Exploring the role of the α-carboxyphosphonate moiety in the HIV-RT activity of α-carboxy nucleoside phosphonates

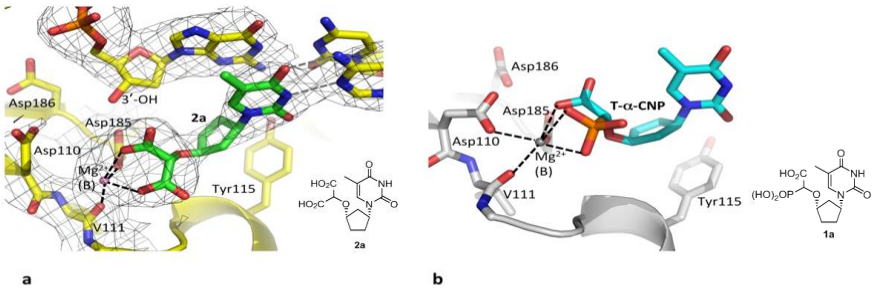
Nicholas D. Mullins,[†] Nuala M. Maguire,[†] Alan Ford,[†] Kalyan Das,[‖] Edward Arnold, [‖] Jan Balzarini[‡] and Anita R. Maguire[†§]*

[†]Department of Chemistry and [§]School of Pharmacy, Analytical and Biological Chemistry Research Facility, Synthesis and Solid State Pharmaceutical Centre, University College, Cork, Ireland

[‡]KU Leuven, Rega Institute for Medical Research, Minderbroedersstraat 10, B-3000 Leuven, Belgium

[‖]Center for Advanced Biotechnology and Medicine and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ 08854-8020; USA

Table of Contents Graphic and Synopsis



The importance of the α-carboxy phosphonoacetic acid moiety to achieve RT inhibition, without the need for prior phosphorylation, is confirmed.

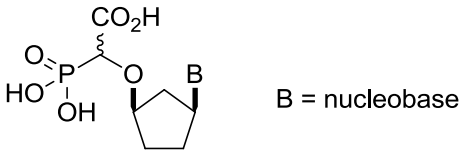
ABSTRACT

As α-carboxy nucleoside phosphonates (α-CNPs) have demonstrated a novel mode of action of HIV-1 reverse transcriptase inhibition, structurally related derivatives were synthesized, namely the malonate **2**, the unsaturated and saturated bisphosphonates **3** and **4**, respectively and the amide **5**. These compounds were evaluated for inhibition of HIV-1 reverse transcriptase in cell-free assays. The importance of the α-carboxy phosphonoacetic acid moiety for achieving reverse transcriptase inhibition, without the need for prior phosphorylation, was confirmed. The malonate derivative **2** was less active by two orders of magnitude than the original α-CNPs, while displaying the same pattern of kinetic behavior; interestingly the activity resides in the "L"-enantiomer of **2**, as seen with the earlier series of α-CNPs. A crystal structure with an RT/DNA complex at 2.95 Å resolution revealed the binding of the "L"-enantiomer of **2**, at the polymerase active site with a weaker metal ion chelation environment compared to **1a** (T-α-CNP) which may explain the lower inhibitory activity of **2**.

INTRODUCTION

World Health Organization statistics state that 2.1 million people were infected with the human immunodeficiency virus (HIV) in 2013, bringing the number of people now living with the virus to approximately 35 million.[1] While a recent study claims that the retrovirus is slowly becoming milder in its virulence due to a number of factors (i.e. mutations, increased drug resistance),[2] other work has predicted increasing rates of disease transmission.[3] Thus, the quest for novel drugs that possess anti-HIV properties (in terms of mechanism of action, resistance profile, inhibition potency, side effects profile, etc.) is of utmost importance.

Recently, we disclosed the synthesis and evaluation of a novel class of α-carboxy nucleoside phosphonates (α-CNPs) **1** which exhibit potent HIV-1 reverse transcriptase (RT) inhibitory properties in cell free assays (Figure 1).[4,5] Most importantly, these compounds unlike nucleoside-analog drugs do not require intracellular activation (*i.e.* phosphorylation) in order to display activity against the virus-encoded RT. Interestingly, the α-carboxy phosphonate moiety is acting as a triphosphate mimic as demonstrated by crystallographic analysis. In fact, the carboxylate oxygen mimics the interaction of a nucleotide α-phosphate with a Mg^{2+} ion, while the two phosphonate oxygens of the α-CNPs mirror the interaction of β- and γ-phosphate oxygens of nucleotides with the same Mg^{2+} ion.[5]



1

Figure 1. α-carboxynucleoside phosphonate (α-CNP)

Intrigued by this novel mode of action, we decided to investigate the influence of altering the carboxy phosphonate group on HIV-1 RT inhibition while keeping the remainder of the α-CNP structure intact. Accordingly, the malonate **2**, bisphosphonates **3** and **4** and the amide **5** were chosen as potential targets. In addition, previously synthesized esters[4] **6** and **7** were also included in the evaluation (Figure 2). The key aspect we wish to investigate is whether the activity of the α-carboxy phosphonate could be mimicked through the diacid **2** or the bisphosphonates, or indeed whether the carboxylic acid could be replaced by a primary amide in **5** or an ester in **6** and **7**, thereby determining whether the ionizable acid is essential, or whether the metal ion coordinating ability of a primary amide is sufficient to maintain the inhibitory activity of the α-CNPs against HIV-1 RT.

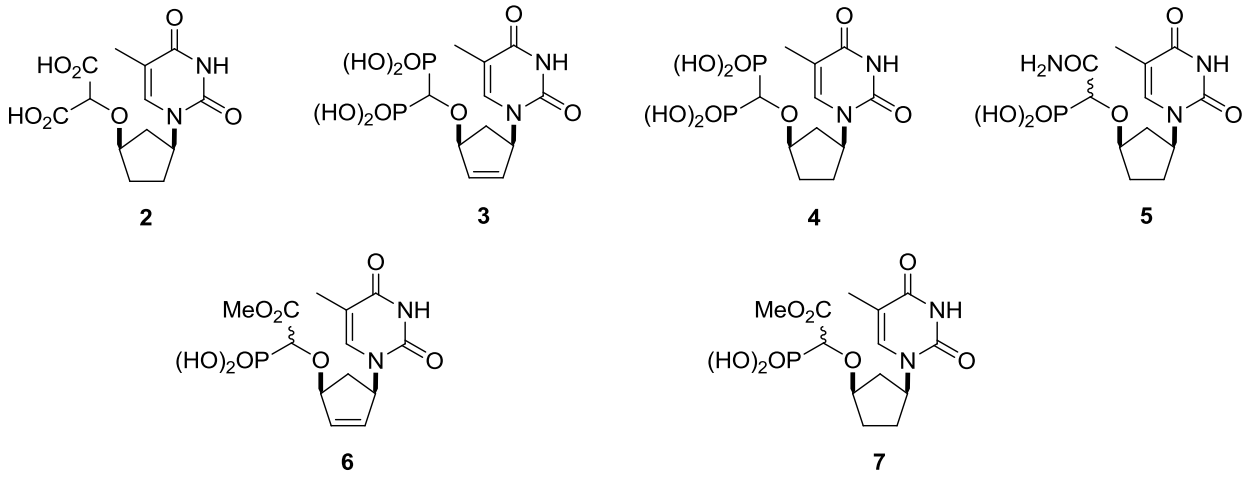
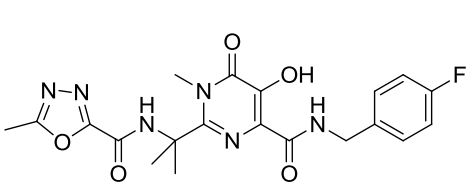


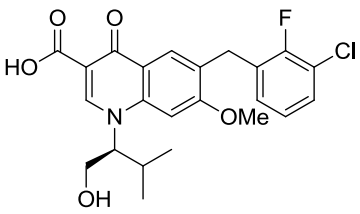
Figure 2. Targets selected for HIV-RT inhibition

Compound **2** possesses a diacid moiety which is related to the diketoacid (DKA) structural feature which has been implicated in the activity of HIV-1 integrase (IN) inhibitors,[6] namely raltegravir **8**,[7,8] elvitegravir[9] **9** and the recently approved dolutegravir **10**.[10] The planar geometry of the DKA moiety is believed to facilitate chelation of two Mg^{2+} ions in the active site. [11,12]

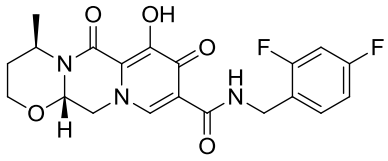
Compound **11** exhibits moderate inhibition of wild-type HIV-1 RT polymerase.[13] The bisphosphonate moiety of **3** and **4** is a key structural motif in compounds such as pamidronic acid **12** which is used in the treatment of bone-related diseases.[14] Recently, advocacy for their use in the treatment of HIV-related osteoporosis has increased.[15] Inorganic pyrophosphate (PPi) **13** is formed during the catalytic cycle of nucleotide monophosphate incorporation into a growing nucleotide copolymer and it can induce a conformational change in polymerases from an open to a closed conformation by binding to the active site.[16] Investigations have centered on the ability of bioisosteres of PPi to inhibit HIV-1 RT. Foscarnet (phosphonoformic acid)[17-19] **14** acts by trapping the pre-translocational state of HIV-1 RT, and while phosphonoacetic acid (PAA) **15** is a potent anti-herpetic agent, it is virtually ineffective against HIV.[20-22] Oxophosphonoacetate **16** is an inhibitor of some nucleic acid polymerases.[22,23] Carbonylbisphosphonate (COBP) **17,** first isolated by Quimby[24] and further investigated by McKenna, is efficient at inhibiting HIV-1 RT.[23,25,26] Compound **18** is non-inhibitory towards HIV-1 RT, however, and despite extensive efforts having been devoted to the synthesis of small molecule substituted derivatives of **15** and **18**, these efforts have usually resulted in less effective HIV-1 RT inhibitors than foscarnet.[26] The preparation and characterization of AZT 5´-COBP **19** has been described, but no biological data has been reported.[27] More recently, purine-like bisphosphonates **20** and **21** have been shown to inhibit HIV-1 RT-catalyzed synthesis.[28] Finally, the amide **5** is reminiscent of anti-HIV pronucleosides **22** and **23** documented by Shirokova, with **22** inhibiting virus replication by 90% at significantly lower concentrations than AZT.[29]



