.5, v/v/v; 7:1:1, v/v/v; 7:1:2, v/v/v) to provide compound **1h** as a white solid (yield 38 %). ¹H NMR (600 MHz, D₂O): δ 8.47 (s, 1H, H₈), 8.17 (s, 1H, H₂), 6.44 (t, 1H, *J* = 7.8 Hz, H_{1'}), 4.67-4.65 (m, 1H, H₃), 4.43-4.40 (m, 1H, CH-Asp), 4.23-4.22 (m, 1H, H_{4'}), 4.08-3.97 (m, 2H, H₅), 3.82-3.72 (m, 2H, NCH₂CO₂H), 3.70-3.47 (dd, 2H, *J*₁ = 11.5 Hz, *J*₁ = 17.9 Hz, NCH₂CONH), 2.81-2.76 (m, 1H, H_{2a'}), 2.69-2.65 (m, 1H, CH_{2a}-Asp), 2.57-2.50 (m, 4H, H_{2b'} + CH_{2b}-Asp); ¹³C NMR (150 MHz, D₂O): δ 178.2, 178.1, 178.0, 173.1, 155.1, 152.3, 148.3, 139.6, 118.1, 85.8, 83.2, 71.0, 64.1, 52.9, 51.5, 51.0, 39.4, 38.5; ³¹P NMR (121 MHz, D₂O) δ 8.05; HRMS (ESI): *m/z* calcd. for (C₁₉H₂₄N₇O₁₀P⁻) 560.1142, found 560.1153.

Methyl 3-(benzylamino)propanoate (8). A solution of benzylamine (1.0 g, 9.33 mmol) in methanol (1 mL) was stirred at room temperature for 1 day, in the presence of methylacrylate (756 μL, 8.40 mmol). The reaction mixture was then concentrated under reduced pressure, the resulting crude material was purified by chromatography on silica gel (Hexane/EtOAc 5:1, v/v) to afford the desired compound **8** as a colorless oil (yield 93 %). ¹H NMR (300 MHz, CDCl₃): δ 7.34-7.26 (m, 5H, *Ph*), 3.82 (s, 2H, NHCH₂Ph),

3.70 (s, 3H, OCH₃), 3.92 (t, 2H, $J = 6.5$ Hz, NHCH₂CH₂CO₂Me), 2.55 (t, 2H, $J = 6.5$ Hz, NHCH₂CH₂CO₂Me); ¹³C NMR (75 MHz, CDCl₃): 173.1, 140.0, 128.4, 128.1, 127.0, 53.7, 51.6, 44.4, 34.5; HRMS (ESI): m/z calcd for (C₁₁H₁₆NO₂⁺) 194.1181, found 194.1190.

Methyl 3-[benzyl-N-{3-(tert-butoxy)-3-oxopropyl}amino] propanoate (9). A solution of **8** (1.67 g, 8.65 mmol) in methanol (1 mL) was stirred at room temperature for 1 day in the presence of *tert*-butyl acrylate (1.40 mL, 9.52 mmol). The reaction mixture was then concentrated under reduced pressure and the resulting crude material was purified by chromatography on silica gel (Hexane/EtOAc 5:1, v/v) to afford the desired compound **9** as a colorless oil (yield 66 %). ¹H NMR (300 MHz, CDCl₃): δ 7.27-7.22 (m, 5H, *Ph*), 3.61 (s, 3H, OCH₃), 3.56 (s, 2H, NCH₂Ph), 2.79-2.73 (m, 4H, NCH₂CH₂CO₂Me + NCH₂CH₂CO₂*t*Bu), 2.44 (t, 2H, $J = 7.2$ Hz, NCH₂CH₂CO₂Me), 2.36 (t, 2H, $J = 7.2$ Hz, NCH₂CH₂CO₂*t*Bu), 1.41 (s, 9H, *t*Bu); ¹³C NMR (75 MHz, CDCl₃): 172.9, 171.8, 139.1, 128.7, 128.1, 126.9, 80.2, 58.3, 51.4, 49.4, 49.1, 33.8, 32.5, 28.0; HRMS (ESI): m/z calcd for (C₁₈H₂₈NO₄⁺) 322.4192, found 322.2016.

