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

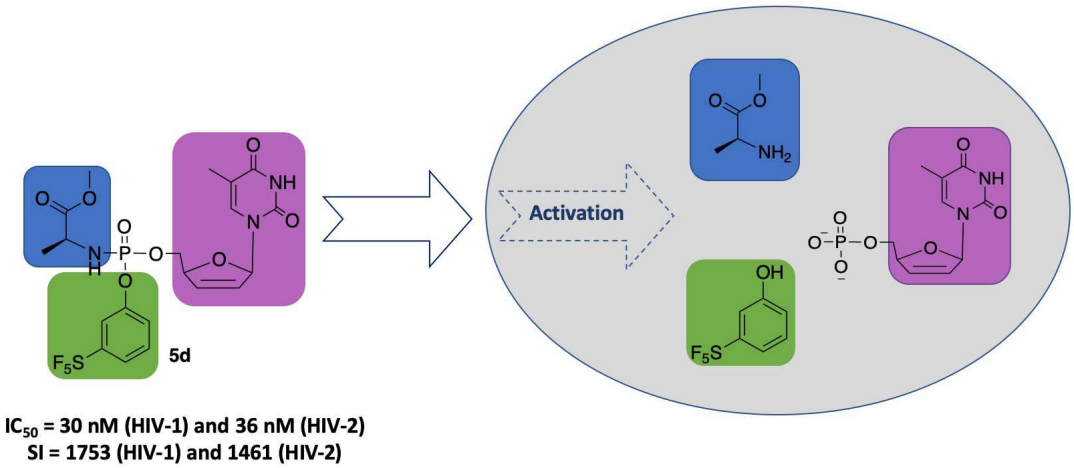
Sahar Kandil[a], Christophe Pannecouque[b], Fiona M. Chapman[a], Andrew D. Westwell[a] and Christopher McGuigan[a]

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

Keywords:

Polyfluoroaromatic, pentafluorosulfanyl, SF_5, Stavudine, d4T, ProTide, Phosphoramidate, Antiviral, Nucleoside, HIV.

Graphical abstract



Highlights

- Synthesis of polyfluoroaromatic Stavudine (d4T) ProTides.
- Anti-retroviral activity against two strains; HIV-1, (III_B) 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-SF_5 aromatic ring; IC_{50} = 30 nM, HIV-1 and IC_{50} = 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 ^{31}P and ^{19}F 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 III_B) and HIV-2 (strain ROD). ProTide **5d** containing a *meta*-substituted pentafluorosulfanyl (3-SF_5) aryl group[17] showed superior antiviral activity over the parent d4T and the non-fluorinated analogue **5a**. ProTide **5d** has low nanomolar antiviral activity; (IC_{50} = 30 nM, HIV-1) and (IC_{50} = 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 ^{31}P and ^{19}F 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[1]. 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[2]. 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[3]. Nucleoside / nucleotide reverse-transcriptase inhibitors (NRTIs/NtRTIs) represent the first major class of antiretroviral drugs developed [4].

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 [5-7]. 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 [5,6]. 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 [8,9]. This is demonstrated clinically by sofosbuvir (anti-HCV, FDA approved) [6] and acelarin (anticancer, Phase III) [9], shown in **Figure 1**.