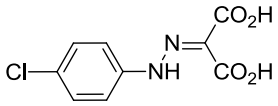
8



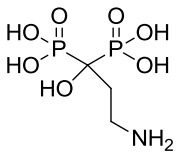
9



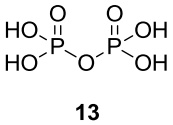
10



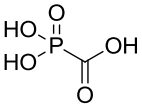
11



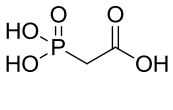
12



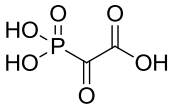
13



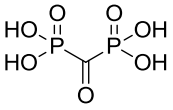
14



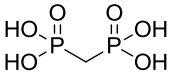
15



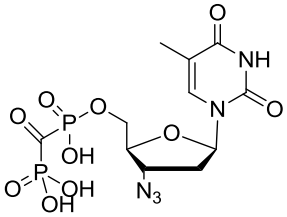
16



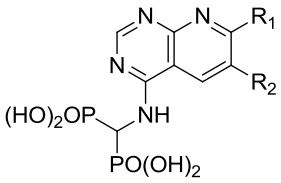
17



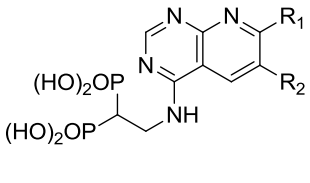
18



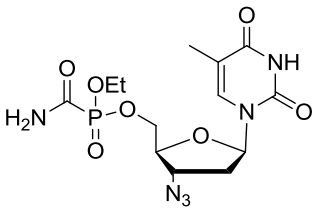
19



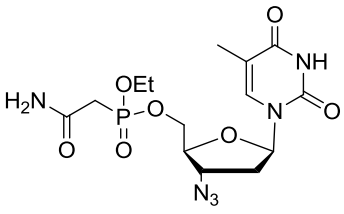
20



21



22

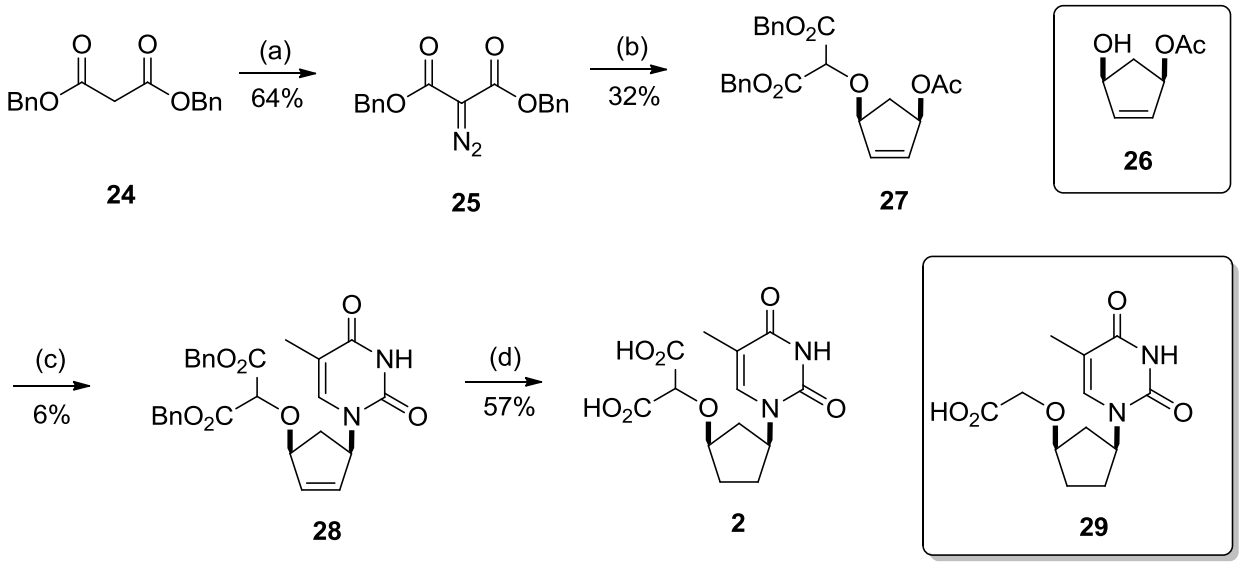


23

In this context, evaluation of the properties of the malonate **2**, bisphosphonates **3** and **4** and the amide **5** was warranted.

RESULTS AND DISCUSSION

Scheme 1. Synthesis of malonate derivative **2**.[a]



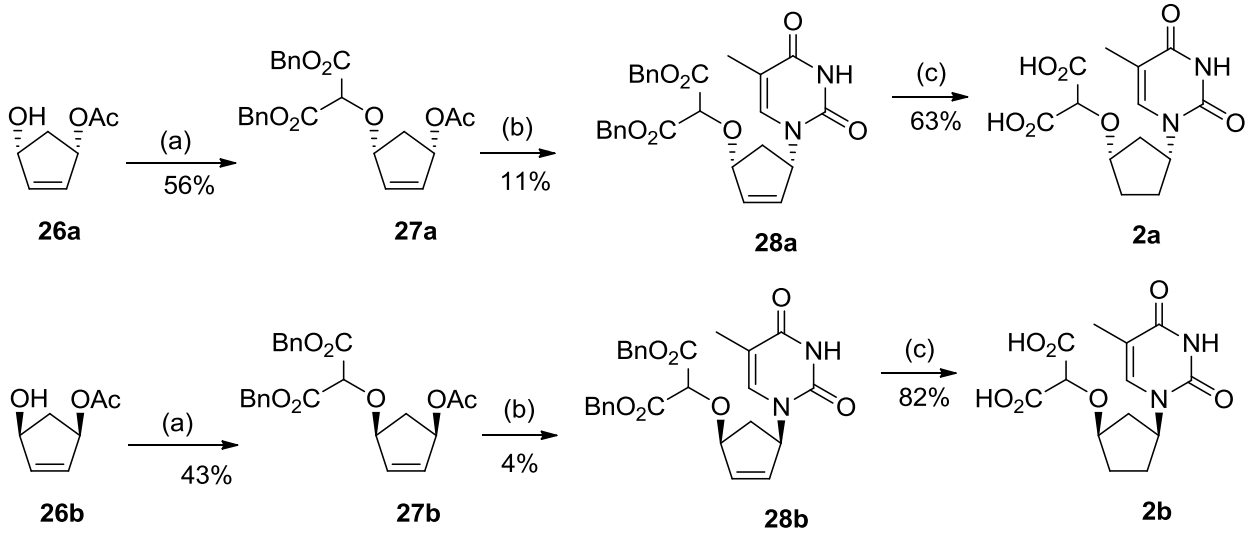
[a]Conditions: (a) ABSA, K_2CO_3, MeCN, 48 h; (b) 26, $Rh_2(OAc)_4$, C_6H_6, 80 °C, 12 h; (c) thymine, $Pd(dba)_2$, dppb, Na_2CO_3, aq. MeCN, Microwave (MW) 50 °C; (d) H_2, Pd/C 1 atm, MeOH, 15 h.

The synthesis of **2** was accomplished in four steps from commercially available dibenzyl malonate (Scheme 1). Diazo transfer to **24** was achieved in the presence of 4-acetamidobenzenesulfonyl azide (ABSA) to give the novel dibenzyl 2-diazomalonate **25**. Building upon our previous work with rhodium insertion reactions,[30,31] rhodium (II)-catalyzed OH insertion of **25** with the allylic alcohol **26**[32-34] furnished **27** in 48% yield. Thymine base insertion of **27** in a palladium-catalyzed Tsuji-Trost reaction[35] was carried out under microwave conditions to afford **28**. Finally, simultaneous hydrogenation of the cyclopentyl alkene and hydrogenolysis of the benzyl esters of **28** was achieved under atmospheric pressure of hydrogen in the presence of Pd/C to provide the target malonate **2** in racemic form (57% yield). Despite our initial concerns that **2** may undergo decarboxylation during isolation or upon standing, a solution of **2** in MeOH-d_4 was found to be stable at room temperature for up to 6 months. We

have already described the synthesis and characterization of the decarboxylated compound **29**[4] and were unable to detect its presence by NMR in the solution of **2** after 6 months. Furthermore, storage of malonate **2** in solid state proved easier than anticipated, with little or no evidence of decarboxylation on storage for several months.

