Polyfluoroaromatic Stavudine (d4T) ProTides Exhibit Enhanced Anti-HIV Activity

Sahar Kandil⁷, Christophe Pannecouque⁷, Fiona M. Chapman⁷, Andrew D. Westwell⁷ and Christopher McGuigan⁷

⁷School of Pharmacy and Pharmaceutical Sciences, Cardiff University, King Edward VII Avenue, Cardiff CF10 3NB, ⁷Rega Institute for Medical Research - Laboratory of Virology and Chemotherapy, K.U. Leuven, Minderbroedersstraat 10, B-3000 Leuven, Belgium

Keywords:
Polyfluoroaromatic, pentafluorosulfanyl, SF₅, Stavudine, d4T, ProTide, Phosphoramidate, Antiviral, Nucleoside, HIV.

Graphical abstract

Highlights

- Synthesis of polyfluoroaromatic Stavudine (d4T) ProTides.
- Anti-retroviral activity against two strains; HIV-1, (IIIb) and HIV-2, (ROD) in MT4 cell line.
- Evaluation of anti-HIV activity against wild type (C8166) as well as thymidine kinase deficient (C8166-TK⁻) cells.
- Nanomolar antiviral activity of novel ProTide 5d containing 3-SF₅ aromatic ring; IC₅₀ = 30 nM, HIV-1 and IC₅₀ = 36 nM, HIV-2, more than tenfold that of d4T.
- High selectivity index of 5d, SI = 1753 (HIV-1) and 1461 (HIV-2) twice that of d4T.
- Enzymatic activation of 5d is monitored by ³¹P and ¹⁹F NMR.
- Polyfluoroaromatic ProTides have high in vitro aqueous solubility.
Human Immunodeficiency Virus (HIV) damages the immune system and leads to the life-threatening acquired immunodeficiency syndrome (AIDS). Despite the advances in the field of antiretroviral treatment, HIV remains a major public health challenge.

Nucleosides represent a prominent chemotherapeutic class for treating viral infections, however their cellular uptake, kinase-mediated activation and catabolism are limiting factors. Herein, we report the synthesis and in vitro evaluation of stavudine (d4T) ProTides containing polyfluorinated aryl groups against two strains; HIV-1 (strain IIIb) and HIV-2 (strain ROD). ProTide 5d containing a meta-substituted pentafluorosulfonyl (3-SF5) aryl group showed superior antiviral activity over the parent d4T and the non-fluorinated analogue 5a. ProTide 5d has low nanomolar antiviral activity; (IC50 = 30 nM, HIV-1) and (IC50 = 36 nM, HIV-2) which is over tenfold more potent than d4T. Interestingly, ProTide 5d showed significantly high selectivity indices (SI) = 1753 (HIV-1) and 1461 (HIV-2) which is more than twice that of the d4T. All ProTides were screened in wild type as well as thymidine kinase deficient (TK-) cells. Enzymatic activation of ProTide 5d using carboxypeptidase Y enzyme and monitored using both 31P and 19F NMR is presented.

Human Immunodeficiency Virus (HIV) is a retrovirus, of the family Retroviridae, which is characterised by the presence of the reverse transcriptase (RT) enzyme. HIV damages the immune system and causes higher susceptibility to life-threatening opportunistic infections, known as acquired immunodeficiency syndrome (AIDS). Many advances have been made in antiretroviral therapy that have turned AIDS into a chronic condition rather than a fatal illness. However, HIV remains a major public health challenge because of the high mutation rate of HIV-RT, the adverse side effects of the antiretroviral treatment and the poor patient compliance in clinical applications. Nucleoside / nucleotide reverse-transcriptase inhibitors (NRTIs/NtRTIs) represent the first major class of antiretroviral drugs developed.

Nucleoside analogues are a well-established class of antiviral and anticancer therapeutic agents. Enzymatic phosphorylation is required to convert the nucleoside analogues to the 5'-monophosphates and subsequently to the triphosphate derivatives. Often the three-step intracellular phosphorylation is inefficient and rate-limited by the initial monophosphorylation step. Further, reduced expression of nucleoside kinases leads to resistance to the nucleoside analogue treatment. Therefore, notable efforts have focused on the delivery of monophosphate prodrugs that are capable of masking the negative charge.
of the phosphate group. The monophosphate phosphoramidate pro-nucleotides, known as ProTides, represent a successful strategy for bypassing the dependence on active transport and nucleoside kinase-mediated activation. This is demonstrated clinically by sofosbuvir (anti-HCV, FDA approved) and acelarin (anticancer, Phase III), shown in Figure 1.

Figure 1: Chemical structures of sofosbuvir (anti-HCV, FDA approved), acelarin (anticancer, Phase III) ProTides and stavudine (d4T) nucleoside.

Stavudine (2’,3’-didehydro-2’,3’-dideoxythymidine, d4T) is an approved antiviral nucleoside for the prevention or treatment of HIV infection, usually as part of antiretroviral combination therapy. In an attempt to overcome the first intracellular thymidine kinase-directed phosphorylation step of d4T, phosphoramidate derivatives were prepared and evaluated against HIV.

Generally, introduction of fluorinated substituents can provide a unique combination of electronegativity, size and lipophilicity and can also greatly affect physicochemical properties. In this work we synthesised phosphoramidate derivatives of d4T (Scheme 1) and compared the effect of introducing different polyfluorinated variants of the aromatic ring (4-CF₃ / 3-CF₃ / 3-SF₅) and investigated whether the lipophilic and electronegative properties of the corresponding ProTides would affect their anti-HIV activity and toxicity profiles compared to the non-fluorinated ProTide.