N-benzyl-N'-[3-methoxy-3-oxopropyl] β -alanine trifluoroacetate (10). The *tert*-butyl ester **9** (1.83 g, 5.69 mmol) was treated by TFA (4.23 mL, 56.9 mmol) in CH₂Cl₂ (40mL) for 2 days at room temperature. After removal of solvent and excess of TFA, the crude residue was triturated with Et₂O and filtered, affording **10** as a white solid of TFA salt (yield 97 %). ¹H NMR (300 MHz, DMSO d₆): δ 7.55-7.46 (m, 5H, *Ph*), 4.40 (s, 2H, NCH₂Ph), 3.62 (s, 3H, OCH₃), 3.35-3.27 (m, 4H, NCH₂CH₂CO₂Me + NCH₂CH₂CO₂H), 2.89 (t, 2H, $J = 7.4$ Hz, NCH₂CH₂CO₂Me), 2.79 (t, 2H, $J = 7.4$ Hz, NCH₂CH₂CO₂H); ¹³C NMR (75 MHz, DMSO d₆): 171.5, 170.4, 131.2, 129.8, 129.7, 129.0, 56.6, 51.9,

47.7, 47.4, 28.2, 28.0; HRMS (ESI): m/z calcd for ($C_{14}H_{20}NO_4^+$) 266.3129, found 266.1386.

Methyl 3-[benzyl-{3-[(2-methoxy-2-oxoethyl)amino]-3-oxopropyl}amino]propanoate

(**11a**). The general procedure (IV) was applied using compound **10** (TFA salt) (500.0

mg, 1.32 mmol), glycine methyl ester hydrochloride (166.0 mg, 1.32 mmol),

triethylamine (550 μ L, 3.95 mmol), 1-hydroxybenzotriazole (267.1 mg, 1.97 mmol) and

N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (379.0 mg, 1.97 mmol).

After purification, the resulting **11a** was obtained as a white solid (yield 57 %). 1H NMR

(300 MHz, $CDCl_3$): δ 8.08 (broad s, *NH*), 7.38-7.30 (m, 5H, *Ph*), 3.95 (d, 2H, $J = 5.4$ Hz,

$NHCH_2CO_2Me$), 3.70 (s, 3H, $NCH_2CH_2CO_2CH_3$), 3.57 (s, 3H, $NHCH_2CO_2CH_3$), 3.55

(s, 2H, NCH_2Ph), 2.80 (t, 2H, $J = 6.9$ Hz, NCH_2CH_2CONH), 2.73 (t, 2H, $J = 6.2$ Hz,

$NCH_2CH_2CO_2Me$), 2.50 (t, 2H, $J = 6.9$ Hz, NCH_2CH_2CONH), 2.40 (t, 2H, $J = 6.2$ Hz,

$NCH_2CH_2CO_2Me$); ^{13}C NMR (75 MHz, $CDCl_3$): 172.6, 172.5, 170.2, 137.6, 129.0,

128.2, 127.1, 66.0, 51.8, 51.4, 49.8, 48.9, 40.7, 32.8, 31.8; HRMS (ESI): m/z calcd for

($C_{17}H_{25}N_2O_5^+$) 337.3908, found 337.1763.

Dimethyl 2-[{3-(benzyl(3-methoxy-3-oxopropyl)amino)propanoyl}amino]succinate

(**11b**). The general procedure (IV) was applied using compound **10** (TFA salt) (500.0

mg, 1.32 mmol), L-aspartic acid dimethyl ester hydrochloride (260.8 mg, 1.32 mmol),

triethylamine (550 μ L, 3.95 mmol), 1-hydroxybenzotriazole (267.1 mg, 1.97 mmol) and

N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (379.0 mg, 1.97 mmol).