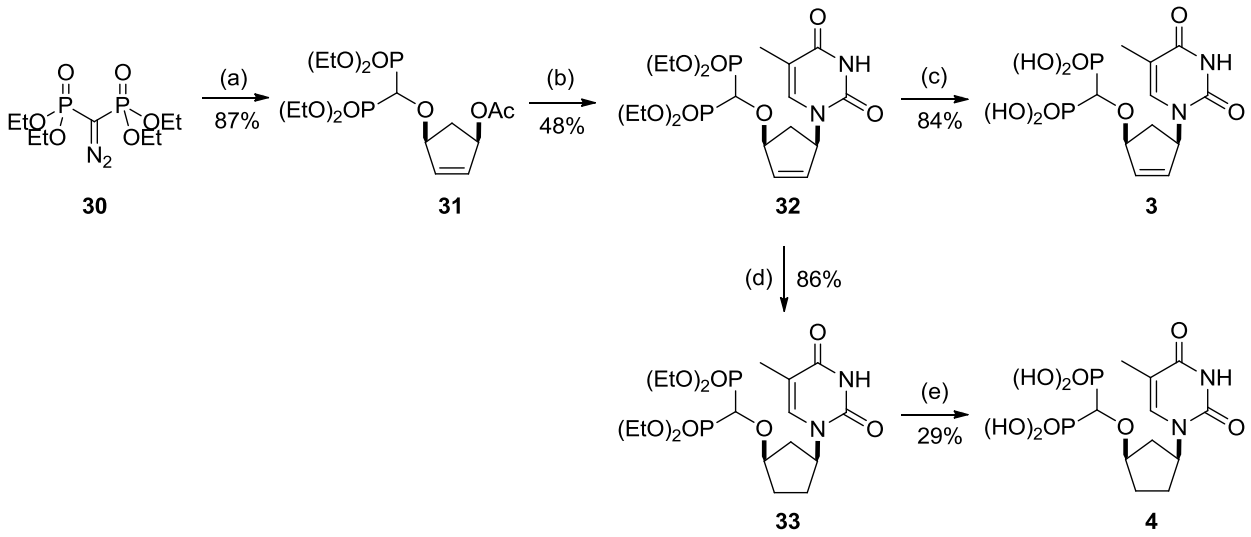
The malonates were also synthesized in both enantiopure forms **2a** and **2b**. This was achieved using the same sequence of reactions described for the racemic series starting from the commercially available enantiomers of the acetoxy alcohol **26a** and **26b** (Scheme 2). Yields were consistent with those obtained for the racemic series.

Scheme 2. Synthesis of enantiopure forms of **2**.[a]



[a]Conditions: (a) **25**, $Rh_2(OAc)_4$, C_6H_6, 80 °C; (b) thymine, $Pd(dba)_2$, dppb, Na_2CO_3, aq. MeCN, MW 50 °C; (c) H_2, Pd/C 1 atm, MeOH.

Scheme 3. Synthesis of bisphosphonates **3** and **4**.[a]

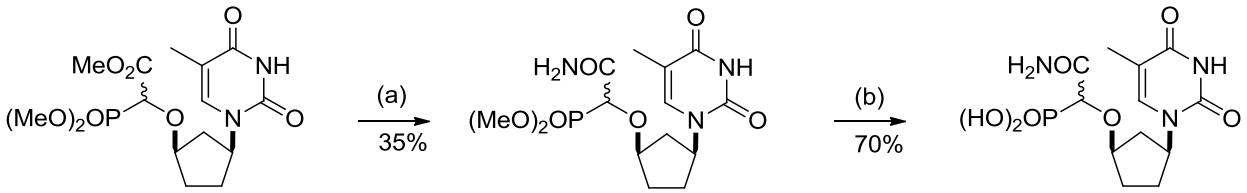


[a]Conditions: (a) **26**, $Rh_2(tfacam)_4$, C_6H_6, 80 °C; (b) thymine, $Pd(dba)_2$, dppb, Na_2CO_3, aq. MeCN, 50 °C MW; (c) (i) TMSBr, lutidine, MeCN, MW 50 °C, (ii) MeOH/H_2O 1 h, (iii) charcoal chromatography; (d) H_2, Pd/C 1 atm, MeOH, 15 h; (e) (i) TMSBr, MeCN, MW 50 °C, (ii) MeOH/H_2O 1 h, (iii) charcoal chromatography.

A similar strategy was employed for the preparation of the racemic bisphosphonates **3** and **4** (Scheme 3). It was envisaged that O-H insertion of tetraethyl diazomethylenebisphosphonate **30**[36] with alcohol **26** would require more forcing conditions than the analogous reaction with diazomalonate **25**. Therefore, in this case, the more active rhodium trifluoroacetamide was employed as catalyst and it was added portionwise over the course of the reaction. Gratifyingly, under these conditions, the O-H insertion product **31** was obtained in 87% yield after correction for the recovered starting material **26**. Introduction of thymine under Tsuji-Trost conditions proceeded efficiently in the microwave to afford bisphosphonate **32** in 48% yield. Finally, cleavage of the phosphonate ethyl esters of **32** was achieved using bromotrimethylsilane (TMSBr) in the presence of lutidine and the desired unsaturated bisphosphonate **3** was isolated as its ammonium salt in 84% yield following charcoal chromatography. Interestingly both the O-H insertion and base insertion steps were effective despite the alteration from the phosphonoacetic

acid moiety to the malonate and bisphosphonate confirming the generality of the synthetic route developed for the α-CNPs. Palladium mediated hydrogenation of **32** afforded **33** and subsequent hydrolysis of the phosphonate esters of **33** followed by charcoal chromatography gave **4** as the ammonium salt.

Scheme 4. Synthesis of amide **5**.[a]



34 **35** **5**

[a]Conditions: (a) 7N NH_3 in MeOH, 65 h; (b) (i) TMSBr, MeCN, MW 50 °C, (ii) MeOH/H_2O 30 min, (iii) charcoal chromatography.

The synthesis of the amide (Scheme 4) derivative **5** commenced with the methyl ester **34**, which was also a key intermediate in the synthesis of α-CNPs which we described earlier.[4] Aminolysis of the carboxy methyl ester was achieved by stirring **34** for 65 h in a 7N solution of ammonia in methanol to afford **35** in 35% yield. The low yield may be explained by the fact that attack at the phosphonate ester groups is also possible under the reaction conditions. Following isolation of **35** by chromatography, the phosphonate esters were hydrolyzed and **5** was isolated as the ammonium salt in 70% yield.

The eight structurally related compounds (malonates **2**, **2a**, **2b**, bisphosphonates **3** and **4**, amide **5** and esters **6** and **7**) were evaluated for their inhibitory activity using a cell-free HIV-1-RT assay and the results compared with those obtained for the parent α-CNPs (**1**, **1a**, **1b** and **36**).[4,5] Poly rA.oligo dT was used as the homopolymeric template/primer, and [^{3}H]dTTP as the radiolabeled

substrate. The IC_{50} (50% inhibitory concentration) was determined as the compound concentration required to inhibit RT-catalysed dTTP incorporation into the growing DNA strain.

For the parent α-CNPs, our research has already shown that the "L"-enantiomer **1a** is the more active of the two enantiomeric compounds (Table 1, Entry 2 vs. 3). The effect of unsaturation of the racemic α-CNPs (i.e. compound **36** compared to **1**) is to decrease the activity by ~ 5-fold (Table 1, Entry 4 vs. 1).[4]

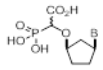
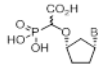
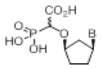
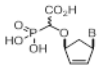
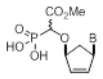
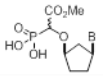
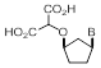
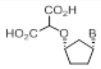
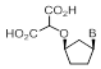
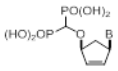
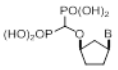
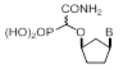
Importantly, the presence of a methyl ester in place of the carboxylic acid effectively shuts down activity as evidenced by results obtained for **6** and **7**. This effect is further highlighted by the amide **5** which shows poor, if any, relevant inhibitory activity against HIV-1 RT (Table 1, Entry 12). The lack of significant activity of the amide and esters **6** and **7** confirms that two ionizable moieties are required to efficiently coordinate the magnesium ions to display efficient RT inhibition.

The unsaturated bisphosphonate **3** is poorly active (IC_{50}: 204 μM) (Table 1, Entry 10), while the saturated analog **4** is similarly ineffective to inhibit HIV-1 RT.

The racemic malonate **2** (Table 1, Entry 7) displays an activity approximately 100-fold less than observed for racemic α-CNP **1** (Table 1, Entry 1). However, the results obtained for the enantiopure malonates are interesting in terms of mirroring the behavior observed for the α-CNPs. The malonate **2a** ("L"-enantiomer) is 2-fold more active (IC_{50}: 20 μM) than the racemic mixture **2** (IC_{50}: 42 μM) [Table 1, Entry 8 vs. Entry 7], while the "D"-enantiomer lacked any relevant inhibitory activity against HIV-1 RT (IC_{50}: > 640 μM) [Table 1, Entry 9]. Such a selective activity profile (predominant activity by the "L"-enantiomer) has also been observed for the prototype α-CNPs as mentioned above. Also, it should be noticed that the malonate **2a**, as

also found for the α-CNPs, does not need metabolic conversion to a higher phosphate metabolite to exert anti-HIV-1 RT activity.

Table 1. Inhibitory activity of the compounds against HIV-1 RT

Entry	Structure[a]	Compound	IC_{50}[b] (μM)
1[c]		1	0.41±0.08
2[c]		1a	0.44 ± 0.06
3[c]		1b	35±5
4[c]		36	1.9 ± 1.1
5[c]		6	>100
6[c]		7	>40
7		2	42 ± 3
8		2a	20±5
9		2b	> 640
10		3	204±102
11		4	>270±30
12		5	223 ± 67

[a]Base = Thymine. [b]Fifty percent inhibitory concentration, or compound concentration required to inhibit HIV-RT-catalyzed incorporation of [^{3}H]dNTP in the homopolymeric template/primer. Data are the mean ± SD of at least 3 to 4 independent experiments. [c]Synthesis (**1**, **1a**, **1b**, **36**, **6** and **7**) and IC_{50} data (**1**, **1a**, **1b**, **36**) see *Ref 4*.