Stavudine (d4T) ProTides were prepared using previously described phosphoro dichloridate chemistry. The phenyl phosphorodichloridates 2a-d were prepared by the reaction of substituted phenols 1a-d with POCl₃. The subsequent reaction with methyl alanine 3 to form the corresponding phosphoro dichloridates 4a-d was monitored by ³¹P NMR. Next the arylaminoacyl phosphoro dichloridates 4a-d were reacted with d4T in the presence of t-BuMgCl as a hindered base (Scheme 1). Stirring for 24h at ambient temperature generated the crude ProTides 5a-d that were purified by column chromatography to provide the characteristic low to moderate yields associated with ProTide chemistry, for the final products (20–34%). Each
of the phosphoramidates was generated as a pair of diastereoisomers at the phosphorus centre, in a roughly 1:1 ratio, as revealed by the two closely spaced peaks in the $^{31}$P NMR spectrum (see Supporting Information). It is worth mentioning that due to the instability of the intermediates it was not feasible to prepare the 4-SF$_5$ analogue.

Scheme 1. Synthesis of d4T phosphoramidate analogues; Reagents and conditions: i) Et$_3$N, anhydrous Et$_2$O, -78 °C to rt, 2 h, ii) Et$_3$N, anhydrous DCM, -78 °C, 3 h, iii) t-BuMgCl, anhydrous THF, -78 °C to rt, 24 h.

ProTides need to be metabolised to release the monophosphate form, which will then undergo further intracellular phosphorylation steps to generate di- and triphosphate active form. The proposed intracellular activation pathway of the most active ProTide 5d is described in Scheme 2, in accordance with previous literature$^{15}$. Metabolic breakdown starts with the hydrolysis of the ProTide carboxylate ester, mediated by a carboxyesterase-type enzyme to form intermediate A, which is followed by a spontaneous cyclisation displacing the aryl moiety via an internal nucleophilic attack of the carboxylate residue on the phosphorus centre to yield intermediate B. In the third step, the unstable cyclic mixed anhydride is thought to be hydrolysed to release the intermediate C. The final step then involves a phosphoramidase-type enzyme, which would cleave off the amino acid to generate the corresponding monophosphate D, trapped within the cell due to the polar nature of the monophosphate.
Scheme 2. Proposed activation pathway of d4T ProTide 5d initiated by esterase or carboxypeptidase enzyme.

To probe the activation process of ProTide 5d, Scheme 2, an enzymatic study using carboxypeptidase Y enzyme was carried out. ProTide 5d was incubated with carboxypeptidase Y to verify that the enzymatic cleavage of the ester motif is sufficient to initiate the activation process and generate intermediate C. Compound 5d was dissolved in acetone-d₆ in the presence of Trizma buffer (pH 7.6), treated with carboxypeptidase Y and monitored by ³¹P NMR chemical shift changes over time, Figure 2. Two peaks at δ 3.61 and 3.27 ppm in the blank spectrum recorded at 25 °C correspond to the diastereoisomers of the parent compound 5d. Within 7 mins of incubation with carboxypeptidase Y, a single peak at around δ 6 ppm started appearing, which is consistent with the single chemical shift expected for the formation of achiral intermediate C via the putative intermediate B. The peak height increased over time suggesting the accumulation of a stable product. The spectra recorded during the enzymatic reaction suggest that within 2 h of the start of the assay 5d was almost fully converted into metabolite C.

Figure 2. ³¹P NMR spectra of d4T ProTide 5d over time after carboxypeptidase Y treatment showing the signals of the metabolite C over time.
Moreover, due to the presence of the 3-SF$_5$ group in 5d, we considered that the activation process of ProTide 5d using carboxypeptidase Y enzyme can be also monitored using $^{19}$F NMR chemical shift changes over time, Figure 3. The appearance of two doublet peaks at $\delta$ 62.15 and 62.48 ppm in the blank spectrum corresponds to the diastereoisomers of the parent compound 5d. The new doublet signal appearing at $\delta$ 61.86 and 62.21 (corresponding to the four fluorine atoms of the SF$_5$ group) indicating the release of 3-SF$_5$ phenol (1d).

![Figure 3. $^{19}$F NMR spectra of ProTide 5d over time after carboxypeptidase Y treatment showing the signals of the metabolites over time.](image)

Both $^{31}$P NMR and $^{19}$F NMR monitoring experiments, Figure 2 and 3, indicate that ProTide 5d activation is initiated by the carboxypeptidase enzyme Y and that after 1.25h of incubation most of the ProTide is metabolised. Quantification of intracellular drug levels would be the focus of future investigation.

Nucleoside reverse transcriptase inhibitors (NRTIs) and nucleotide analogue reverse-transcriptase inhibitors (NtRTIs) act as chain terminators for both viral and host DNA synthesis. The former describes their antiviral activity, while the latter describes their drug toxicity/side effects. Hence, d4T ProTides 5a-d were screened for both antiviral activity and cytotoxicity.

The antiviral activity profile of compounds 5a-d was evaluated in MT4 cells against two HIV strains; HIV-1 (IIIb) and HIV-2 (ROD) in parallel with stavudine (d4T) as a positive control. The biological results are expressed as IC$_{50}$, CC$_{50}$ and SI (selectivity index; CC$_{50}$/IC$_{50}$ ratio).