After purification the resulting compound **11b** was obtained as a white solid (yield

57.5 %). 1H NMR (300 MHz, $CDCl_3$): δ 8.20 (broad s, *NH*), 7.27-7.21 (m, 5H, *Ph*),

4.89-4.83 (m, 1H, NHCHCO_2Me), 3.70 (s, 3H, $\text{NHCHCO}_2\text{CH}_3$), 3.62 (s, 3H, $\text{NHCHCH}_2\text{CO}_2\text{CH}_3$), 3.60 (s, 5H, $\text{NCH}_2\text{CH}_2\text{CO}_2\text{CH}_3 + \text{NCH}_2\text{Ph}$), 2.97-2.83 (m, 2H, $\text{NHCHCH}_2\text{CO}_2\text{Me}$), 2.81-2.71 (m, 4H, $\text{NCH}_2\text{CH}_2\text{CONH} + \text{NCH}_2\text{CH}_2\text{CO}_2\text{Me}$), 2.49 (t, 2H, $J = 7.2$ Hz, $\text{NCH}_2\text{CH}_2\text{CO}_2\text{Me}$), 2.38 (t, 2H, $J = 6.0$ Hz, $\text{NCH}_2\text{CH}_2\text{CONH}$); ^{13}C NMR (75 MHz, CDCl_3): 172.5, 171.9, 171.0, 170.5, 137.6, 128.8, 128.1, 127.0, 57.8, 52.3, 51.6, 51.3, 49.8, 48.5, 48.1, 36.0, 33.0, 31.3; HRMS (ESI): m/z calcd for ($\text{C}_{20}\text{H}_{29}\text{N}_2\text{O}_7^+$) 409.1975, found 409.1969.

Methyl 3-[(3-[(2-methoxy-2-oxoethyl)amino]-3-oxopropyl)amino]propanoate (12a).

The general procedure (V) was applied using compound **11a** (250.0 mg, 0.74 mmol) and 10 % Pd/C (26 mg) under H_2 atmosphere. After filtration and concentration, product **12a** was obtained as colourless oil (yield 96 %). ^1H NMR (300 MHz, CDCl_3): δ 8.23 (broad s, NH), 3.98 (s, 2H, $\text{NHCH}_2\text{CO}_2\text{Me}$), 3.70 (s, 6H, $\text{NHCH}_2\text{CO}_2\text{CH}_3 + \text{NHCH}_2\text{CH}_2\text{CO}_2\text{CH}_3$), 3.35-3.23 (m, 4H, $\text{NHCH}_2\text{CH}_2\text{CO}_2\text{Me} + \text{NHCH}_2\text{CH}_2\text{CONH}$), 3.01-2.94 (m, 4H, $\text{NHCH}_2\text{CH}_2\text{CO}_2\text{Me} + \text{NHCH}_2\text{CH}_2\text{CONH}$); ^{13}C NMR (75 MHz, CDCl_3): 171.0, 170.9, 170.5, 52.3, 52.2, 44.5, 43.6, 40.9, 31.1, 30.2; HRMS (ESI): m/z calcd for ($\text{C}_{10}\text{H}_{19}\text{N}_2\text{O}_5^+$) 247.2683, found 247.1285.

Dimethyl 2-[(3-[(3-methoxy-3-oxopropyl)amino]propanoyl)amino]succinate (12b). The general procedure (V) was applied using compound **11b** (310.0 mg, 0.76 mmol) and 10 % Pd/C (26 mg) under H_2 atmosphere. After filtration and concentration, product **12b** was obtained as colourless oil (yield 71 %). ^1H NMR (300 MHz, CDCl_3): δ 8.24 (d, 1H, NH), 4.82-4.76 (m, 1H, NHCHCO_2Me), 4.66 (broad s, 1H, NH), 3.67 (s, 3H, $\text{NHCHCO}_2\text{CH}_3$), 3.63 (s, 3H, $\text{NHCHCH}_2\text{CO}_2\text{CH}_3$), 3.61 (s, 3H, $\text{NHCH}_2\text{CH}_2\text{CO}_2\text{CH}_3$), 3.00-2.93 (m, 4H, $\text{NHCH}_2\text{CH}_2\text{CONH} + \text{NHCH}_2\text{CH}_2\text{CO}_2\text{Me}$), 2.92-2.74 (m, 2H, $\text{NHCHCH}_2\text{CO}_2\text{Me}$), 2.64