Finally, when the active malonate enantiomer **2a** was investigated for its kinetic interaction with HIV-1 RT, reciprocal velocity *versus* substrate Lineweaver-Burk plots revealed a competitive inhibition of the enzyme by **2a** with respect to the natural substrate dTTP (Fig. 3). The inhibitory constant Ki of **2a** was 6.6 μM. The Km of dTTP was 5.9 μM. This kinetic behavior was also earlier observed for the prototype α-CNP compound **1**.[5]

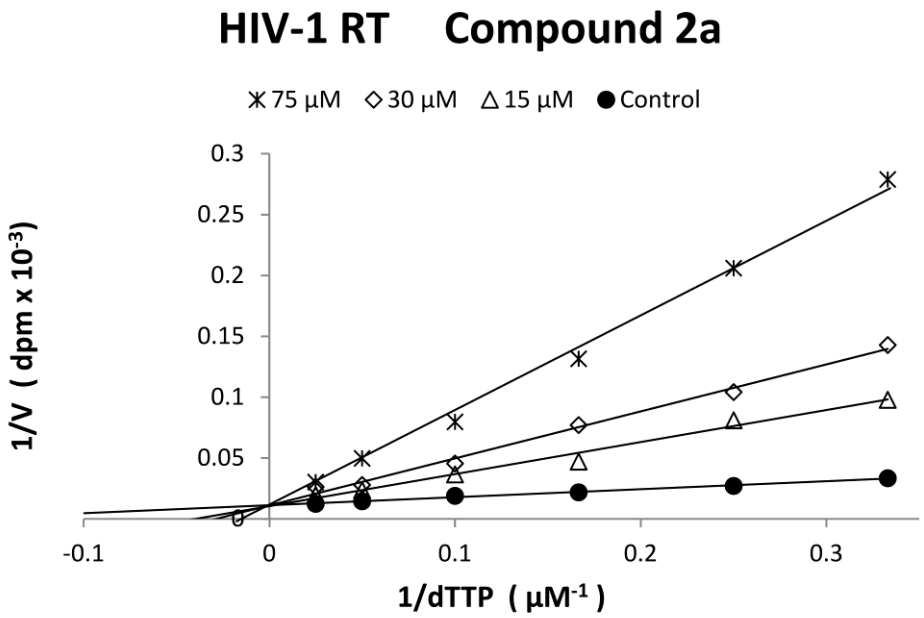


Figure 3. Lineweaver-Burk plots for the inhibition of HIV-1 RT by compound **2a**.

A crystal structure of **2a** was determined in complex with HIV-1 RT/38-mer DNA aptamer;[37] the 38-mer DNA template-primer aptamer mimics the binding of a double-stranded (ds) DNA at the polymerase active site, and contains a single-stranded (ss) DNA template overhang.[37] The pre-formed crystals were soaked in a 2mM solution of **2a** for the formation of RT/DNA/**2a** ternary complex, and the crystal structure revealed the binding of **2a** at the polymerase active site (Fig. 4). The structure contains two copies of RT/DNA, and only the first one has ordered electron density for **2a**. The general mode of binding of **2a** is similar to that of **1a** (T-α-CNP).[5] The thymine base of **2a** is base-paired with the adenine base of 1st overhang nucleotide, and the

cyclopentyl ring is positioned over the aromatic side chain of Tyr115; both interactions of **2a** are analogous to those observed with the binding of **1a** (T-α-CNP) to RT. However, metal chelation at the polymerase active site differs for the two inhibitors. One oxygen from each of the carboxyl groups of **2a** chelate with a Mg^{2+} ion (metal B) at the polymerase active site, and the ion also chelates to the main-chain carbonyl oxygen of V111 and a side-chain oxygen of the catalytic aspartate Asp185. Interestingly, the catalytic residue Asp110 is switched away from the active site, and does not participate in the Mg^{2+} chelation in the structure of the **2a** complex. Presumably, the chelation of the malonate group of **2a** is not compatible with the chelation of all three catalytic aspartates (Asp110, Asp185, and Asp186), which is observed for the binding of a dNTP and for **1a** (T-α-CNP).[5] The binding of **2a** involves only 4 observed chelations for the Mg^{2+} ion B in contrast to a complete octahedral coordination environment when T-α-CNP **1a** was bound. We expect that two water molecules might be involved to complete the coordination environment for the Mg^{2+} ion when **2a** is bound, however, no clear electron density was observed for any potentially chelating ordered water molecule in the crystal structure of RT/DNA/**2a** complex. In summary, the structure suggests that the significantly perturbed coordination environment due to the malonate group is a primary contributor to the lower RT inhibition by **2a** compared to the inhibition by **1a** (T-α-CNP). Thereby, a potent CNP class of nucleoside competing RT inhibitor (NcRTI) should chelate the catalytic Mg^{2+} ions in a mode analogous to that observed in the chelation of a dNTP substrate.

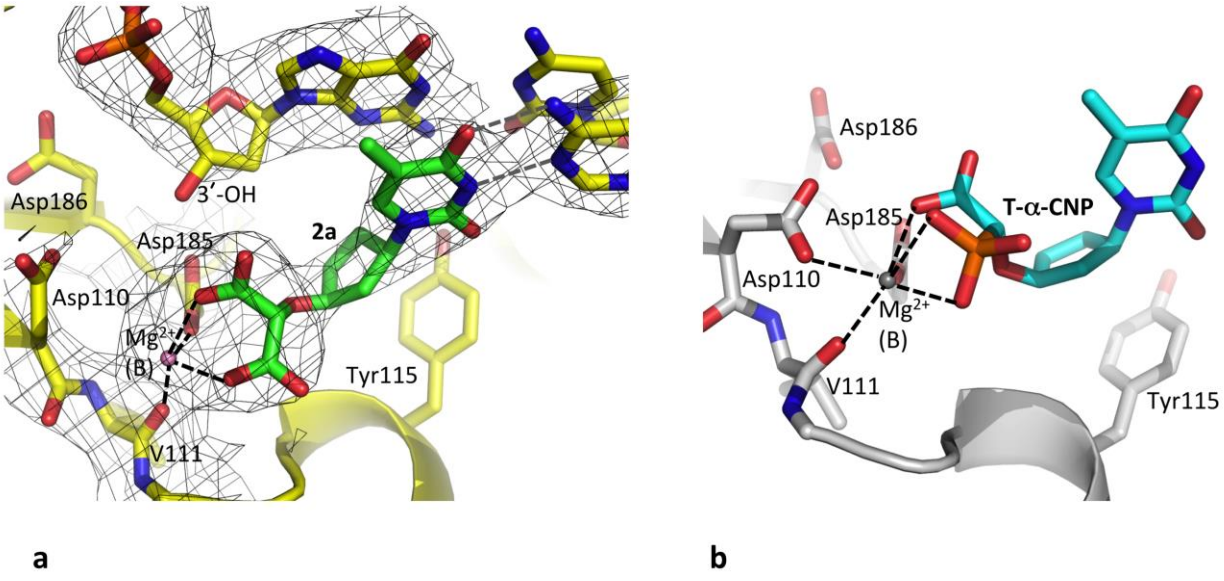


Figure 4. Structure of **2a** in complex with HIV-1 RT/DNA. (**a**) 2.95 Å resolution electron density map defines the binding of **2a** (green) at the polymerase active site of RT. The inhibitor binding affinity has contributions from the base pairing (gray dotted lines) and base stacking, hydrophobic stacking of the cyclopentyl ring with the aromatic side chain of Tyr115, and the chelation with a Mg^{2+} ion at the polymerase active site (black dotted lines). The chelation environment for **2a** is significantly different from the octahedral coordination environment observed for **1a** (T-α-CNP) (**b**).

When evaluated for their anti-HIV activity in cell culture, the compounds were found to be antivirally inactive at concentrations up to 100 μM. Also, they were devoid of any significant cytotoxicity. Presumably, the compounds might be too polar to be efficiently taken up by intact cells. Therefore, a prodrug approach should be considered to deliver the active molecules directly into the intact virus-infected cells. The design and synthesis of such prodrug derivatives are currently under investigation.

CONCLUSION

We designed and synthesized structurally related analogues of α-CNPs and investigated their ability to inhibit HIV-1 RT in cell-free enzyme assays. The amide derivative **5** was virtually inactive while the bisphosphonates **3** and **4** also exhibited poor inhibition of HIV-1 RT. The requirement for two ionizable functional groups to be present in order to exhibit inhibition was demonstrated through the lack of activity of both esters **6** and **7** and also the amide **5**. The malonate **2** was the most active of the compounds albeit significantly less active that the original α-CNPs. Enantiopure forms of the malonate behaved in a similar fashion to the parent α-CNP in that the activity resides in the "L"-enantiomer. The crystal structure together with the RT inhibition data confirms the significance of the α-carboxy phosphonoacetic acid moiety in affording efficient RT inhibition without need for any further phosphorylation.

EXPERIMENTAL SECTION

General Information

Solvents were distilled prior to use as follows: dichloromethane was distilled from phosphorus pentoxide; ethyl acetate was distilled from potassium carbonate. Benzene was dried before use with activated 4Å molecular sieves. For O-H and base insertion reactions, solvents were degassed by purging with nitrogen for 5 min. Organic phases were dried using anhydrous magnesium sulfate. All commercial reagents were used without further purification. Microwave reactions were carried out using a CEM Discover in conjunction with Synergy software, reaction temperatures were measured by IR sensor. ^{1}H, ^{13}C, and ^{31}P spectra were recorded at 20 °C on

300 or 400 MHz spectrometers. Chemical shifts are given in ppm relative to tetramethylsilane (TMS) as an internal standard. ^{31}P chemical shifts are referenced to H_3PO_4 (external standard). Coupling constants (*J*) are given in hertz (Hz). In some cases the signal for the PCH was not observed in the carbon-13 NMR spectrum. Infrared spectra were recorded neat on a PerkinElmer Spectrum Two; operating in Universal Attenuated Total Reflectance (UATR) mode. Melting points were measured using a capillary melting point apparatus and are not corrected. Optical rotations were measured at 20 °C at 589 nm in a 10 cm cell; concentrations (c) are expressed in g/100 ml, [α] is expressed in units of 10^{-1} deg cm^2 g^{-1}. Low resolution mass spectra were recorded in electrospray ionization (ESI) mode, high resolution mass spectra (HRMS) were recorded on a Time of Flight spectrometer in electrospray ionization (ESI) mode. Column chromatography was performed using silica gel 60. Thin layer chromatography (TLC) was carried out on precoated silica gel plates (60 PF254). Visualization was achieved by UV (254 nm) detection and/or staining with vanillin or permanganate. Compounds **6**,[4] **7**,[4] **26**[32-34] and **30**[36] were prepared by literature methods.