All ProTides showed markedly enhanced antiviral activity compared to the parent d4T except the non-fluorinated ProTide 5a which is slightly more potent than d4T. The best activity was
observed across both HIV types in ProTide 5d with the meta-pentafluorosulfanyl (3-SF₅) substituent attached to its aromatic ring. ProTide 5d has low nanomolar antiviral activity; (IC₅₀ = 30 nM, HIV-1) and (IC₅₀ = 36 nM, HIV-2) which is over tenfold more potent than d4T, Table 1. Additionally, ProTide 5d showed significantly high selectivity indices (SI) = 1753 (HIV-1) and 1461 (HIV-2) which is more than twice that of the d4T. This could be attributed to its enhanced lipophilic and/or electronegative properties imparted by the presence of five fluorine atoms. Also, the position of the CF₃ substituent in the phenyl ring of 5b and 5c has little influence on the anti-HIV activity but more effect on the toxicity profile which was in favour of the meta-CF₃ (5c) over the para-CF₃ (5b) as observed in the selectivity index values; SI (5c) = 1612 (HIV-1) and 1349 (HIV-2) compared to SI (5b) = 1159 (HIV-1) and 921 (HIV-2), Table 1.

<table>
<thead>
<tr>
<th>comp</th>
<th>HIV-1 (IIIb)</th>
<th>HIV-2 (ROD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IC₅₀ (µM)</td>
<td>CC₅₀ (µM)</td>
</tr>
<tr>
<td>d4T</td>
<td>0.40±0.09</td>
<td>313±7</td>
</tr>
<tr>
<td>5a</td>
<td>0.21±0.09</td>
<td>251±7</td>
</tr>
<tr>
<td>5b</td>
<td>0.037±0.002</td>
<td>42.9±12.6</td>
</tr>
<tr>
<td>5c</td>
<td>0.041±0.006</td>
<td>66.1±24.8</td>
</tr>
<tr>
<td>5d</td>
<td>0.030±0.012</td>
<td>52.6±4.9</td>
</tr>
</tbody>
</table>

Table 1. Anti-HIV-1 and 2 activity and cytotoxicity of stavudine (d4T) and ProTide analogues 5a-d.

Further, we tested our compounds in the C8166 cell line and a variant thymidine kinase deficient variant (C8166-TK) to probe the effect of TK deficiency on the activity of the test compounds and the degree to which the ProTides could bypass this dependence.

As shown in Table 2, the fluorinated ProTides 5b-d displayed greatly enhanced antiviral potency than the non-fluorinated analogue 5a or the parent d4T across. ProTide 5b (4-CF₃) has around 180-fold higher activity than d4T, while ProTide 5d (3-SF₅) is around 75-fold more active than d4T. It is interesting to observe that compounds 5a, 5b and 5d can retain their activity in TK deficient cells while d4T has completely lost its activity. While ProTide 5c (IC₅₀ = 0.039 µM) shows better activity than 5a (IC₅₀ = 0.30 µM) and d4T (IC₅₀ = 4.06 µM), it is not active in TK- cells. This is most probably because 5c uptake is enhanced but it falls back to its free nucleoside form and cannot be re-phosphorylated in these cells, Table 2.
Table 2. Antiviral activity of stavudine (d4T) and ProTide analogues 5a-d against HIV-1 strain IIIb in the C8166 cell line and the thymidine kinase deficient (C8166-TK-) cell line. NA: not active

<table>
<thead>
<tr>
<th>ID</th>
<th>IC50 WT (µM)</th>
<th>IC50 TK (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>d4T</td>
<td>4.06±0.18</td>
<td>NA</td>
</tr>
<tr>
<td>5a</td>
<td>0.30±0.04</td>
<td>0.32±0.15</td>
</tr>
<tr>
<td>5b</td>
<td>0.022±0.006</td>
<td>0.075±0.000</td>
</tr>
<tr>
<td>5c</td>
<td>0.039±0.007</td>
<td>NA</td>
</tr>
<tr>
<td>5d</td>
<td>0.054±0.039</td>
<td>0.335±0.004</td>
</tr>
</tbody>
</table>

Introduction of fluorine atoms into molecular structures increases their hydrophobicity. Thus, we investigated the in vitro turbidimetric aqueous solubility of the polyfluorinated ProTides 5b-d using five concentrations (1, 3, 10, 30 and 100 µM) in phosphate buffered saline (PBS; pH 7.4) solution (1% DMSO), which were incubated at 37°C for 2 hr. The estimated precipitation range (lower and upper bound) and a mid-range value were calculated for compounds 5b-d. Nicardipine and pyrene were used as control compounds. The results show that ProTides with polyfluoroaromatic substituents 5b-d retain high aqueous solubility with estimated precipitation range 100 µM (lower bound) and >100 µM (upper bound), Table 3, confirming their potential advantage for intestinal absorption.

Table 3. Aqueous solubility of 5b-d versus control compounds (nicardipine and pyrene) in PBS (pH=7.4) buffered aqueous solution. *Calculated by PerkinElmer ChemDraw 16.0 software.