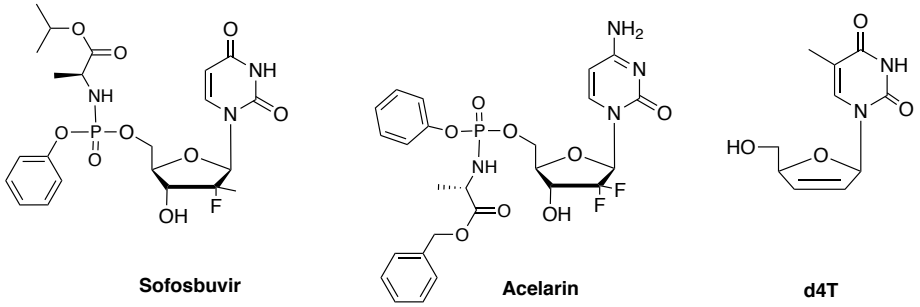


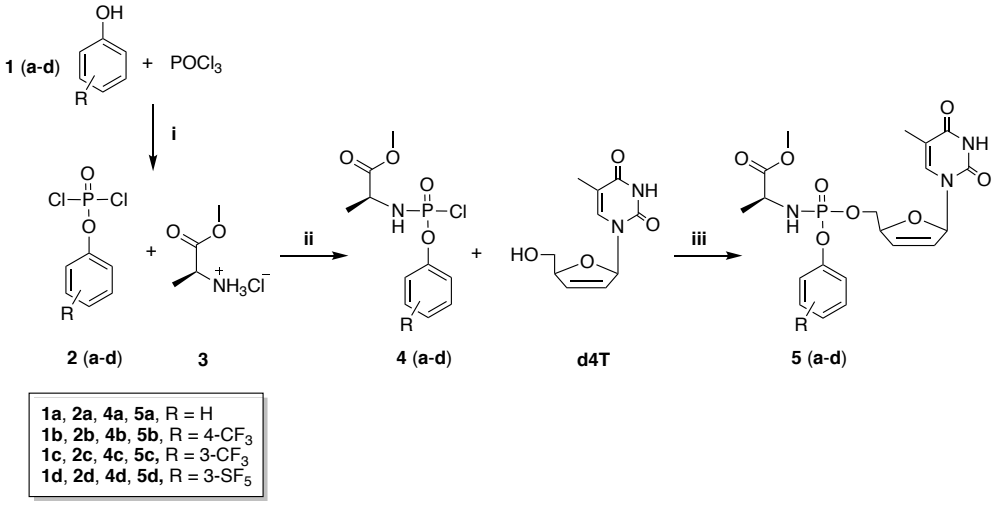
Figure 1: Chemical structures of sofosbuvir (anti-HCV, FDA approved), acelarin (anticancer, Phase III) ProTides and stavudine (d4T) nucleoside.

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[10]. In an attempt to overcome the first intracellular thymidine kinase-directed phosphorylation step of d4T, phosphoramidate derivatives were prepared and evaluated against HIV[11].

Generally, introduction of fluorinated substituents can provide a unique combination of electronegativity, size and lipophilicity and can also greatly affect physicochemical properties[12-14]. 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-CF_3 / 3-CF_3 / 3-SF_5) 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.

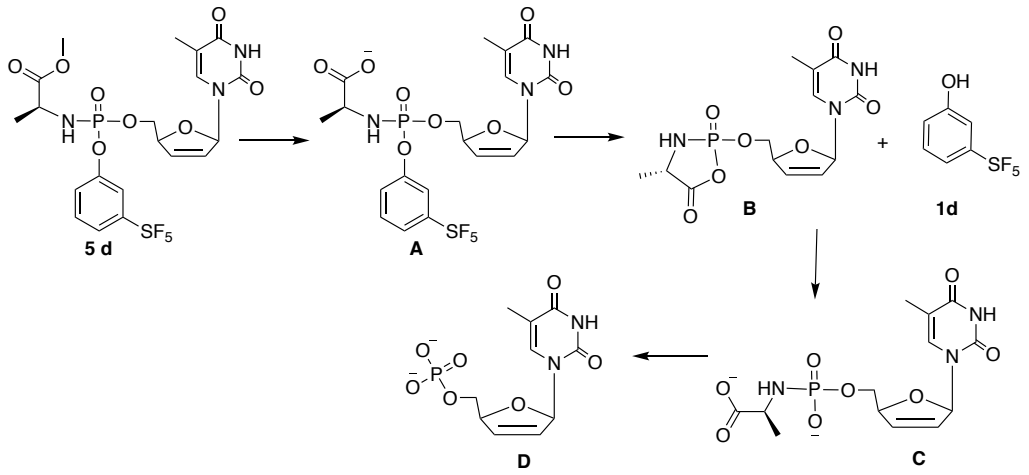
Stavudine (d4T) ProTides were prepared using previously described phosphorochloridate chemistry[15]. The phenyl phosphorodichloridates **2a**-**d** were prepared by the reaction of substituted phenols **1a**-**d** with $POCl_3$. The subsequent reaction with methyl alanine **3** to form the corresponding phosphorochloridates **4a**-**d** was monitored by ^{31}P NMR. Next the arylaminoacyl phosphorochloridates **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%)[16]. 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_3N, anhydrous Et_2O, -78 °C to rt, 2 h, ii) Et_3N, 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-d_6 in the presence of Trizma buffer (pH 7.6), treated with carboxypeptidase Y and monitored by ^{31}P 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**.

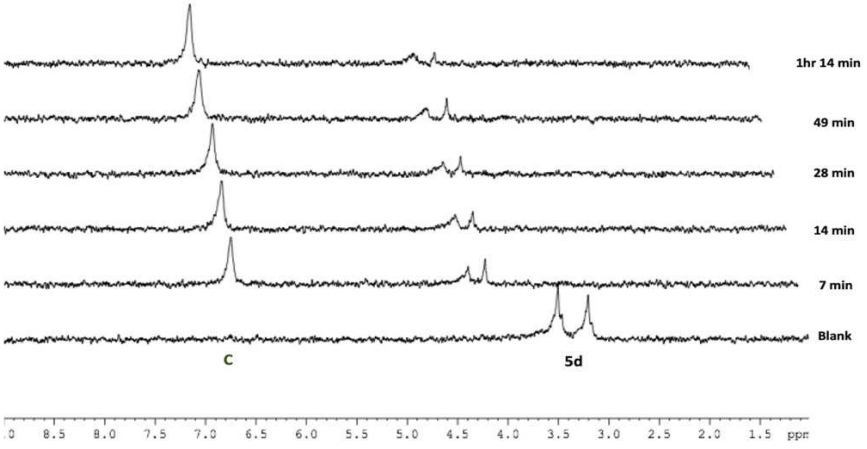


Figure 2. ^{31}P NMR spectra of **d4T** ProTide **5d** over time after carboxypeptidase Y treatment showing the signals of the metabolite **C** over time.