Reverse transcription assays

HIV-I RT assays were carried out in the presence of homopolymeric template/primers. Poly(A) and dT12-18 were from Pharmacia (Uppsala, Sweden). To prepare the template/primers for the RT experiments, 0.15 mM poly(A) was mixed with an equal volume of 0.0375 mM oligo(dT). The reaction mixture (50 µl) contained 50 mM Tris.HC1 pH 7.8, 5 mM dithiothreitol, 300 mM glutathione, 500 µM EDTA, 150 mM KCl, 5 mM $MgCl_2$, 1.25 µg of bovine serum albumin, an appropriate concentration of the tritium-labeled substrate [CH3-3H]dTTP (1 µCi/assay; 60 Ci/mmole), a fixed concentration of the template/primer poly(A).oligo(dT) (0.015 mM), 0.06% Triton X-100, 10 µl of α-CNP inhibitor solution (containing various concentrations of the

compounds), and 1 µl (being equivalent to 1 pmole (0.13 ng) enzyme or 20 pM final concentration) of the HIV-1 RT preparation. The reaction mixtures were incubated at 37°C for 30 minutes, at which time 100 µl of calf thymus DNA (150 µg/ml), 2 ml of $Na_4P_2O_7$ (0.1 M in 1 M HCl), and 2 ml of trichloroacetic acid (10% v/v) were added. The solutions were kept on ice for 30 minutes, after which the acid-insoluble material was washed and analyzed for radioactivity. For the experiments in which the 50% inhibitory concentration (IC_{50}) of the test compounds was determined, a fixed concentration of 1.25 µM [^{3}H]dTTP was used. For the experiments in which the kinetic nature of the malonate derivative **2a** was determined, different [^{3}H]dTTP concentrations (3, 4, 6, 10, 20 and 40 µM) were used in the presence of different inhibitor concentrations (75, 30 and 15 µM). The activity values were plotted in a reciprocal velocity *versus* substrate Lineweaver-Burk diagram. It should be emphasized that under the experimental conditions, the RT reaction was linear for up to 30 min, and at the end of the incubation period, less than 5% of the available dTTP substrate was consumed (incorporated).

X-ray crystallography

The crystals of RT/38-mer DNA aptamer complex were obtained using a previously defined protocol.[37] A crystal soaked in a solution containing 2.5mM of 2 and 20mM $MgCl_2$ for 15 minutes diffracted the best up to 2.95Å resolution (Supplementary Information Table-1). X-ray data were collected at APS 23ID-D beam line. Due to anisotropy diffractions and radiation sensitivity of the crystal, the collected data is only 84.8% complete with 87.5% in the highest resolution cell. The dataset was processed using Mosflm[38] and the RT/38-mer DNA aptamer binary complex structure (PDB Id. 5D3G) was used as the starting model for obtaining the molecular replacement solution. There are two copies of RT/DNA complex in the asymmetric unit. The difference (|Fo| - |Fc|) electron density map clearly revealed the binding of the

compound **2** at the polymerase active site of one of two RT/DNA copies in the asymmetric unit. The crystallography software Phenix[39] and Coot[40] were used for structure refinement and model building, respectively. Final refinement statistics are listed in Supplementary Information Table-1, and the coordinates and structure factors are deposited in PDB with the accession number 5HLF.

Synthesis

Dibenzyl 2-diazomalonate (25)

Potassium carbonate (3.34 g, 0.024 mol) was added to a solution of dibenzyl malonate (3.19 g, 0.011 mol) and 4-acetamidobenzenesulfonyl azide (2.65 g, 0.011 mol) in MeCN (50 mL). The resulting suspension was stirred for 24 h at room temperature, filtered and concentrated *in vacuo*. The residue was taken up in DCM (50 mL), and the solution was washed with aqueous NaOH (50 mL, 1M), water (50 mL) and brine (50 mL), dried ($MgSO_4$) and concentrated *in vacuo*. The residue was purified by chromatography (SiO_2, ether) to afford the title compound as a white solid (2.21 g, 64%): m.p. = 56-57 °C; ν_{max}/cm^{-1}: 3407, 2940, 2145, 2129, 1758, 1708, 1690; ^1H NMR (300 MHz, $CDCl_3$) δ: 5.27 (s, 2H), 7.29-7.39 (m, 5H); ^{13}C NMR (75.5 MHz, $CDCl_3$) δ: 67.1, 128.2, 128.4, 128.6, 135.3, 160.8; signal for **C**N_2 not seen; Anal. calcd for $C_{17}H_{14}N_2O_4$: C. 65.80; H, 4.55; N, 9.03. Found: C, 65.82; H, 4.62; N, 8.83; HRMS (ES+): calcd for $C_{17}H_{15}N_2O_4$ $(M+H)^+$ 311.1032, found 311.1021; MS (ES+) *m/z*: 311 $(M + H)^+$.

(−)-(1S, 4R)-Dibenzyl 2-[(1-acetoxycyclopent-2-en-4-yl)oxy]malonate (27a)

Rhodium(II) acetate (15 mg, 0.033 mmol, 1 mol%) was added to a degassed solution of (1*S*,4*R*)-*cis*-4-hydroxy-2-cyclopentenyl acetate **26a** (0.50 g, 3.52 mmol) and **25** (1.20 g, 3.87 mmol) in benzene (25 mL). The reaction mixture was stirred while heating under reflux for 24 h under a nitrogen atmosphere. The mixture was filtered, concentrated and the residue purified by

chromatography (SiO_2, 20% ethyl acetate/hexane) to give **(−)-(1*S*, 4*R*)-27a** as a pale yellow oil (0.82 g, 56%): ν_{max}/cm^{-1}: 2925, 1730, 1608, 1587, 1235; $[\alpha]_D^{20}$ - 24.38 (*c* 0.08, $CHCl_3$); ^1H NMR (300 MHz, $CDCl_3$) δ: 1.73-1.84 (m, 1H), 1.99 (s, 3H), 2.67-2.81 (m, 1H), 4.62-4.72 (m, 2H), 5.19 (s, 4H), 5.39-5.47 (m, 1H), 5.94-6.16 (m, 2H), 7.19-7.41 (m, 10H); ^{13}C NMR (75.5 MHz, $CDCl_3$) δ: 21.0, 37.1, 67.6, 76.2, 76.8, 82.9, 128.3, 128.53, 128.59, 134.3, 134.8, 135.0, 166.6, 170.6; HRMS (ES+): calcd for $C_{24}H_{24}NaO_7$ $(M+Na)^+$ 447.1420, found 447.1426.

(+)-(1*R*, 4*S*)-Dibenzyl 2-[(1-acetoxycyclopent-2-en-4-yl)oxy]malonate (27b)

Synthesized following the procedure described for **27a** using: rhodium(II) acetate (15 mg, 0.033 mmol, 1 mol%), (1*R*,4*S*)-*cis*-4-hydroxy-2-cyclopentenyl acetate **26b** (0.50 g, 3.52 mmol) and **25** (1.20 g, 3.87 mmol). Purification by chromatography (SiO_2, 20% ethyl acetate/hexane) afforded **(+)-(1*R*, 4*S*)-27b** as a colorless oil (0.65 g, 43%): $[\alpha]_D^{20}$ + 20.00 (*c* 0.05, $CHCl_3$).

***cis*-Dibenzyl 2-[(1-acetoxycyclopent-2-en-4-yl)oxy]malonate (27)**

Synthesized following the procedure described for **27a** using: rhodium(II) acetate (15 mg, 0.033 mmol, 0.8 mol%) , *cis*-4-hydroxy-2-cyclopentenyl acetate **26** (0.58 g, 4.08 mmol) and **25** (1.40 g, 4.51 mmol). Purification by chromatography (SiO_2, 20% ethyl acetate/hexane) gave **27** as a pale yellow oil (0.56 g, 32%, 48% when corrected for recovered alcohol **26**).