(t, 2H, $J = 6.2$ Hz, $\text{NHCH}_2\text{CH}_2\text{CO}_2\text{Me}$), 2.56-2.52 (m, 2H, $\text{NHCH}_2\text{CH}_2\text{NCONH}$); ^{13}C NMR (75 MHz, CDCl_3): 172.2, 171.6, 171.1, 171.0, 52.5, 51.8, 51.7, 48.3, 44.7, 44.0, 35.9, 33.7, 32.6; HRMS (ESI): m/z calcd for $(\text{C}_{13}\text{H}_{23}\text{N}_2\text{O}_7^+)$ 319.1505, found 319.1501.

O-(2'-deoxyadenosin-5'-yl)-*N*-[2-carboxyethyl]-*N'*-[3-{(carboxymethyl)amino}-3-oxopropyl] phosphoramidate (**1i**). The general procedure (III) was applied using 2'-deoxyadenosine-5'-monophosphate (dAMP) (100 mg, 0.30 mmol), compound **12a** (369 mg, 1.5 mmol) and DCC (433 mg, 2.10 mmol). After purification the observed crude material was treated with 0.4 M sodium hydroxide solution in MeOH/H₂O (v/v, 1:4) (5 mL) for 3 hours. Solvent was removed under reduced pressure, and the resulting crude product was purified by chromatography on silica gel (gradient: iPrOH/H₂O/NH₃ 7:0.5:0.5, v/v/v; 7:1:1, v/v/v; 7:1:2, v/v/v) to provide compound **1i** as a white solid (yield 28 % for two steps). ^1H NMR (600 MHz, D₂O): δ 8.42 (s, 1H, H₈), 8.18 (s, 1H, H₂), 6.45 (t, 1H, $J = 6.8$ Hz, H_{1'}), 4.70-4.69 (m, 2H, H_{3'}), 4.21-4.20 (m, 1H, H_{4'}), 3.89-3.85 (m, 2H, H_{5'}), 4.64-3.58 (m, 2H, NHCH_2COOH), 3.14-3.12 (m, 2H, $\text{NCH}_2\text{CH}_2\text{CONH}$), 3.05-3.02 (m, 2H, $\text{NCH}_2\text{CH}_2\text{COOH}$), 2.88-2.79 (m, 2H, $\text{CH}_2\text{CH}_2\text{COOH} + \text{H}_{2\text{a}}'$), 2.59-2.52 (m, 1H, H_{2\text{b}}'}), 2.34-2.56 (m, 2H, $\text{CH}_2\text{CH}_2\text{COOH}$), 2.24-2.17 (m, 1H, $\text{NCH}_2\text{CH}_2\text{CONH}$); ^{13}C NMR (150 MHz, D₂O): δ 178.6, 176.0, 173.7, 156.9, 152.8, 149.7, 141.3, 119.9, 85.8, 85.1, 70.6, 63.4, 42.7 (3C), 38.0, 35.5, 35.3; ^{31}P NMR (121 MHz, D₂O) δ 8.48; HRMS (ESI): m/z calcd. for $\text{C}_{18}\text{H}_{25}\text{N}_7\text{O}_{10}\text{P}^-$ 530.1401, found 530.1409.

O-(2'-deoxyadenosin-5'-yl)-*N*-[2-carboxyethyl]-*N'*-[2-{(carboxymethyl)amino}-2-oxoethyl] phosphoramidate (**1j**). The general procedure (III) was applied using 2'-deoxyadenosine-5'-monophosphate (dAMP) (100 mg, 0.30 mmol), compound **12b** (477