<table>
<thead>
<tr>
<th>Compound</th>
<th>MW</th>
<th>ClogP*</th>
<th>Estimated Precipitation range (µM)</th>
<th>Estimated solubility (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lower bound</td>
<td>Upper bound</td>
</tr>
<tr>
<td>d4T</td>
<td>224.22</td>
<td>-0.4875</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>5b</td>
<td>533.40</td>
<td>1.9788</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>5c</td>
<td>533.40</td>
<td>1.9788</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>5d</td>
<td>591.44</td>
<td>2.3288</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Nicardipine</td>
<td>-</td>
<td>-</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>Pyrene</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>10</td>
</tr>
</tbody>
</table>

In this study, a series of polyfluoroaromatic ProTides of d4T were synthesised. The antiviral activity was evaluated in vitro against the two HIV types; HIV-1 (strain IIIb) and HIV-2 (strain ROD). Our findings identified compound 5d featuring a meta-pentafluorosulfanyl (SF₅) in its aromatic ring as promising lead. ProTide 5d has low nanomolar antiviral activity; (IC50 = 30 nM, HIV-1) and (IC50 = 36 nM, HIV-2) which is over tenfold more potent than d4T. Moreover,
**5d** has the best activity / toxicity profile with a selectivity index (SI) = 1753 (HIV-1) and 1461 (HIV-2), which is more than twice that of the parent d4T (SI) = 782 (HIV-1) and 695 (HIV-2). Additionally, ProTides **5a, 5c and 5d** retained some activity in the thymidine kinase deficient (C8166-TK⁻) cell line. Furthermore, the polyfluorinated ProTides **5b-d** retained aqueous solubility comparable to that of the parent d4T. Considering the above findings, it is concluded that **5d** represents a promising basis for further development of NtRTI anti-HIV therapies.

**Acknowledgment**

The authors would like to thank the Welsh Government (A4B-Academic Expertise for Business). The technical assistance of Mrs. Kristien Erven, Mr. Kris Uyttersprot and Mrs. Cindy Heens for the HIV experiments is gratefully acknowledged. The authors also would like to dedicate this article to the memory of Prof. Christopher McGuigan.

**Supplementary data**

Supplementary data including chemistry and carboxypeptidase Y enzymatic assay procedures. $^1$H, $^{31}$P, $^{19}$F and $^{13}$C NMR spectra of Protides **5a-d** as well as HPLC analysis reports associated with this article can be found, in the online version.

**References and notes**


16. General experimental procedure for the preparation of ProTides (5a-d)

D4T phosphoramidates (5a-d) were prepared applying the previously described phosphochloridate chemistry. To a stirring solution of d4T (1.0 eq.) in THF, tert-butylmagnesium chloride (1.0 M in THF) (1.1 eq) was added dropwise over 1 min, and the appropriate phosphochloridate (3.0 eq.) in anhydrous THF was added at room temperature and stirred for 16 h. The solvent was removed under reduced pressure and the yellow oil obtained was dissolved in DCM and washed with 0.5 M HCl and water. The organic layer was dried over MgSO₄, filtered, reduced to dryness and purified by flash chromatography.

17. Spectral data of ProTide 2ʹ,3ʹ-Didehydro-2ʹ,3ʹ-dideoxythymidine-5ʹ-(3-(penta fluorosulfanyl) phenyl methoxy alaninyl phosphate) (5d);

**¹H NMR** (CDCl₃): δ 9.11 (bs, 1H, NH), 7.54-7.49 (m, 2H, ArH), 7.41-7.31 (m, 2H, ArH), 6.97-6.93 (2m, 1H, H₁), 6.28-6.20 (2m, 1H, H-2‘), 5.88 -5.82 (2m, 1H, H-3‘), 5.00-4.92 (2m, 1H, H-4‘), 4.35-4.17 (2m, 2H, H-5’), 3.98-3.88 (m, 1H, Ala CH), 3.64, 3.64 (2s, 3H, OCH₃), 1.82 (bs, 1H, NH), 1.78, 1.74 (2d, J = 1 Hz, 3H, CH₃- 5), 1.32, 1.27 (2d, J = 6.5 Hz, 3H, Ala CH₃); ³¹P NMR (CDCl₃): δ 3.10, 2.60; ¹⁹F NMR (CDCl₃): δ 83.21 (p, J = 149.9 Hz, 1F), 62.82, 62.80 (2d, J = 149.9, 150.4 Hz, 4F); ¹³C NMR (CDCl₃): 173.88, 173.66 (2d, J= 6.5, 7.4 Hz, C=O ester), 163.68, 163.63 (ArC, C-4), 154.41 (m, ArC, CSF₅), 150.76 (C=O, C-2), 150.18, 150.04 (2d, J= 5.8 Hz, ArC ipso phenyl), 135.64, 135.44 (ArCH, C-6), 133.05, 132.77 (C-3’), 129.91 (ArCH)
,127.73, 127.63 (C-2'), 123.80, 123.60 (2d, \(^2J_{C-P} = 3.8\), 5 Hz, ArCH), 122.66 (m, ArCH), 118.55 (m, ArCH), 111.45, 111.27 (ArC, C-5), 89.96, 89.68 (C-1'), 84.48, 84.42 (2d, \(^3J_{C-P} = 2.5\) Hz, C-4'), 67.58, 66.86 (2d, \(^2J_{C-P} = 4\), 3.8 Hz, C-5'), 52.69 (OCH\(_3\)), 50.25, 50.16 (Ala CH), 20.93, 20.82 (2d, 5, 6.3 Hz, Ala CH\(_3\)), 12.35, 12.25 (CH\(_3\), C-5). MS (ES+) m/z: 614.1 (M+Na+), C\(_{29}H_{23}F_5N_3O_8\)PS, Reverse-phase HPLC, eluting with H\(_2\)O/CH\(_3\)CN from 90/10 to 0/100 in 30 min; flow = 1 mL/min, \(t_R = 16.01\) min. yield 34%.
Supporting Information

Polyfluoroaromatic Stavudine (d4T) ProTides Exhibit Enhanced Anti-HIV Activity

Sahar Kandil, Christophe Pannecouque, Fiona M. Chapman, Andrew D. Westwell and Christopher McGuigan

School of Pharmacy and Pharmaceutical Sciences, Cardiff University, King Edward VII Avenue, Cardiff CF10 3NB, Rega Institute for Medical Research - Laboratory of Virology and Chemotherapy, K.U. Leuven, Minderbroedersstraat 10, B-3000 Leuven, Belgium

Contents

Chemistry and carboxypeptidase Y enzymatic assay procedures. $^1$H, $^{31}$P, $^{19}$F and $^{13}$C NMR spectra of Protides 5a-d as well as HPLC analysis reports.