(−)-(1*S*, 4*R*)-Dibenzyl 2-[(1-thymin-1-yl-cyclopent-2-en-4-yl)oxy]malonate (28a)

A microwave vial containing a degassed suspension of thymine (190 mg, 1.51 mmol) and sodium carbonate (115 mg, 1.07 mmol) in water (1 mL) and acetonitrile (1 mL) was heated under microwave conditions (50°C, 200 W) for 30 min. A degassed solution of **(−)-(1*S*, 4*R*)-27a** (420 mg, 0.99 mmol) in acetonitrile (1 mL), bis(dibenzylideneacetone)palladium(0) $Pd(dba)_2$ (30

mg, 5 mol%) and 1,4-bis(diphenylphosphino)butane (dppb) (40 mg, 9 mol%) was added to the vial. The resulting solution was irradiated (50°C, 200 W) for 30 min whereupon a second portion of $Pd(dba)_2$ (30 mg) and dppb (40 mg) was added followed by irradiation (50°C, 200 W) for a further 30 min. The reaction mixture was cooled to room temperature, gravity filtered and concentrated under vacuum to give a purple residue which was purified by chromatography (SiO_2, 3% methanol/dichloromethane) to afford compound **(−)-(1*S*, 4*R*)-28a** as an oil (52 mg, 11%): ν_{max}/cm^{-1}: 3177, 3034, 2924, 1762, 1742, 1686; ^1H NMR (300 MHz, $CDCl_3$) δ: 1.63-1.78 (m, 1H), 2.64-2.86 (m, 1H), 4.56-4.67 (m, 1H), 4.71 (s, 1H), 5.06-5.32 (m, 4H), 5.55-5.75 (m, 1H), 5.80-5.96 [m, 1H), 6.15-6.29 (m, 1H), 7.21-7.39 (m, 11H), 9.31-9.52 (bs, 1H); ^{13}C NMR (75.5 MHz, $CDCl_3$) δ: 12.3, 37.2, 57.7, 67.8, 77.9, 83.4, 111.6, 128.41, 128.44, 128.63, 128.64, 128.69, 134.4, 134.65, 134.69, 135.9, 137.1, 151.1, 164.0, 166.1, 166.3; HRMS (ES+): calcd for $C_{27}H_{27}N_2O_7$ $(M+H)^+$ 491.1818, found 491.1829; MS (ES+) *m/z*: 491 $(M + H)^+$.

(+)-(1*R*, 4*S*)-Dibenzyl 2-[(1-thymin-1-yl-cyclopent-2-en-4-yl)oxy]malonate (28b)

Prepared following the procedure described for **28a** using: thymine (190 mg, 1.51 mmol), sodium carbonate (115 mg, 1.07 mmol), **(+)-(1*R*, 4*S*)-27b** (420 mg, 0.99 mmol), $Pd(dba)_2$ (30 mg, 5 mol%) and dppb (40 mg, 9 mol%). Purification by chromatography (SiO_2, 3% methanol/dichloromethane) gave **(+)-(1*R*, 4*S*)-28b** as an oil (18 mg, 4%).

***cis*-Dibenzyl 2-[(1-thymin-1-yl-cyclopent-2-en-4-yl)oxy]malonate (28)**

Prepared following the procedure described for **28a** using: thymine (160 mg, 1.27 mmol), sodium carbonate (97 mg, 0.92 mmol), **27** (355 mg, 0.84 mmol), $Pd(dba)_2$ (25 mg, 5 mol%) and dppb (35 mg, 10 mol%). Purification by chromatography afforded compound **28** as an oil (25 mg, 6%).

(+)-(1*R*, 4*S*)- 2-[(1-thymin-1-yl-cyclopentan-4-yl)oxy]malonic acid (2a)

A mixture of **(−)-(1*S*, 4*R*)-28a** (30 mg, 0.061 mmol) and palladium on carbon (10 mg, 10%) in methanol (10 mL) was stirred under a balloon of hydrogen at atmospheric pressure for 24 h at room temperature. The mixture was filtered on Celite and the cake was rinsed with methanol. The filtrate was concentrated to give **(+)-(1*R*, 4*S*)-2a** (12 mg, 63%): $[\alpha]_D^{20}$ + 4.17 (*c* 0.24, MeOH); ^{1}H NMR (300 MHz, MeOH-D_4) δ: 1.48-2.15 (m, 8H), 2.19-2.36 (m, 1H), 4.02-4.17 (m, 1H), 4.55 (s, 1H), 5.01-5.18 (m, 1H), 7.88 (s, 1H); ^{13}C NMR (75.5 MHz, MeOH-d4) δ: 10.8, 29.7, 30.7, 37.8, 53.6, 76.5, 80.5, 110.7, 139.2, 151.8, 165.1, 169.0; HRMS (ES+): calcd for $C_{13}H_{17}N_2O_7$ $(M+H)^+$ 313.1036, found 313.1030; MS (ES+) *m/z*: 313 $(M + H)^+$.

(−)-(1*S*, 4*R*)- 2-((1-thymin-1-yl-cyclopentan-4-yl)oxy)malonic acid (2b)

Synthesized using procedure described for **2a** using: **(+)-(1*R*, 4*S*)-28b** (25 mg, 0.051 mmol) and palladium on carbon (10 mg, 10%) to give **(−)-(1*S*, 4*R*)-2b** (13 mg, 82%): $[\alpha]_D^{20}$ - 6.67 (*c* 0.06, MeOH).

***cis*-2-[(1-thymin-1-yl-cyclopentan-4-yl)oxy]malonic acid (2)**

Synthesized using procedure described for **2a** using: **28** (25 mg, 0.051 mmol) and palladium on carbon (10 mg, 10%) to give **2** (9 mg, 57%).

***cis*-1-[bis(Diethylphosphono)methoxy]-4-acetoxycyclopent-2-ene (31)**

cis-4-Hydroxy-2-cyclopentenyl acetate **26** (0.69 g, 4.7 mmol) and tetraethyl diazomethylenebisphosphonate **30** (2.95 g, 9.4 mmol) were dissolved in degassed benzene (15 mL). Rhodium (II) trifluoroacetamide (30 mg, 1 mol %) was added and the reaction mixture was refluxed for 18 h, after which time a second portion (30 mg, 1 mol %) of catalyst was added and reflux was continued for a further 24 h and then a third portion (30 mg, 1 mol %) was added and reflux was continued for a further 28 h. After removal of benzene under reduced pressure the residue was subjected to flash chromatography (EtOAc to 5% MeOH:95% EtOAc) to afford starting material (0.20 g, 29%) then the title bisphosphonate **31** as an oil (1.25 g, 62%); ν_{max}/cm^{-1} 2984, 1733, 1237, 1013; ^{1}H NMR (300 MHz, $CDCl_3$) δ: 1.36 (t, J = 7.1, 12H), 1.90 (dt, J = 14.5, 4.2, 1H), 2.04 (s, 3H), 2.79 (overlapping dt, J = 14.5, 7.4, 1H), 4.11 (t, J_{PH} = 17.6, 1H), 4.13-4.33 (m, 8H), 4.76-4.85 (m, 1H), 5.44-5.53 (m, 1H), 5.99-6.06 (m, 1H), 6.16-6.23 (m, 1H); ^{13}C NMR (75.5 MHz, $CDCl_3$) δ: (75.5 MHz, $CDCl_3$) 16.2-16.5 (m), 21.0, 37.0, 63.2-63.7 (m), 71.8 (t, J_{PC} = 157.8), 76.4, 85.6 (t, J_{PC} = 4.7), 133.6, 135.3, 170.6; ^{31}P NMR (121.5 MHz, $CDCl_3$) δ: 15.65, 15.80 (q, J = 31.7); HRMS (ES+) Exact mass calculated for $C_{16}H_{31}O_9P_2$ $[M+H]^+$ 429.1443; found 429.1448.

***cis*-1-{4-[bis(Diethylphosphono)methoxy]cyclopent-2-en-1-yl}thymine (32)**

Thymine (95 mg, 0.75 mmol) and sodium carbonate (80 mg, 0.75 mmol) were suspended in degassed water (0.8 mL) and degassed acetonitrile (0.8 mL) in a microwave vial and the mixture was subjected to microwave irradiation at 50 °C for 30 mins. Bisphosphonate **31** (270 mg, 0.63 mmol), $Pd(dba)_2$ (18 mg, 5 mol %) and dppb (27 mg, 10 mol %) were suspended in degassed acetonitrile (1 mL), transferred to the microwave vial and the resultant mixture was irradiated at 50 °C for 30 mins. TLC analysis showed residual starting material, so a second portion of catalysts was added together with more acetonitrile (1 mL) and the reaction mixture was

irradiated for a further 30 mins at 50 °C. After cooling, dichloromethane (5 mL) was added and the mixture was filtered through Celite, washed with dichloromethane (20 mL) and the filtrate was concentrated under reduced pressure to give a residue which was purified by flash chromatography (5-10% MeOH in EtOAc) to afford **32** as a foamy solid (148 mg, 48 %); ν_{max}/cm^{-1} 3169, 2983, 1683, 1251, 1015; ^{1}H NMR (400 MHz, $CDCl_3$) δ: 1.28-1.45 (m, 12H), 1.86 (dt, J = 15.2, 2.8, 1H), 1.92 (s, 3H), 2.81 (overlapping dt, J = 15.6, 7.9, 1H), 4.14 (t, J_{PH} = 17.8, 1H), 4.17-4.36 (m, 8H), 4.85-4.93 (m, 1H), 5.63-5.73 (m, 1H), 5.87-5.95 (m, 1H), 6.29-6.38 (m, 1H), 7.23 (s, 1H), 9.79 (s, 1H); ^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 12.3, 16.0-16.7 (m), 36.7, 57.8, 63.1-63.6 (m), 72.0 (t, J_{PC} = 158.1), 86.1 (t, J_{PC} = 4.5), 111.2, 134.0, 136.1, 137.2, 151.2, 164.1; ^{31}P NMR (162.0 MHz, $CDCl_3$) δ: 15.72, 15.92 (ABq, J_{AB} = 31.1); HRMS (ES+) Exact mass calculated for $C_{19}H_{33}N_2O_9P_2$ $[M+H]^+$ 495.1661; found 495.1660.