mg, 1.5 mmol) and DCC (433 mg, 2.10 mmol). After purification the resulting crude product was treated with 0.4 M sodium hydroxide solution in MeOH/H₂O (v/v, 1:4) (5 mL) for 3 hours. Solvent was removed under reduced pressure, and the resulting crude product was purified by chromatography on silica gel (gradient: iPrOH/H₂O/NH₃ 7:0.5:0.5, v/v/v; 7:1:1, v/v/v; 7:1:2, v/v/v) to provide compound **1j** as a white solid (yield 32 %). ¹H NMR (600 MHz, D₂O): δ 8.43 (s, 1H, H₈), 8.19 (s, 1H, H₂), 6.43 (t, 1H, J = 6.6 Hz, H_{1'}), 4.70-4.68 (m, 1H, H_{3'}), 4.39-4.37 (m, 1H, CH-Asp), 4.19-4.17 (m, 1H, H_{4'}), 3.87-3.83 (m, 2H, H_{5'}), 3.06-3.00 (m, 4H, NCH₂CH₂CO), 2.86-2.81 (m, 1H, H_{2a'}), 2.68-2.65 (m, 1H, CH_{2a}-Asp), 2.58-2.53 (m, 2H, CH_{2b}-Asp+H_{2b'}), 2.31-2.29 (m, 4H, NCH₂CH₂CO); ¹³C NMR (150 MHz, D₂O): δ 179.7, 178.6, 178.0, 175.0, 155.6, 150.4, 149.7, 139.9, 119.9, 85.8, 85.6, 83.6, 70.6, 63.3, 51.3, 42.8, 42.4, 38.5, 35.6, 34.8; ³¹P NMR (121 MHz, D₂O) δ 8.48; HRMS (ESI): m/z calcd. for C₂₀H₂₇N₇O₁₂P⁻ 588.1455, found 588.1462.

Diethyl 2-[(2-((benzyloxy)carbonyl)amino)acetyl]amino]malonate (13). The general procedure (IV) was applied using diethyl aminomalonate hydrochloride (1.0 g, 4.72 mmol), triethylamine (722 μ L, 5.20 mmol), carbobenzyloxy glycine (988.4 mg, 4.72 mmol), 1-hydroxybenzotriazole (957.5 mg, 7.809 mmol) and ethyldiaminopropylcarbodiimide (1.36 g, 7.09 mmol). After purification the resulting compound **13** was obtained as a white solid (yield 87 %). ¹H NMR (300 MHz, CDCl₃): δ 7.38-7.29 (m, 5H, *Ph*), 6.91 (broad s, 1H, *NH*), 5.36, (s, 1H, *NH*), 5.16-5.14 (m, 3H, CH₂Ph + CHCO₂Et), 4.33-4.23 (m, 4H, CO₂CH₂CH₃), 4.00-3.98 (m, 2H, NHCH₂), 1.30 (t, 6H, J = 7.0 Hz, CO₂CH₂CH₃); ¹³C NMR (75 MHz, CDCl₃): 168.7, 167.3, 166.0,

139.9, 128.5, 128.2, 128.1, 67.3, 62.8, 55.1, 44.3, 14.0 ; HRMS (ESI): m/z calcd for ($C_{17}H_{23}N_2O_7^+$) 367.1500, found 367.1508.

Diethyl 2-[(2-aminoacetyl)amino]malonate hydrochloride (14). Compound **13** (300 mg, 0.82 mmol) was dissolved in ethanol (5 mL) and reduced under H_2 atmosphere for 6 hours, in the presence of 10 % Pd/C (29 mg) and hydrochloric acid (50 μ L). The reaction mixture was then filtered on Celite and concentrated under reduced pressure. The title compound **14** was obtained as a colourless oil without additional purification (yield 88 %). 1H NMR (300 MHz, $CDCl_3$): δ 9.12 (broad s, 1H, *NH*), 5.20, (s, 1H, *CHCO*₂Et), 4.27-4.20 (m, 4H, *CO*₂*CH*₂*CH*₃), 3.78 (s, 2H, *CH*₂*NH*₂), 1.29 (t, 6H, $J = 7.0$ Hz, *CO*₂*CH*₂*CH*₃); ^{13}C NMR (75 MHz, $CDCl_3$): 167.4, 167.3, 63.7, 57.9, 41.5, 14.3; HRMS (ESI): m/z calcd for ($C_9H_{17}N_2O_5^+$) 233.1132, found 233.1135.