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 δ 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 δ 61.86 and 62.21 (corresponding to the four fluorine atoms of the SF5 group) indicating the release of 3-SF_5 phenol (**1d**).

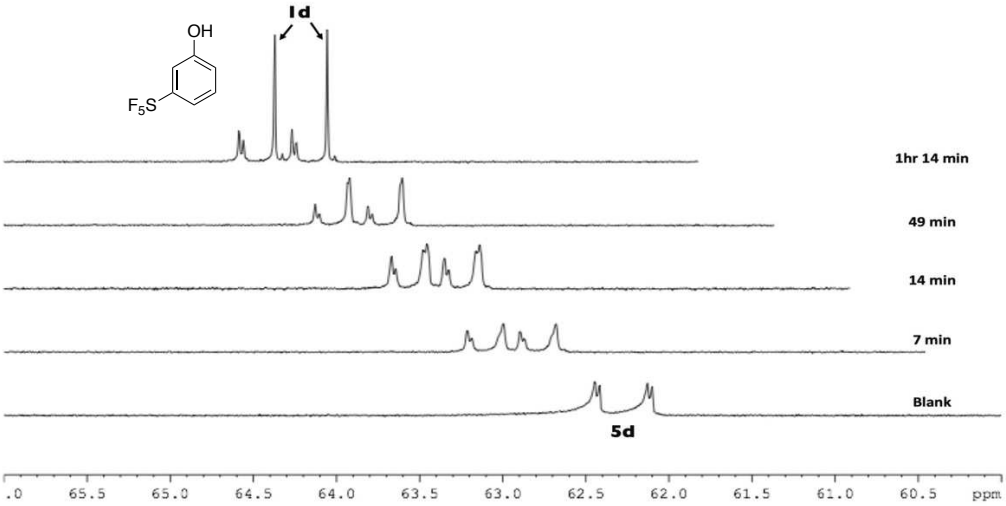


Figure 3. ^{19}F NMR spectra of ProTide **5d** over time after carboxypeptidase Y treatment showing the signals of the metabolites over time.

Both ^{31}P NMR and ^{19}F NMR monitoring experiments, **Figure 2** and **3**, indicate that ProTide **5d** activation is initiated by the carboxpeptidase 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[4]. 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 (III_B) 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-SF_5) substituent attached to its aromatic ring. ProTide **5d** has low nanomolar antiviral activity; (IC_{50} = 30 nM, HIV-1) and (IC_{50} = 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 CF_3 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*-CF_3 (**5c**) over the *para*-CF_3 (**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**.

	HIV-1 (III_B)			**HIV-2 (ROD)**		
comp	**IC_{50} (μM)**	**CC_{50} (μM)**	**SI**	**IC_{50} (μM)**	**CC_{50} (μM)**	**SI**
d4T	0.40±0.09	313±7	782	0.45±0.09	313±7	695
5a	0.21±0.09	251±7	1195	0.27±0.11	251±7	930
5b	0.037±0.002	42.9±12.6	1159	0.047±0.022	42.9±12.6	912
5c	0.041±0.006	66.1±24.8	1612	0.049±0.017	66.1±24.8	1349
5d	0.030±0.012	52.6±4.9	1753	0.036±0.020	52.6±4.9	1461

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-CF_3) has around 180-fold higher activity than d4T, while ProTide **5d** (3-SF_5) 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** (IC_{50} = 0.039 μM) shows better activity than **5a** (IC_{50} = 0.30 μM) and d4T (IC_{50} = 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**.

	HIV-1 (III_B)

ID	IC_{50} WT (μM)	IC_{50} TK^- (μM)
d4T	4.06±0.18	NA
5a	0.30±0.04	0.32±0.15
5b	0.022±0.006	0.075±0.000
5c	0.039±0.007	NA
5d	0.054±0.039	0.335±0.004

Table 2. Antiviral activity of stavudine (d4T) and ProTide analogues **5a**-**d** against HIV-1 strain III_B in the C8166 cell line and the thymidine kinase deficient (C8166-TK^-) cell line. NA: not active

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.