***cis*-1-{4-[(Diphosphono)methoxy]cyclopent-2-en-1-yl}thymine (3)**

Bisphosphonate **32** (32 mg, 0.065 mmol) was suspended in acetonitrile (1.5 mL) in a microwave vial and 2,6-lutidine (75 µL, 0.64 mmol) then TMSBr (85 µL, 0.64 mmol) were added. The reaction mixture was irradiated at 50 °C for 30 mins then methanol (1.5 mL) and water (100 µL) were added and stirring was continued for 1 h. Solvents were removed under reduced pressure and the residue was subjected to charcoal chromatography, eluting with water then 20% aqueous ammonia, to give the title compound as the ammonium salt (22 mg, 84%); ^{1}H NMR (400 MHz, D_2O) δ: 1.64-1.78 (m, 1H), 1.83 (s, 3H), 2.81-3.02 (m, 1H), 3.67-3.98 (m, 1H), 4.85-4.95 (m, 1H), 5.36-5.46 (m, 1H), 5.80-5.90 (m, 1H), 6.31-6.42 (m, 1H), 7.49 (s, 1H); ^{13}C NMR (100.6 MHz, D_2O) δ: 11.3, 37.2, 59.1, 84.9-85.1 (m), 111.3, 131.9, 137.3, 140.0, 152.4, 166.8, (P**C**H

not seen); ^{31}P NMR (162.0 MHz, D_2O) δ: 13.4; HRMS (ES-) Exact mass calculated for $C_{11}H_{15}N_2O_9P_2$ $[M-H]^-$ 381.0253; found 381.0240.

***cis*-1-{4-[bis(Diethylphosphono)methoxy]cyclopentan-1-yl}thymine (33)**

cis-1-{4-[bis(Diethylphosphono)methoxy]cyclopent-2-en-1-yl}thymine (123 mg, 0.25 mmol) was dissolved in methanol (10 mL), 5% palladium on carbon (40 mg) was added and the mixture was stirred under hydrogen for 17 h then filtered through Celite. After washing with MeOH (2 x 10 mL) then dichloromethane (2 x 10 mL) the filtrate was concentrated under reduced pressure and the residue subjected to flash chromatography (1:9, MeOH:EtOAc) to afford the title product (106 mg, 86%); ν_{max}/cm^{-1} 3171, 2982, 1667, 1240, 1011; ^{1}H NMR (400 MHz, $CDCl_3$) δ: 1.31-1.45 (12H, m), 1.49-1.65 (1H, m), 1.77-1.92 (2H, m), 1.98 (3H, s), 2.10-2.23 (2H, m), 2.33-2.46 (1H, m), 4.04 (t, J_{PH} = 17.6, 1H), 4.15-4.36 (8H, m), 4.41-4.50 (1H, m), 5.15-5.28 (1H, m), 7.65 (1H, s), 8.81 (1H, s); ^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 12.3, 16.3-16.6 (m), 30.2, 30.8, 38.3, 53.5, 63.1-63.7 (m), 70.8 (t, J_{PC} = 158.6), 83.4 (t, J_{PC} = 4.2), 111.5, 138.0, 151.3, 163.7; ^{31}P NMR (162.0 MHz, $CDCl_3$) δ: 16.25, 16.31 (ABq, J_{AB} = 31.2); HRMS (ES+) Exact mass calculated for $C_{19}H_{35}N_2O_9P_2$ $[M+H]^+$ 497.1818; found 497.1821.

***cis*-1-{4-[(Diphosphono)methoxy]cyclopentan-1-yl}thymine (4)**

cis-1-{4-[bis(Diethylphosphono)methoxy]cyclopentan-1-yl}thymine (106 mg, 0.21 mmol) was suspended in acetonitrile (2 mL) in a microwave vial and TMSBr (420 µL, 0.64 mmol, 15 eq) was added. The reaction mixture was irradiated at 50 °C for 1 h then methanol (2 mL) and water (0.5 mL) were added and stirring was continued for 1 h. Solvents were removed under reduced pressure and the residue was subjected to charcoal chromatography, eluting with water then 20% aqueous ammonia, to give the ammonium salt of the title compound as a cream solid (24 mg,

29%); ν_{max}/cm^{-1} 2948, 1655, 1272, 1041; ^{1}H NMR (400 MHz, D_2O) δ: 1.62-2.11 (m, 8H), 2.31-2.51 (1H, m), 3.72 (t, J_{PH} = 13, 1H), 4.31-4.51 (1H, m), 4.80-4.95 (1H, m), 7.84 (1H, s); ^{13}C NMR (100.6 MHz, D_2O) δ: 11.5, 29.1, 30.0, 37.3, 54.8, 81.1-81.3 (m), 111.3, 140.6, 152.6, 166.7; ^{31}P NMR (162.0 MHz, D_2O) δ: 13.76, 13.96; MS (ES+) *m/z*: 407 $(M + Na)^+$; HRMS (ES-) Exact mass calculated for $C_{11}H_{17}N_2O_9P_2$ $[M-H]^-$ 383.0409; found 383.0402.

***cis*-1-{4-[(Aminocarbonyl)dimethylphosphonomethoxy]cyclopentan-1-yl}thymine (35)**

cis-1-{4-[(Methoxycarbonyl)dimethylphosphonomethoxy]cyclopentan-1-yl}thymine **34** (155 mg, 0.40 mmol) was dissolved in 7N ammonia in methanol (20 mL, 140 mmol). The flask was tightly stoppered and the solution stirred at RT for 65 h. After concentration under reduced pressure, flash chromatography (1:9, MeOH:CH_2Cl_2) afforded **35** as a white solid as an essentially equimolar mixture of two diastereomers (52 mg, 35%); ν_{max}/cm^{-1} 3188, 2958, 1661, 1247, 1026, 727; ^{1}H NMR (400 MHz, $CDCl_3$) δ: 1.56-2.24 (m, 8H), 2.35-2.50 (m, 1H), 3.80-3.93 (m, 6H), 4.18-4.26 (m, 1H), 4.32 (2 overlapping d which appear as a t, J_{PH} = 18.8), 5.00-5.22 (m, 1H), 6.67-6.78 (br s, 1H), 6.94 (br s, 0.5H), 7.06 (br s, 0.5H), 7.55 (s, 0.5H), 7.57 (s, 0.5H), 9.95 (br s, 1H); ^{13}C NMR (100.6 MHz, $CDCl_3$) δ: 12.29, 12.33, 29.6, 30.4, 31.6, 37.4, 38.5, 53.5, 54.39, 53.8 (d, J_{PC} = 6.8), 54.0 (d, J_{PC} = 6.8), 54.2 (d, J_{PC} = 6.8), 54.3 (d, J_{PC} = 6.9), 74.1 (d, J_{PC} = 158.1), 75.3 (d, J_{PC} = 157.3), 81.8 (d, J_{PC} = 10.8), 82.7 (d, J_{PC} = 8.8), 111.4, 111.5, 137.91, 137.93, 151.49, 151.55, 164.2, 169.07 (d, J_{PC} = 3.0), 169.35 (d, J_{PC} = 3.3); ^{31}P NMR (162.0 MHz, $CDCl_3$) δ: 18.47, 18.65; HRMS (ES+) Exact mass calculated for $C_{14}H_{23}N_3O_7P$ $[M+H]^+$ 376.1274; found 376.1265.

***cis*-1-{4-[(Aminocarbonyl)phosphonomethoxy]cyclopentan-1-yl}thymine (5)**

cis-1-{4-[(Aminocarbonyl)dimethylphosphonomethoxy]cyclopentan-1-yl}thymine **35** (66 mg, 0.17 mmol) was suspended in acetonitrile (2 mL) in a microwave vial and TMSBr (93 μL, 0.70 mmol) was added. The reaction mixture was irradiated at 50 °C for 10 mins then methanol (1.9 mL) and water (100 μL) were added and stirring was continued for 10 minutes. Solvents were removed under reduced pressure and the residue was subjected to charcoal chromatography, eluting with water then 10:10:3, H_2O:MeOH:20% aq NH_3 to afford the ammonium salt of **5** as a fawn solid (45 mg, 70%); ν_{max}/cm^{-1} 3186, 2957, 1561, 1270, 1055, 765; ^1H NMR (400 MHz, D_2O) δ: 1.52-2.01 (m, 8H), 2.19-2.39 (m, 1H), 3.91 (2 overlapping d which appear as a t, J_{PH} = 17.2), 4.01 (br s, 1H), 4.66-4.85 (m, 1H), 7.61 (s, 0.5H), 7.64 (s, 0.5H); ^{13}C NMR (100.6 MHz, D_2O) δ: 11.46, 11.49, 28.7, 28.9, 29.7, 30.5, 36.5, 37.4, 54.7, 55.0, 77.4 (d, J_{PC} = 138.3), 78.0 (d, J_{PC} = 138.7), 80.6 (d, J_{PC} = 9.2), 81.1 (d, J_{PC} = 8.8), 111.2, 111.3, 140.1, 140.2, 152.49, 152.51, 166.49, 166.52, 175.8, 176.0; ^{31}P NMR (162.0 MHz, D_2O) δ: 9.7 br; HRMS (ES+) Exact mass calculated for $C_{12}H_{19}N_3O_7P$ $[M+H]^+$ 348.0961; found 348.0953.

ASSOCIATED CONTENT

Electronic Supplementary Information (ESI) available: ^1H and ^{13}C NMR spectra of all novel compounds. Analytic purity of compounds **2**, **2a**, **2b**, **3**, **4**, and **5** determined by HPLC.

AUTHOR INFORMATION

Corresponding Author

*A.R.M.: Fax, +(353)0214274097; Tel, +(353)0214901694; e-mail, a.maguire@ucc.ie

Author Contributions

All authors contributed equally to this work and have given their approval to the final version of the manuscript.

ACKNOWLEDGMENTS

This work was supported by Science Foundation Ireland (05/PICA/B802 and SFI 14/TIDA/2402). EA acknowledges support from NIH grant R37 MERIT Award AI27690. The research of JB was supported by the KU Leuven (GOA (GOA 15/19 TBA). We are grateful to Mrs. Lizette van Berckelaer and Mrs. Ria Van Berwaer for excellent technical assistance, Matthew T. Miller and Mena Issa for help with crystallization, Argon Photon Source (APS) beam line 23ID-D for X-ray data collection, and Prof. Pat Guiry, University College Dublin, Ireland for mass spectrometry.