O-(2'-deoxyadenosin-5'-yl)-*N*-[2-{(carboxymethyl)amino}-2-oxoethyl]phosphoramidate (**1k**). 2'-Deoxyadenosine-5'-O-mono- phosphoric acid (50 mg, 0.15 mmol), compound **14** (194 mg, 0.76 mmol) and DCC (156 mg, 0.76 mmol) were dissolved in a mixture of tert-butanol (3 mL) and H_2O (1mL). The reaction was heated at 85 °C for 6 hours under argon atmosphere. After completion, the reaction mixture was cooled down and solvent was removed under reduced pressure. The resulting crude material was purified by chromatography on silica gel (gradient: iPrOH/ H_2O / NH_3 7:0.5:0.5, v/v/v; 7:1:1, v/v/v; 7:1:2, v/v/v), the obtained product treated with 0.4 M sodium hydroxide solution in MeOH/ H_2O (1:4, v/v) for 4 hours, solvent was removed under reduced pressure, the resulting crude material was purified by HPLC, to give compound **1k** as white solid (yield 32 % for two steps). 1H NMR (500 MHz, D_2O): δ 8.39 (s, 1H, *H*₈), 8.20 (s, 1H, *H*₂), 6.45 (t, 1H, $J = 5.8$ Hz, *H*_{1'}), 4.70-4.67 (m, 1H, *H*_{3'}), 4.17-4.14 (m, 1H, *H*_{4'}), 4.03-

3.92 (m, 2H, H_{5'}), 3.64 (s, 2H, NHCH₂CO₂H), 3.54 (d, 2H, NHCH₂CONH), 2.84-2.52 (m, 2H, H_{2'}); ¹³C (75 MHz, D₂O): δ 178.2, 171.6, 155.5, 152.6, 148.6, 139.7, 118.6, 85.6, 83.6, 71.2, 64.1, 44.3, 42.8, 38.6; ³¹P NMR (121 MHz, D₂O) δ 7.73; HRMS (ESI): *m/z* calcd. for (C₁₄H₁₉N₇O₈P⁻) 444.1033, found 444.1042.

Enzymatic reactions. - *DNA duplexes:* Oligonucleotides P₁ and T₁ were purchased from Sigma Genosys. The concentrations were determined with a Varian Cary-300-Bio UV Spectrophotometer. The primer P₁ was 5'-³³P-labeled with 5'-[γ³³P]-ATP (Perkin Elmer) using T4 polynucleotide kinase (New England Biolabs) according to the procedure provided by the supplier. The labeled oligonucleotide was further purified using Illustra™ Microspin TM G-25 columns (GE Healthcare).

DNA polymerase reactions: End-labeled primer was annealed to its template by combining primer and template in a molar ratio of 1:2 and heating the mixture to 75 °C for 10 min followed by slowly cooling to room temperature over a period of 1.5 h. For the incorporation of compound **1a-1k**, a series of 20 μL-batch reactions were performed with the enzyme (HIV-1 RT, Taq, Vent (exo⁻), Terminator DNA polymerase). The final reaction mixture contained 125 nM primer-template complex, reaction buffer (50 mM Tris.HCl, 50 mM KCl, 10 mM MgCl₂, 0.5 mM spermidine, 10 mM dithiothreitol (DTT); pH 8.3 for HIV-1 RT; 20 mM Tris.HCl, 10 mM KCl, 2 mM MgSO₄, 0.1 % Triton X-100, pH 8.8 for thermostable polymerase), appropriate concentration of enzyme, and different concentrations of phosphoramidate building blocks (1 mM, 500 μM, 200 μM and 100 μM). In the control reaction a 10 μM dATP was used as reference. The reaction mixture was incubated at 37 °C or 75 °C respectively, aliquots were quenched after 10, 20, 30, 60 and 120 min.