Compound	MW	ClogP[a]	Estimated Precipitation range (μM)			Estimated solubility (μM)
			Lower bound	Upper bound	Calculated mid-range	
d4T	224.22	-0.4875	100	100	100	> 100
5b	533.40	1.9788	100	100	100	> 100
5c	533.40	1.9788	100	100	100	> 100
5d	591.44	2.3288	100	100	100	> 100
Nicardipine	-	-	10	30	20	10-30
pyrene	-	-	3	10	6.5	3-10

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

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 III_B) and HIV-2 (strain ROD). Our findings identified compound **5d** featuring a *meta*-pentafluorosulfanyl (SF_5) in its aromatic ring as promising lead. ProTide **5d** has low nanomolar antiviral activity; (IC_{50} = 30 nM, HIV-1) and (IC_{50} = 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

1. U. Sandkovsky, K. Robertson, J. Meza, HIV Clinical Trials, 14(4) (2014) 165–174.
2. Hontelez JA, Vlas SD, R. Baltussen, AIDS, (2012); 26(1): S19–S30.
3. Z. Zhou, T. Liu, G. Wu, D. Kang, Z. Fu, Z. Wang, E. De Clercq, C. Pannecouque, P. Zhan and X. Liu, *Org. Biomol. Chem.*, 17 (2019), 3202-3217.
4. S. Pawar, S. Waghmare, B. Gawde, P. Kale and Y. Johri, *Journal of Drug Delivery and Therapeutics*. 9, 3 (2019), 599-608.
5. F.M. Oliveira, L.C.A. Barbosa, F.M.D. Ismail, RSC Adv. 4 (2014) 18998.
6. U. Pradere, E.C. Garnier-Amblard, S.J. Coats, F. Amblard, R.F. Schinazi, Chem. Rev. 114 (2014) 9154.
7. A.J. Wiemer, D.F. Wiemer, Top. Curr. Chem. 360 (2015) 115.
8. C. McGuigan, R.N. Pathirana, N. Mahmood, A.J. Hay, Bioorg. Med. Chem. Lett. 2(1992) 701–704.
9. M. Slusarczyk, M.H. Lopez, J. Balzarini, M. Mason, W.G. Jiang, S. Blagden, E. Thompson, E. Ghazaly, C. McGuigan, J. Med. Chem. 57 (4) (2014) 1531.

10. M. J. Browne, K. H. Mayer, S. B. D. Chafee, M. N. Dudley, M. R. Posner, S. M. Steinberg, K. K. Graham, S. M. Geletko, S. H. Zinner, S. L. Denman, L. M. Dunkle, S. Kaul, C. McLaren, G. Skowron, N. M. Kouttab, T. A. Kennedy, A. B. Weitberg, G. A. Curt, *The Journal of Infectious Diseases*, 167, 1 (1993) 21–29.
11. C. McGuigan, L. Bidois, A. Hiouni, C. Ballatore, E.D. Clercq and J. Balzarini, Antiviral Chemistry & Chemotherapy 11 (2000) 111–116.
12. C. McGuigan. PCT. Int. Appl. 2005 WO 2005012327 A2 20050210.
13. P. Shah, A. D. Westwell J. Enz. Inhib. Med. Chem., 22 (2007), 527-540.
14. S. Kandil, A. D. Westwell, C. McGuigan, Bioorg. Med. Chem. Lett. 26, 8 (2016) 2000-2004.
15. S. Kandil, J. Balzarini, S. Rat, A. Brancale, A. D. Westwell, C. McGuigan, Bioorg. Med. Chem. Lett. 26 (2016) 5618–5623.

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 $MgSO_4$, filtered, reduced to dryness and purified by flash chromatography.

17. **Spectral data of ProTide 2',3'-Didehydro-2',3'-dideoxythymidine-5'-(3-(penta fluorosulfanyl) phenyl methoxy alaninyl phosphate) (5d);** ^{1}H NMR ($CDCl_3$): δ 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); ^{31}P NMR ($CDCl_3$): δ 3.10, 2.60; ^{19}F NMR ($CDCl_3$): δ 83.21 (p, J = 149.9 Hz, 1F), 62.82, 62.80 (2d, J = 149.9, 150.4 Hz, 4F); ^{13}C NMR ($CDCl_3$): 173.88, 173.66 (2d, $^{3}J_{C-P}$ = 6.5, 7.4 Hz, C=O ester), 163.68, 163.63 (ArC, C-4), 154.41 (m, ArC, CSF_5), 150.76 (C=O, C-2), 150.18, 150.04 (2d, $^{2}J_{C-P}$ = 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, $^{2}J_{C-P}$ = 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, $^{3}J_{C-P}$ = 2.5 Hz, C-4'), 67.58, 66.86 (2d, $^{2}J_{C-P}$ = 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+), $C_{20}H_{23}F_5N_3O_8PS$, 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%.