REFERENCES

1. UNAIDS factsheet 2014

2. R. Payne, M. Muenchhoff, J. Mann, H. E. Roberts, P. Matthews, E. Adland, A. Hempenstall, K.-H. Huang, M. Brockman, Z. Brumme, M. Sinclair, T. Miura, J. Frater, M. Essex, R. Shapiro, B. D. Walker, T. Ndung'u, A. R. McLean, J.M. Carlson and P.J.R. Goulder, *Proceedings of the National Academy of Sciences* 2014, **111**, E5393-E5400. doi:10.1073/pnas.1413339111.

3. N. Pantazis, K. Porter, D. Costagliola, A. De Luca, J. Ghosn, M. Guiguet, A. M. Johnson, A. D. Kelleher, C. Morrison, R. Thiebaut, L. Wittkop and G. Touloumi, *The Lancet HIV* 2014, **1**, e119-e126. doi: 10.1016/S2352-3018(14)00002-2.

4. S. J. Keane, A. Ford, N. D. Mullins, N. M. Maguire, T. Legigan, J. Balzarini and A. R. Maguire, *J. Org. Chem.* 2015, **80**, 2479-2493. doi: 10.1021/jo502549y.

5. J. Balzarini, K. Das, J. A. Bernatchez, S. E. Martinez, M. Ngure, S. Keane, A. Ford, N. Maguire, N. Mullins, J. John, Y. Kim, W. Dehaen, J. Vande Voorde, S. Liekens, L. Naesens, M. Götte, A. R. Maguire and E. Arnold, *Proceedings of the National Academy of Sciences* 2015, **112**, 3475-3480. doi: 10.1073/pnas.1420233112.

6. M. Huang, G. H. Grant and W. G. Richards, *J. Mol. Graph. Model* 2011, **29**, 956-964. doi: 10.1016/j.jmgm.2011.04.002.

7. A. Savarino, *Expert Opinion on Investigational Drugs* 2006, **15**, 1507-1522. doi:10.1517/13543784.15.12.1507.

8. V. Summa, A. Petrocchi, F. Bonelli, B. Crescenzi, M. Donghi, M. Ferrara, F. Fiore, C. Gardelli, O. Gonzalez Paz, D. J. Hazuda, P. Jones, O. Kinzel, R. Laufer, E. Monteagudo, E. Muraglia, E. Nizi, F. Orvieto, P. Pace, G. Pescatore, R. Scarpelli, K. Stillmock, M. V. Witmer and M. Rowley, *J. Med. Chem.* 2008, **51**, 5843-5855. doi: 10.1021/jm800245z.

9. K.; Shimura, E.; Kodama, Y. Sakagami, Y. Matsuzaki, W. Watanabe, K. Yamataka, Y. Watanabe, Y. Ohata, S. Doi, M. Sato, M. Kano, S. Ikeda and M. Matsuoka, *J. Virol.* 2008, **82**, 764-774. doi: 10.1128/jvi.01534-07.

10. S. Hare, S. J. Smith, M. Métifiot, A. Jaxa-Chamiec, Y. Pommier, S. H. Hughes and P. Cherepanov, *Mol. Pharmacol.* 2011, **80**, 565-572. doi: 10.1124/mol.111.073189.

11. J. A. Grobler, K. Stillmock, B. Hu, M. Witmer, P. Felock, A. S. Espeseth, A. Wolfe, M. Egbertson, M. Bourgeois, J. Melamed, J. S. Wai, S. Young, J. Vacca and D. J. Hazuda,

Proceedings of the National Academy of Sciences 2002, **99**, 6661-6666. doi: 10.1073/pnas.092056199.

12. G. C. G. Pais, X. Zhang, C. Marchand, N. Neamati, K. Cowansage, E. S. Svarovskaia, V. K. Pathak, Y. Tang, M.; Nicklaus, Y. Pommier and T. R. Burke, *J. Med. Chem.* 2002, **45**, 3184-3194. doi: 10.1021/jm020037p.

13. C. A. Shaw-Reid, V. Munshi, P. Graham, A. Wolfe, M. Witmer, R. Danzeisen, D. B. Olsen, S. S. Carroll, M. Embrey, J. S. Wai, M. D. Miller, J. L. Cole and D. J. Hazuda, *J. Biol. Chem.* 2003, **278**, 2777-2780. doi: 10.1074/jbc.C200621200.

14. E. F. Eriksen, A. Díez-Pérez and S.Boonen, *Bone* 2014, **58**, 126-135. doi: 10.1016/j.bone.2013.09.023.

15. F. Vescini and F. Grimaldi, *Endocrine* 2015, **48**, 358-359. doi: 10.1007/s12020-014-0488-3.

16. Y. W. Yin, T. A. Steitz, *Cell* 2004, **116**, 393-404. doi: 10.1016/s0092-8674(04)00120-5.

17. P. R. Meyer, W. Rutvisuttinunt, S. E. Matsuura, A. G. So and W. A. Scott, *J. Mol. Biol.* 2007, **369**, 41-54. doi: 10.1016/j.jmb.2007.03.006.

18. P. R. Meyer, S. E.; Matsuura, D. Zonarich, R. R. Chopra, E. Pendarvis, H. Z. Bazmi, J. W. Mellors and W. A. Scott, *J. Virol.* 2003, **77**, 6127-6137. doi: 10.1128/jvi.77.11.6127-6137.2003.

19. B. Marchand, E. P. Tchesnokov and M. Götte, *J. Biol. Chem.* 2007, **282**, 3337-3346. doi: 10.1074/jbc.M607710200.

20. B. Eriksson, B. Öberg and B.Wahren, *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* 1982, **696**, 115-123. doi: 10.1016/0167-4781(82)90018-5.

21. C. E. McKenna, L. A. Khawli, A. Bapat, V. Harutunian and Y. C. Cheng, *Biochem. Pharmacol.* 1987, **36**, 3103-3106. doi: 10.1016/0006-2952(87)90618-6.

22. C. E. McKenna and J. N. Levy, *J. Chem. Soc. Chem. Comm.* 1989, 246-247. doi: 10.1039/C39890000246.

23. Z.-Y. Peng, J. M. Mansour, F. Araujo, J.-Y. Ju, C. E. McKenna and T. E. Mansour, *Biochem. Pharmacol.* 1995, **49**, 105-113. doi: 10.1016/0006-2952(94)00437-Q.

24. O. T. Quimby, J. B. Prentice and D. A. Nicholson, *J. Org. Chem.* 1967, **32**, 4111-4114. doi: 10.1021/jo01287a104.

25. R. V. Talanian, N. C. Brown, C. E. McKenna, T. G. Ye, J. N. Levy and G. E. Wright, *Biochemistry* 1989, **28**, 8270-8274. doi: 10.1021/bi00447a002.

26. C. E. McKenna, J. N. Levy, L. A. Khawli, V. Harutunian, T. G. Ye, M. C. Starnes, A. Bapat and Y. C. Cheng, In Nucleotide Analogues as Antiviral Agents; American Chemical Society: 1989; Vol. 401, p 1.

27. C. E. McKenna, B. A. Kashemirov and C. N. Rozé, *Bioorg. Chem.* 2002, **30**, 383-395. doi: 10.1016/S0045-2068(02)00521-7.

28. C. M. Lacbay, J. Mancuso, Y.-S. Lin, N. Bennett, M. Götte and Y. S. Tsantrizos, *J. Med. Chem.* 2014, **57**, 7435-7449. doi: 10.1021/jm501010f.

29. E. A. Shirokova, M. V. Jasko, A. L. Khandazhinskaya, A. V. Ivanov, D. V. Yanvarev, Y. S. Skoblov, V. A. Mitkevich, E. V. Bocharov, T. R. Pronyaeva, N. V. Fedyuk, M. K. Kukhanova and A. G. Pokrovsky, *J. Med. Chem.* 2004, **47**, 3606-3614. doi: 10.1021/jm0310176.

30. S. Debarge, A.R. Maguire and J. Balzarini, *J. Org. Chem.* 2011, **76**,105-126. doi: 10.1021/jo101738e.

31. I. Hladezuk, V. Chastagner, S. G. Collins, S. J. Plunkett, A. Ford, S. Debarge and A. R. Maguire, *Tetrahedron*, 2012, **68** 1894-1909. doi: 10.1016/j.tet.2011.12.077.

32. J. K. Crandall, D. B. Banks, R. A. Colyer, R. J. Watkins and J. P. Arrington, *J. Org. Chem.* 1968, **33**, 423-425. doi: 10.1021/jo01265a089.

33. D. R. Deardorff and D. C. Myles, *Org. Synth.* 1989, **67**, 114. doi: 10.15227/orgsyn.067.0114.

34. M. Korach, D. R. Nielson and W. H. Rideout, *Org. Synth.* 1962, **42**, 50. doi: 10.15227/orgsyn.042.0050.

35. B. M. Trost, G. H. Kuo and T. Benneche, *J. Am. Chem. Soc.* 1988, **110**, 621. doi: 10.1021/ja00210a064.

36. A. B. Khare and C. E. McKenna, *Synthesis* 1991, 405-405. doi: 10.1055/s-1991-26478.

37. M. T. Miller, S. Tuske, K. Das, J. Jeffrey DeStefano and E. Arnold, *Protein Sc.* 2015. In press. doi: 10.1002/pro.2776.

38. A.G.W. Leslie and H.R. Powell, (2007), Evolving Methods for Macromolecular Crystallography, 245, 41-51. Processing Diffraction Data with Mosflm. ISBN 978-1-4020-6314-5.

39. Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart. P. H. *Acta Cryst.* 2010, **D66**, 213-221.

40. P. Emsley, B. Lohkamp, W. G. Scott and L. Cowtan, *Acta Crystallogr.* 2010, **D66**, 486-501.