S1. Chemistry.

S2. Carboxypeptidase Y enzymatic assay procedure.

S1. Chemistry

The anhydrous solvents and any other commercially available reagents were purchased from Sigma-Aldrich (U.K.) and used without further purification. Amino acid esters were purchased from Carbosynth (U.K.). Carboxypeptidase Y and buffers were purchased from Sigma-Aldrich (U.K.). All reactions were carried out under nitrogen atmosphere. Reactions were monitored with analytical TLC on silica gel 60-F254 precoated aluminium plates and visualised under UV (254 nm) and/or using $^{31}$P NMR spectra. Column chromatography was performed on silica gel (35-70 μM). Preparative TLC plates (20 cm × 20 cm, 500-2000 μm) were purchased from Merck. Proton ($^1$H), carbon ($^{13}$C), Fluorine ($^{19}$F) and phosphorus ($^{31}$P) NMR spectra were recorded on a Bruker Avance 500 MHz spectrometer at 25 °C. Spectra were autocalibrated to the deuterated solvent peak, and all $^{13}$C NMR and $^{31}$P NMR were proton decoupled. The purity of the final compounds was verified to be >95% by HPLC analysis using either i) ThermoSCIENTIFIC, SPECTRA SYSTEM P4000, detector SPECTRA SYSTEM UV2000, Varian Pursuit XRs 5 C18, 150 x 4.6 mm (as an analytic column) or ii) Varian Prostar (LC Workstation-Varian Prostar 335 LC detector), Thermo SCIENTIFIC Hypersil Gold C18, 5 μm, 150 x 4.6 mm (as an analytical column) with a gradient elution of H$_2$O/ CH$_3$CN from 100/0 to 0/100 in 35
min, flow = 1 mL/min, λ = 275 nm. Mass spectra were performed on Bruker Daltonics microTOF-LC, (atmospheric pressure ionisation, electron spray mass spectroscopy) in either positive or negative mode.

2',3'-Dideoxy-2',3'-didehydrothymidine 5'-(phenyl methoxyalaninyl phosphate) (5a)

{H NMR (CDCl₃): δ 8.80, 8.76 (2s, 1H, NH), 7.36-7.13 (m, 6H, ArH, H-6), 7.05-6.99 (2m, 1H, H-1'), 6.37-6.26 (2dt, J = 6, 2 Hz, 1H, H-2'), 5.93 -5.85 (2m, 1H, H-3'), 5.06-4.98 (2m, 1H, H-4'), 4.42-4.26 (2m, 2H, H-5'), 4.04-3.93 (m, 1H, Ala CH), 3.71, 3.70 (2s, 3H, OCH₃), 2.00 (bs, 1H, NH), 1.86, 1.82 (2d, J = 0.3 Hz, 3H, CH₃- 5), 1.36, 1.32 (2d, J = 7 Hz, 3H, Ala CH₃); ³¹P NMR (CDCl₃): δ 3.10, 2.49; ¹³C NMR (CDCl₃): 173.96, 173.82 (2d, ³JC-P = 7.1 Hz, C=O ester), 163.69, 163.63 (ArC-4), 150.79, 150.75 (C=O, C-2), 150.42, 150.28 (2d, ³JC-P = 6.3 Hz, ArC ipso phenyl), 135.88, 135.61 (ArCH, C-6), 133.37, 132.07 (C-3'), 129.80, 129.74 (ArCH),127.49, 127.33 (C-2'), 125.25, 125.19 (ArCH), 120.19, 120.05 (2d, ³JC-P = 3.8, 5 Hz, ArCH), 111.41, 111.29 (C-5), 89.85, 89.62 (C-1'), 84.71, 84.64 (C-4'), 67.16, 66.55 (2d, ³JC-P = 5, 3.8 Hz, C-5'), 52.60, 52.58 (OCH₃), 50.21, 50.09 (AlaCH), 20.96, 20.90 (2d, 5 Hz, Ala CH₃), 12.34, 12.30 (CH₃, C-5). MS (ES⁺) m/z: 488.1 (M+Na⁺), C₂₀H₂₄N₃O₅P, Reverse-phase HPLC, eluting with H₂O/CH₃CN from 90/10 to 0/100 in 30 min; flow = 1 mL/min, tᵣ = 15.00 min. yield 28%.

2',3'-Didehydro-2',3'-dideoxythymidine 5'-(3-trifluoromethylphenylmethoxyalaninyl phosphate) (5b)