Supporting Information

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

Sahar Kandil[a], Christophe Pannecouque[b], Fiona M. Chapman[a], Andrew D. Westwell[a] and Christopher McGuigan[a]

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

Contents

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_2O/ CH_3CN 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)

^{1}H NMR ($CDCl_3$): *δ* 8.80, 8.76 (2bs, 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, OCH_3), 2.00 (bs, 1H, NH), 1.86, 1.82 (2d, J = 0.3 Hz, 3H, CH_3- 5), 1.36, 1.32 (2d, J = 7 Hz, 3H, Ala CH_3); ^{31}P NMR ($CDCl_3$): *δ* 3.10, 2.49; ^{13}C NMR ($CDCl_3$): 173.96, 173.82 (2d, $^{3}J_{C-P}$ = 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, $^{2}J_{C-P}$ = 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, $^{2}J_{C-P}$ = 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, $^{2}J_{C-P}$ = 5, 3.8 Hz, C-5′), 52.60, 52.58 (OCH_3), 50.21, 50.09 (AlaCH), 20.96, 20.90 (2d, 5 Hz, Ala CH_3), 12.34, 12.30 (CH_3, C-5). MS (ES^+) m/z: 488.1 ($M+Na^+$), $C_{20}H_{24}N_3O_8P$, Reverse-phase HPLC, eluting with H_2O/CH_3CN from 90/10 to 0/100 in 30 min; flow = 1 mL/min, t_R = 15.00 min. yield 28%.

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

^{1}H NMR ($CDCl_3$): *δ* 9.24, 9.21 (2bs, 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, OCH_3), 1.86, 1.81 (2d, J = 1.5 Hz, 3H, CH_3- 5), 1.38, 1.35 (2d, J = 7 Hz, 3H, Ala CH_3); ^{31}P NMR ($CDCl_3$): *δ* 3.25, 2.63; ^{19}F NMR ($CDCl_3$): δ -62.74, -62.75; ^{13}C NMR ($CDCl_3$): 173.90, 173.74 (2d, $^{3}J_{C-P}$ = 7.5 Hz, C=O ester), 163.83, 163.79 (ArC-4), 150.86 (C=O, C-2), 150.63, 150.50 (2d, $^{2}J_{C-P}$ = 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, $^{2}J_{C-P}$ = 3.8, 5 Hz, ArCH), 123.39 (q, $^{1}J_{C-P}$ = 270.8 Hz, CF_3), 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, $^{2}J_{C-P}$ = 3.8, 5 Hz, C-4′), 67.47, 66.80 (2d, $^{2}J_{C-P}$ = 3.8, 5 Hz, C-5′), 52.64 (OCH_3), 50.24, 50.13 (AlaCH), 20.89, 20.80 (2d, 5. 6.3 Hz, Ala CH_3), 12.34, 12.26 (CH_3, C-5). MS (ES^+) m/z: 556.1 ($M+Na^+$), $C_{21}H_{23}F_3N_3O_8P$, Reverse-phase HPLC, eluting with H_2O/CH_3CN from 90/10 to 0/100 in 30 min; flow = 1 mL/min, t_R = 14.97 min. yield 20%.

2′,3′-Didehydro-2′,3′-dideoxythymidine-5′(4-(trifluoromethyl)phenyl methoxy alaninyl phosphate) (5c).

^{1}H NMR ($CDCl_3$): *δ* 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); ^{31}P NMR ($CDCl_3$): *δ* 2.95, 2.41; ^{19}F NMR ($CDCl_3$): δ -62.21, -62.22; ^{13}C NMR ($CDCl_3$): 173.74, 173.60 (C=O ester), 163.44, 163.41 (ArC-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, $^{2}J_{C-P}$ = 3.8, 5 Hz, C-5′), 52.69 (OCH_3), 50.21,50.14 (AlaCH), 20.98, 20.90 (2d, 5 Hz, Ala CH_3), 12.39, 12.36 (CH_3, C-5). MS (ES^+) m/z: 556.1 ($M+Na^+$), $C_{21}H_{23}F_3N_3O_8P$, 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%.