Polyacrylamide gel electrophoresis. - All polymerase reaction aliquots (2.5 μ L) were quenched by addition of 10 μ L of loading buffer (90 % formamide, 0.05 % bromophenol blue, 0.05 % xylene cyanol and 50 mM ethylenediaminetetraacetic acid (EDTA)). Samples were heated at 85 °C for 3 minutes prior to analysis by electrophoresis for 2.5 h at 2000 V on a 30 cm x 40 cm x 0.4 mm 20 % (19:1 mono:bis) denaturing gel in the presence of a 100 mM Tris-borate, 2.5 mM EDTA buffer; pH 8.3. Products were visualised by phosphorimaging. The amount of radioactivity in the bands corresponding to the products of enzymatic reactions was determined by using the imaging device Cyclone® and the Optiquant image analysis software (Perkin Elmer).

The authors thanks FWO and K.U. Leuven (IDO, GOA) for financial support, Jef Rozenski for providing HRMS, Luc Baudempiez for recoding the NMR spectra and Chantal Biernaux for editorial help.

REFERENCES

- [1] P. Herdewijn, P. Marlière, *Chem. Biodiv.* **2009**, 6, 791.
- [2] V. B. Pinheiro, A. I. Taylor, C. Cozens, M. Abramov, M. Renders, S. Zhang, J.C. Chaput, J. Wengel, S.-Y. Peaak-Chew, S. H. McLaughlin, P. Herdewijn, P. Holliger, *Science* **2012**, 336, 341.
- [3] O. Adelfinskaya, P. Herdewijn, *Angew. Chem., Int. Ed.* **2007**, 46, 4356.
- [4] X.-P. Song, C. Bouillon, E. Lescrinier, P. Herdewijn, *ChemBioChem* **2011**, 12, 1868.

- [5] A. Giraut, P. Herdewijn, *ChemBioChem* **2010**, *11*, 1399.
- [6] A. Giraut, X.-P. Song, M. Froeyen, P. Marlière, P. Herdewijn, *Nucleic Acid Res.* **2010**, *38*, 2541.
- [7] S. Yang, M. Froeyen, E. Lescrinier, P. Marlière, P. Herdewijn, *Org. Biomol. Chem.* **2011**, *9*, 111.
- [8] S. Yang, C. Pannecouque, E. Lescrinier, P. Marlière, A. Giraut, P. Herdewijn, *Org. Biomol. Chem.* **2012**, *10*, 146.
- [9] J. W. Payne, C. Gilvarg, *J. Biol. Chem.* **1968**, *243*, 335.
- [10] N. M. Hooper in, 'Handbook of Proteolytic Enzymes (2nd edition)' Eds. A.J. Barrett, N.D. Rawlings, J.F. Woessner, Elsevier, London, 2004, p. 994.
- [11] T. W. Abraham, T. I. Kalman, E. J. McIntee, C. R. Wagner, *J. Med. Chem.* **1996**, *39*, 4569.
- [12] C. M. Joyce, T. A. Steitz, *Annu. Rev. Biochem.* **1994**, *63*, 777.
- [13] M. S.Boosalis, J. Petruska, M. F. Goodman, *J. Biol. Chem.* **1987**, *262*, 14689.
- [14] S. Creighton, L. B. Bloom, M. F. Goodman, *Methods Enzymol.* **1995**, *262*, 232.
- [15] H. Huang, R. Chopra, G. L. Verdine, S. C. Harrison, *Science* **1998**, *282*, 1669.
- [16] A. F. Gardner, W. E. Jack, *Nucleic Acids Res.* **1999**, *27*, 2545.
- [17] A. F. Gardner, W. E. Jack, *Nucleic Acids Res.* **2002**, *30*, 605.