{H NMR (CDCl₃): δ 9.24, 9.21 (2s, 1H, NH), 7.49-7.38 (m, 4H, ArH), 7.29-7.20 (2m, 1H, H-6), 7.05-7.00 (2m, 1H, H-1'), 6.36-6.27 (2dt, J = 6, 2 Hz, 1H, H-2'), 5.94 -5.87 (2m, 1H, H-3'), 5.07-4.99 (2m, 1H, H-4'), 4.43-4.24 (2m, 2H, H-5'), 4.11-3.94 (m, 2H, Ala CH, NH), 3.71, 3.70 (2s, 3H, OCH₃), 1.86, 1.81 (2d, J = 1.5 Hz, 3H, CH₃- 5), 1.38, 1.35 (2d, J = 7 Hz, 3H, Ala CH₃); ³¹P NMR (CDCl₃): δ 3.35, 2.63; ¹⁹F NMR (CDCl₃): δ -62.74, -62.75; ¹³C NMR (CDCl₃): 173.90, 173.74 (2d, ³JC-P = 7.5 Hz, C=O ester), 163.83, 163.79 (ArC-4), 150.86 (C=O, C-2), 150.63, 150.50 (2d, ³JC-P = 6.3, 5 Hz, ArC ipso phenyl), 135.68, 135.48 (ArCH, C-6), 133.09, 132.83 (C-3'), 130.47 (ArCH), 130.06 (m, ArC), 127.68, 127.58 (C-2'), 123.85, 123.68 (2d, ³JC-P = 3.8, 5 Hz, ArCH), 123.39 (q, ³JC-P = 270.8 Hz, CF₃), 121.96 (m, ArCH), 117.39 (m, ArCH), 111.44, 111.27 (C-5), 89.92, 89.65 (C-1'), 84.54, 84.51 (2d, ³JC-P = 3.8, 5 Hz, C-4'), 67.47, 66.80 (2d, ³JC-P = 3.8, 5 Hz, C-5'), 52.64 (OCH₃), 50.24, 50.13 (AlaCH), 20.89, 20.80 (2d, 5 Hz, 6.3 Hz, Ala CH₃), 12.34, 12.26 (CH₃, C-5). MS (ES⁺) m/z: 556.1 (M+Na⁺), C₂₁H₂₃F₃N₃O₈P, Reverse-phase HPLC, eluting with H₂O/CH₃CN from 90/10 to 0/100 in 30 min; flow = 1 mL/min, tᵣ = 14.97 min. yield 20%. 
2',3'-Didehydro-2',3'-dideoxythymidine-5'(4-(trifluoromethyl)phenyl methoxy alaninyl phosphate) (Sc).

\[
\begin{align*}
\text{H NMR (CDCl}_3\text{): } & \delta 8.49, 8.45 (2bs, 1H, NH), 7.54 (d, J = 8 Hz, 2H, ArH), 7.26, 7.23 (d, J = 8.5, 9 Hz, 2H, ArH), 7.21-7.13 (2m, 1H, H-6), 6.98-6.93 (2m, 1H, H-1'), 6.31-6.21 (2dt, J = 6, 2 Hz, 1H, H-2'), 5.88 -5.82 (2m, 1H, H-3'), 5.01-4.93 (2m, 1H, H-4'), 4.35-4.16 (2m, 2H, H-5'), 3.98-3.86 (m, 1H, Ala CH), 3.76-3.66 (m, 1H, NH), 3.64, 3.63 (2s, 3H, OCH_3), 1.81, 1.76 (2d, J = 1 Hz, 3H, CH_3- 5), 1.31, 1.27 (2d, J = 7 Hz, 3H, Ala CH_3); 31P NMR (CDCl_3): \delta 2.95, 2.41; 19F NMR (CDCl_3): \delta -62.21, -62.22; 13C NMR (CDCl_3): 173.74, 173.60 (C=O ester), 163.44, 163.41 (Ar-C-4), 153.04, 152.84 (ArC ipso phenyl), 150.62 (C=O, C-2), 135.68, 135.46 (ArCH, C-6), 133.18, 132.87 (C-3'), 127.70, 127.55 (C-2'), 127.21 (m, ArCH), 124.83 (ArC), 122.93 (ArC), 120.49 (2d, J = 5 Hz, ArCH), 111.42, 111.27 (C-5), 89.96, 89.71 (C-1'), 84.54, 84.47 (C-4'), 67.49, 66.80 (2d, \text{J}_{CP} = 3.8, 5 Hz, C-5'), 52.69 (OCH_3), 50.21,50.14 (AlaCH), 20.98, 20.90 (2d, J = 5 Hz, Ala CH_3), 12.39, 12.36 (CH_3, C-5). \text{MS (ES') m/z: } 556.1 (M+Na^+), C_{21}H_{23}F_3N_3O_8P, \text{Reverse-phase HPLC, eluting with H}_2O/CH_3CN from 90/10 to 0/100 in 30 min; flow = 1 mL/min, t_r = 17.08 min. yield 24%. \\
2',3'-Didehydro-2',3'-dideoxythymidine-5'-{(3-pentafluorosulfanyl)phenyl methoxy alaninyl phosphate} (5d)
\]