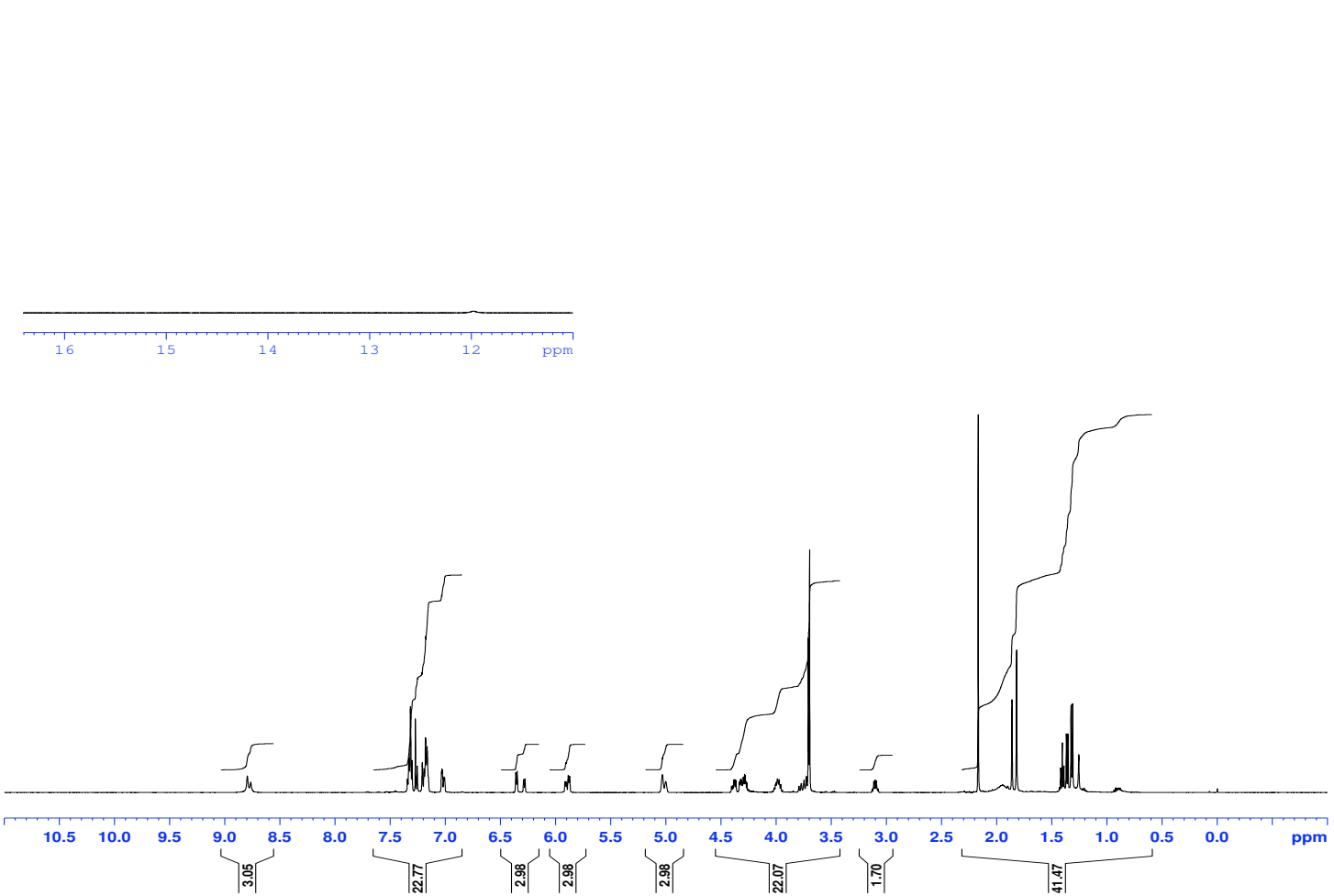
2′,3′-Didehydro-2′,3′-dideoxythymidine-5′-(3-(pentafluorosulfanyl)phenyl methoxy alaninyl phosphate) (5d)

^{1}H NMR ($CDCl_3$): *δ* 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); ^{31}P NMR ($CDCl_3$): *δ* 3.10, 2.60; ^{19}F NMR ($CDCl_3$): δ 83.21 (p, J = 149.9 Hz, 1F), 62.82, 62.80 (2d, J = 149.9, 150.4 Hz, 4F); ^{13}C NMR ($CDCl_3$): 173.88, 173.66 (2d, $^{3}J_{C-P}$ = 6.5, 7.4 Hz, C=O ester), 163.68, 163.63 (ArC, C-4), 154.41 (m, ArC, CSF_5), 150.76 (C=O, C-2), 150.18, 150.04 (2d, $^{2}J_{C-P}$ = 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, $^{2}J_{C-P}$ = 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, $^{3}J_{C-P}$ = 2.5 Hz, C-4′), 67.58, 66.86 (2d, $^{2}J_{C-P}$ = 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+), $C_{20}H_{23}F_5N_3O_8PS$, 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