Table 1. *Overview of compounds 1a-1k*

Compound No.	Structure	Compound No.	Structure
1a	Asp-Asp-dAMP	1g	Gly-IDA-dAMP
1b	Gly-Asp-dAMP	1h	Asp-IDA-dAMP
1c	His-Asp-dAMP	1i	Gly-IDP-dAMP
1d	Asp-His-dAMP	1j	Asp-IDP-dAMP
1e	Asp-Gly-dAMP	1k	Gly-Gly-dAMP
1f	His-IDA-dAMP		

Table 2. *Primer-template complex used in the incorporation assay.*

Oligonucleotide	Sequence
P ₁	5'-CAGGAAACAGCTATGAC-3'
T ₁	3'-GTCCTTTGTCGATACTGT CCC -5'

Table 3. Overview of incorporation results of compound **1a-1k** by HIV-1 RT.

Conditions: primer *P*₁, template *T*₁, [HIV-1 RT]=0.025 U/ μ L

Compound No.	Incorporation result, %	Compound No.	Incorporation result, %
1a	33	1g	0
1b	0	1h	0
1c	20	1i	18
1d	15	1j	5
1e	14	1k	21
1f	20		

LEGENDS

Scheme 1. *Synthesis of compound 1a-1e. Reagents and conditions: i) DCC, HoSu, N-ethylmorpholine, CH₂Cl₂, rt; ii) TFA, CH₂Cl₂, rt; iii) dAMP, DCC, 1,4-dioxane/H₂O, Et₃N, 80 °C; iv) 10 %, Pd/C, H₂, rt. for compound 1a, 1b, and 1e; 0.4 M NaOH (MeOH/H₂O), rt. for compound 1c and 1d*

Scheme 2. *Synthesis of compound 1f-1h. Reagents and conditions: i) DiPEA, ACN, rt.-reflux; ii) TFA, CH₂Cl₂, rt; iii) EDCi, HOBT, CH₂Cl₂, Et₃N, rt for 6a and DCC, HoSu, N-ethylmorpholine, CH₂Cl₂, rt for 6b-6c; iv) H₂, 10 %, Pd/C, MeOH, rt; v) dAMP, DCC, 1,4-dioxane/H₂O, Et₃N, 80 °C for 1g and 1h and dAMP, DCC, t-BuOH/H₂O, Et₃N, 80 °C for 1f; vi) 0.4 M NaOH (MeOH/H₂O), rt.*

Scheme 3. *Synthesis of phosphoramidate analogue 1i and 1j. Reagents and conditions: i) methylacrylate, MeOH, rt; ii) tert-Butyl-acrylate, MeOH, rt.; iii) TFA, CH₂Cl₂, rt.; iv) EDCi, HOBT, CH₂Cl₂, Et₃N, rt; v) H₂, 10 %, Pd/C, MeOH, rt; vi) dAMP, DCC, 1,4-dioxane/H₂O, Et₃N, 80 °C; vii) 0.4 M NaOH (MeOH/H₂O), rt.*

Scheme 4. *Synthesis of phosphoramidate analogue 1k. Reagents and conditions: i) EDCi, HOBT, CH₂Cl₂, Et₃N, rt; ii) H₂, 10 %, Pd/C, MeOH, HCl, rt; iii) dAMP, DCC, t-BuOH/H₂O, Et₃N, 80 °C; iv) 0.4 M NaOH (MeOH/H₂O), rt.*

Figure 1. *Single nucleotide incorporation of compound 1a-1k into P₁T₁ by Terminator polymerase. [Terminator pol]=0.05 U/μL*

Figure 2. *Single incorporation of compound 1c into P₁T₁ by Terminator DNA*

polymerase. [Terminator pol]=0.05 U/ μ L

Time points: 10, 20, 30, 60 and 120 min.

Figure 3. *Single nucleotide incorporation of compound 1a-1k by Taq and Vent (exo⁻)*

DNA polymerase. [Vent(exo⁻) pol]=0.1 U/ μ L, [Taq pol]=0.15 U/ μ L.

Figure 4. *Single incorporation of compound 1f into P₁T₁ by Vent (exo⁻) DNA*

polymerase. [Vent(exo⁻) pol]=0.1 U/ μ L

Time points: 10, 20, 30, 60 and 120 min.

Scheme 1

Scheme 2

Scheme 3

Scheme 4

Figure 1

Figure 2

Figure 3

Figure 4