\[
\begin{align*}
\text{H NMR (CDCl}_3\text{): } & \delta 9.11 (bs, 1H, NH), 7.54-7.49 (m, 2H, ArH), 7.41-7.31 (m, 2H, ArH), 7.21-7.11 (2m, 1H, H-6), 6.97-6.93 (2m, 1H, H-1'), 6.28-6.20 (2m, 1H, H-2'), 5.88 -5.82 (2m, 1H, H-3'), 5.00-4.92 (2m, 1H, H-4'), 4.35-4.17 (2m, 2H, H-5'), 3.98-3.88 (m, 1H, Ala CH), 3.64, 3.64 (2s, 3H, OCH_3), 1.82 (bs, 1H, NH), 1.78, 1.74 (2d, J = 1 Hz, 3H, CH_3- 5), 1.32, 1.27 (2d, J = 6.5 Hz, 3H, Ala CH_3); 31P NMR (CDCl_3): \delta 3.10, 2.60; 19F NMR (CDCl_3): \delta 83.21 (p, J = 149.9 Hz, 1F), 62.82, 62.80 (2d, J = 149.9, 150.4 Hz, 4F); 13C NMR (CDCl_3): 173.88, 173.66 (2d, \text{J}_{CP} = 6.5, 7.4 Hz, C=O ester), 163.68, 163.63 (ArC, C-4), 154.41 (m, ArC, CSF_3), 150.76 (C=O, C-2), 150.18, 150.04 (2d, \text{J}_{CP} = 5.8 Hz, ArC ipso phenyl), 135.64, 135.44 (ArCH, C-6), 133.05, 132.77 (C-3'), 129.91 (ArCH), 127.73, 127.63 (C-2'), 123.80, 123.60 (2d, \text{J}_{CP} = 3.8, 5 Hz, ArCH), 122.66 (m, ArCH), 118.55 (m, ArCH), 111.45, 111.27 (ArC, C-5), 89.96, 89.68 (C-1'), 84.48, 84.42 (2d, \text{J}_{CP} = 2.5 Hz, C-4'), 67.58, 66.86 (2d, \text{J}_{CP} = 4, 3.8 Hz, C-5'), 52.69 (OCH_3), 50.25, 50.16 (Ala CH), 20.93, 20.82 (2d, 5, 6.3 Hz, Ala CH_3), 12.35, 12.25 (CH_3, C-5). \text{MS (ES') m/z: } 614.1 (M+Na^+), C_{20}H_{23}F_3N_3O_8PS, \text{Reverse-phase HPLC, eluting with H}_2O/CH_3CN from 90/10 to 0/100 in 30 min; flow = 1 mL/min, t_r = 16.01 min. yield 34%.
\]
S2. Carboxypeptidase Y enzymatic assay procedure

The enzymatic activation of the protides towards carboxypeptidase Y was studied using in situ $^{31}$P NMR. The experiment was carried out by dissolving 5d (5.0 mg) in $d_6$-acetone (0.15 mL) and adding Trizma buffer pH 7.6 (0.30 mL). The resulting solution was placed in an NMR tube and a $^{31}$P NMR experiment at 25 °C was recorded as the blank experiment. The enzyme carboxypeptidase Y (0.1 mg) was dissolved in Trizma (0.15 mL) and added to the solution of the phosphoramidate derivative in the NMR tube. The $^{31}$P-NMR experiment was performed recording the experiment every 7 min at 25 °C.
$^1$H NMR spectrum of ProTide (5a)
$^{31}$P NMR spectrum of ProTide (5a)
$^{13}$C NMR spectrum of ProTide (5a)
HPLC analysis spectrum of ProTide (5a)

Chromatogram : SK75V3_channel1

<table>
<thead>
<tr>
<th>Index</th>
<th>Name</th>
<th>Time [Min]</th>
<th>Quantity [% Area]</th>
<th>Height [mAU]</th>
<th>Area [mAU.Min]</th>
<th>Area % [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>UNKNOWN</td>
<td>10.52</td>
<td>0.44</td>
<td>5.2</td>
<td>0.8</td>
<td>0.437</td>
</tr>
<tr>
<td>2</td>
<td>UNKNOWN</td>
<td>15.00</td>
<td>95.44</td>
<td>716.6</td>
<td>174.4</td>
<td>95.442</td>
</tr>
<tr>
<td>3</td>
<td>UNKNOWN</td>
<td>15.84</td>
<td>0.12</td>
<td>1.8</td>
<td>0.2</td>
<td>0.117</td>
</tr>
<tr>
<td>4</td>
<td>UNKNOWN</td>
<td>16.93</td>
<td>1.31</td>
<td>15.0</td>
<td>2.4</td>
<td>1.315</td>
</tr>
<tr>
<td>5</td>
<td>UNKNOWN</td>
<td>21.32</td>
<td>2.25</td>
<td>46.8</td>
<td>4.1</td>
<td>2.254</td>
</tr>
<tr>
<td>6</td>
<td>UNKNOWN</td>
<td>22.45</td>
<td>0.17</td>
<td>4.9</td>
<td>0.3</td>
<td>0.173</td>
</tr>
<tr>
<td>7</td>
<td>UNKNOWN</td>
<td>23.24</td>
<td>0.26</td>
<td>5.4</td>
<td>0.5</td>
<td>0.261</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>182.8</td>
<td>100.000</td>
</tr>
</tbody>
</table>
$^1$H NMR spectrum of ProTide (5b)
$^{31}$P NMR spectrum of ProTide (5b)
$^{19}$F NMR spectrum of ProTide (5b)
\( ^{13}\text{C} \) NMR spectrum of ProTide (5b)
HPLC analysis spectrum of ProTide 5b

Chromatogram : SK71V3_channel1

System : HPLC_Analytical
Method : SFU-Blanka ACN 30+5min ok
User : User1
Acquisition time: 8/22/2014 11:30:01 AM
Date: 8/22/2014
Time: 12:16:05 PM
Acquired : 8/22/2014 11:30:01 AM
Processed : 8/22/2014 12:14:03 PM
Printed : 8/22/2014 12:16:05 PM
Method: SFU-Blanka ACN 30+5min ok

SK71V3.DATA - Prostar 335 Absorbance Analog Channel 1 EL04099020

Peak results :