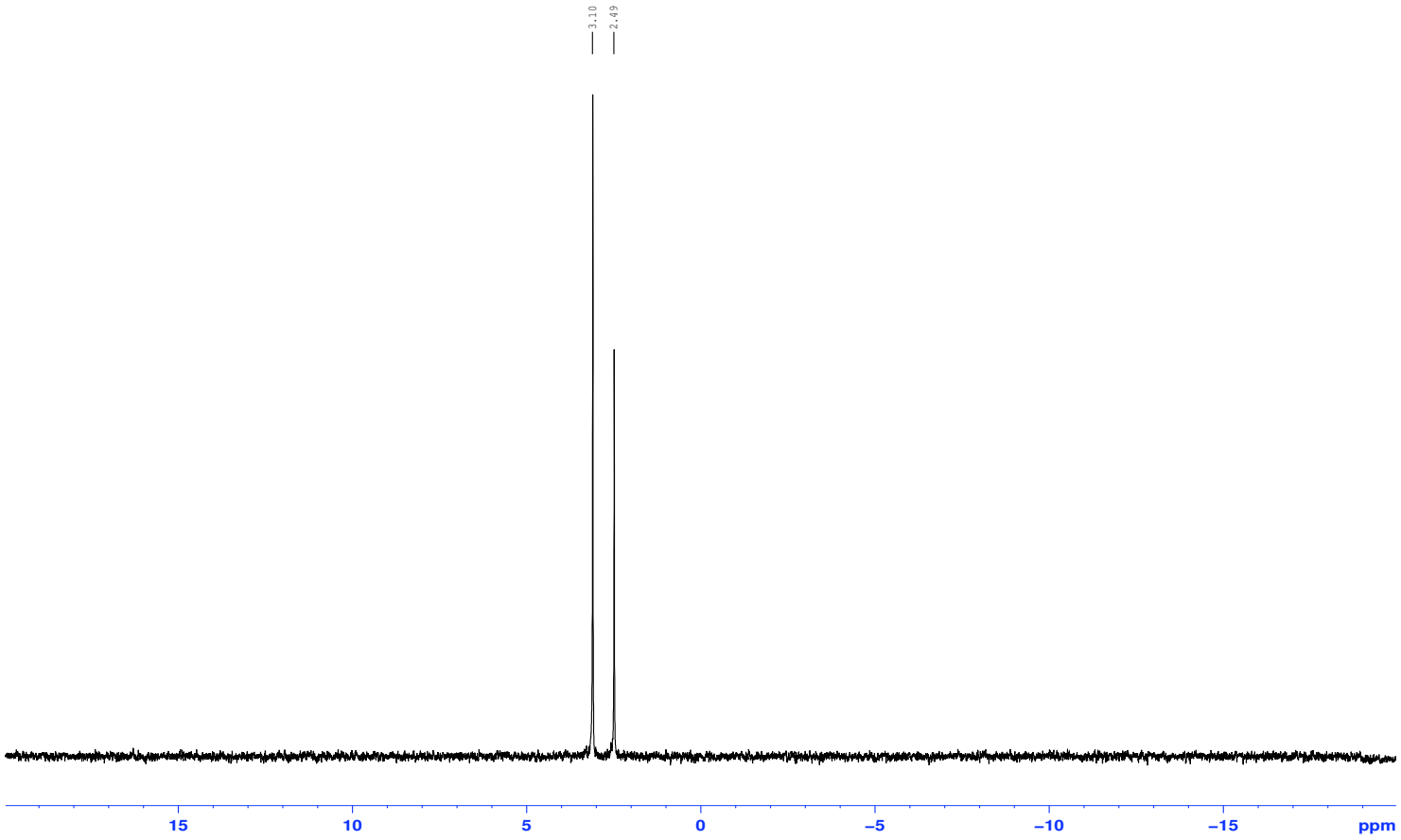
The enzymatic activation of the protides towards carboxypeptidase Y was studied using *in situ* ^{31}P NMR. The experiment was carried out by dissolving **5d** (5.0 mg) in d_6-acetone (0.15 mL) and adding Trizma buffer pH 7.6 (0.30 mL). The resulting solution was placed in an NMR tube and a ^{31}P NMR 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)