<table>
<thead>
<tr>
<th>Index</th>
<th>Name</th>
<th>Time (Min)</th>
<th>Quantity [% Area]</th>
<th>Height (mAU)</th>
<th>Area [mAU.Min]</th>
<th>Area % [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>UNKNOWN</td>
<td>3.15</td>
<td>0.05</td>
<td>2.2</td>
<td>0.2</td>
<td>0.050</td>
</tr>
<tr>
<td>2</td>
<td>UNKNOWN</td>
<td>3.29</td>
<td>0.30</td>
<td>21.9</td>
<td>1.2</td>
<td>0.297</td>
</tr>
<tr>
<td>3</td>
<td>UNKNOWN</td>
<td>17.08</td>
<td>95.15</td>
<td>1706.9</td>
<td>394.5</td>
<td>95.146</td>
</tr>
<tr>
<td>4</td>
<td>UNKNOWN</td>
<td>18.16</td>
<td>0.74</td>
<td>20.9</td>
<td>3.1</td>
<td>0.741</td>
</tr>
<tr>
<td>5</td>
<td>UNKNOWN</td>
<td>18.33</td>
<td>0.85</td>
<td>24.1</td>
<td>3.5</td>
<td>0.852</td>
</tr>
<tr>
<td>6</td>
<td>UNKNOWN</td>
<td>18.88</td>
<td>1.14</td>
<td>34.3</td>
<td>4.7</td>
<td>1.145</td>
</tr>
<tr>
<td>7</td>
<td>UNKNOWN</td>
<td>19.05</td>
<td>1.29</td>
<td>36.3</td>
<td>5.4</td>
<td>1.292</td>
</tr>
<tr>
<td>8</td>
<td>UNKNOWN</td>
<td>19.43</td>
<td>0.48</td>
<td>12.1</td>
<td>2.0</td>
<td>0.478</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>414.7</td>
<td>100.000</td>
</tr>
</tbody>
</table>
$^1$H NMR spectrum of ProTide (5c)
$^{31}$P NMR spectrum of ProTide (5c)
$^{19}$F NMR spectrum of ProTide (Sc)
$^{13}$C NMR spectrum of ProTide (5c)
HPLC analysis spectrum of ProTide 5c

Chromatogram: SK73Vir2_channel1

System: HPLC_Analytical
Method: 5FU-Blanka ACN 30+5min ok
User: User1
Acquisition time: 7/21/2014 11:11:44 AM
Date: 7/21/2014
Time: 11:56:32 AM
Acquired: 7/21/2014 11:11:44 AM
Printed: 7/21/2014 11:56:32 AM
Method: 5FU-Blanka ACN 30+5min ok

Peak results:

<table>
<thead>
<tr>
<th>Index</th>
<th>Name</th>
<th>Time [Min]</th>
<th>Quantity [% Area]</th>
<th>Height [mAU]</th>
<th>Area [mAU Min]</th>
<th>Area [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>UNKNOWN</td>
<td>2.36</td>
<td>0.07</td>
<td>1.1</td>
<td>0.5</td>
<td>0.06</td>
</tr>
<tr>
<td>2</td>
<td>UNKNOWN</td>
<td>2.55</td>
<td>0.05</td>
<td>3.0</td>
<td>0.4</td>
<td>0.050</td>
</tr>
<tr>
<td>3</td>
<td>UNKNOWN</td>
<td>2.91</td>
<td>0.02</td>
<td>1.9</td>
<td>0.1</td>
<td>0.018</td>
</tr>
<tr>
<td>4</td>
<td>UNKNOWN</td>
<td>3.20</td>
<td>0.01</td>
<td>0.9</td>
<td>0.1</td>
<td>0.009</td>
</tr>
<tr>
<td>5</td>
<td>UNKNOWN</td>
<td>13.07</td>
<td>0.03</td>
<td>2.0</td>
<td>0.2</td>
<td>0.028</td>
</tr>
<tr>
<td>6</td>
<td>UNKNOWN</td>
<td>14.97</td>
<td>99.72</td>
<td>3491.3</td>
<td>810.1</td>
<td>99.721</td>
</tr>
<tr>
<td>7</td>
<td>UNKNOWN</td>
<td>25.60</td>
<td>0.11</td>
<td>6.3</td>
<td>0.9</td>
<td>0.107</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>100.00</td>
<td>3517.0</td>
<td>812.3</td>
<td>100.000</td>
<td></td>
</tr>
</tbody>
</table>
$^1$H NMR spectrum of ProTide (5d)
$^{31}\text{P} \text{ NMR spectrum of ProTide (5d)}$
$^{19}$F NMR spectrum of ProTide (5d)
$^{13}$C NMR spectrum of ProTide (5d)
HPLC analysis spectrum of ProTide 5d

Chromatogram: SK57Vir2_channel1

System: HPLC_Analytical
Method: SFU-Blanka ACN 30+5min ok
User: User1
Acquisition time: 7/21/2014 11:59:40 AM

Date: 7/21/2014
Time: 12:57:36 PM

Acquired: 7/21/2014 11:59:40 AM
Processed: 7/21/2014 12:57:23 PM
Printed: 7/21/2014 12:57:36 PM
Method: SFU-Blanka ACN 30+5min ok

Peak results:

<table>
<thead>
<tr>
<th>Index</th>
<th>Name</th>
<th>Time (Min)</th>
<th>Quantity [% Area]</th>
<th>Height [mAU]</th>
<th>Area [mAU Min]</th>
<th>Area [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>UNKNOWN</td>
<td>15.16</td>
<td>2.06</td>
<td>2.9</td>
<td>0.5</td>
<td>2.055</td>
</tr>
<tr>
<td>1</td>
<td>UNKNOWN</td>
<td>16.01</td>
<td>97.15</td>
<td>102.0</td>
<td>24.2</td>
<td>97.145</td>
</tr>
<tr>
<td>2</td>
<td>UNKNOWN</td>
<td>25.81</td>
<td>0.80</td>
<td>2.1</td>
<td>0.2</td>
<td>0.800</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100.000</td>
</tr>
</tbody>
</table>

SK57Vir2.DATA - Prostar 335 Absorbance Analog Channel 1 EL04099020