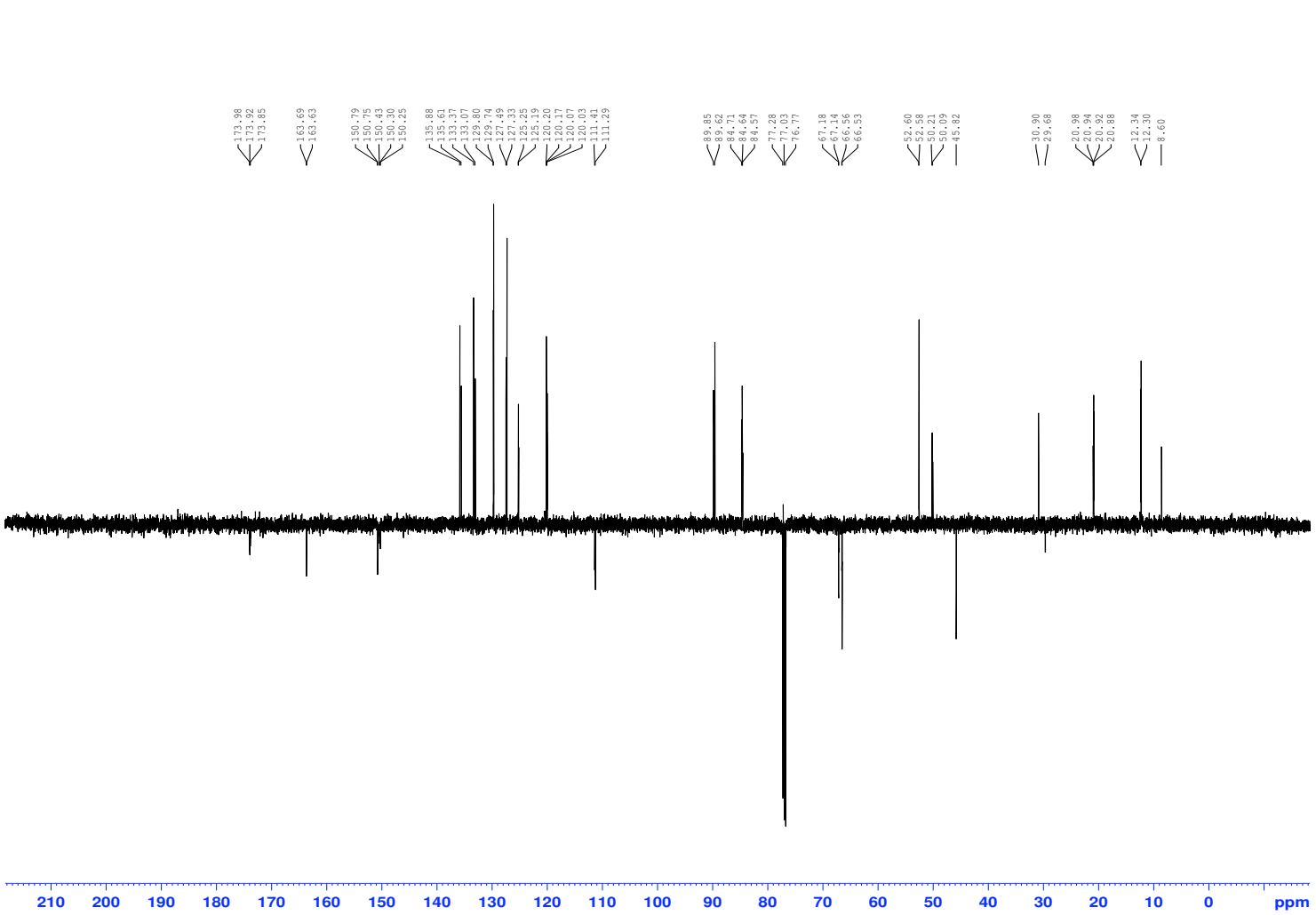

```
NAME            CM-SK75Vir
EXPNO                    5
PROCNO                   1
Date_             20140809
Time                 12.47
INSTRUM         Avance500
PROBHD    5 mm QNP 1H/13
PULPROG               zg30
TD                   65536
SOLVENT              CDCl3
NS                      16
DS                       2
SWH              10330.578 Hz
FIDRES            0.157632 Hz
AQ                3.1719923 sec
RG                     287
DW                  48.400 use
DE                   6.50 use
TE                  298.1 K
D1              1.00000000 sec
TD0                      1

======== CHANNEL f1 =======
NUC1                    1H
P1                   11.50 use
PL1                  -1.00 dB
PL1W          11.38419914 W
SFO1         500.1330885 MHz
SI                   32768
SF           500.1300083 MHz
WDW                     EM
SSB                      0
LB                    0.30 Hz
GB                       0
PC                    1.00
```

^{31}P NMR spectrum of ProTide (5a)



NAME	CM-SK75Vir
EXPNO	1
PROCNO	1
Date_	20140730
Time	13.54
INSTRUM	Avance500
PROBHD	5 mm QNP 1H/13
PULPROG	zgpg30
TD	16384
SOLVENT	CDC13
NS	16
DS	4
SWH	8081.896 Hz
FIDRES	0.493280 Hz
AQ	1.0136734 sec
RG	3250
DW	61.867 use
DE	12.00 use
TE	298.1 K
D1	1.65999997 sec
D11	0.03000000 sec
TD0	1
======== CHANNEL f1 =======	
NUC1	31P
P1	10.40 use
PL1	4.00 dB
PL1W	24.34000015 W
SFO1	202.4563350 MHz
======== CHANNEL f2 =======	
CPDPRG2	waltz16
NUC2	1H
PCPD2	80.00 use
PL2	-2.00 dB
PL12	14.85 dB
PL13	120.00 dB
PL2W	14.33185768 W
PL12W	0.29600734 W
PL13W	0.00000000 W
SFO2	500.1320005 MHz
SI	32768
SF	202.4563350 MHz
WDW	EM
SSB	0
LB	1.00 Hz
GB	0
PC	1.40

^{13}C NMR spectrum of ProTide (5a)



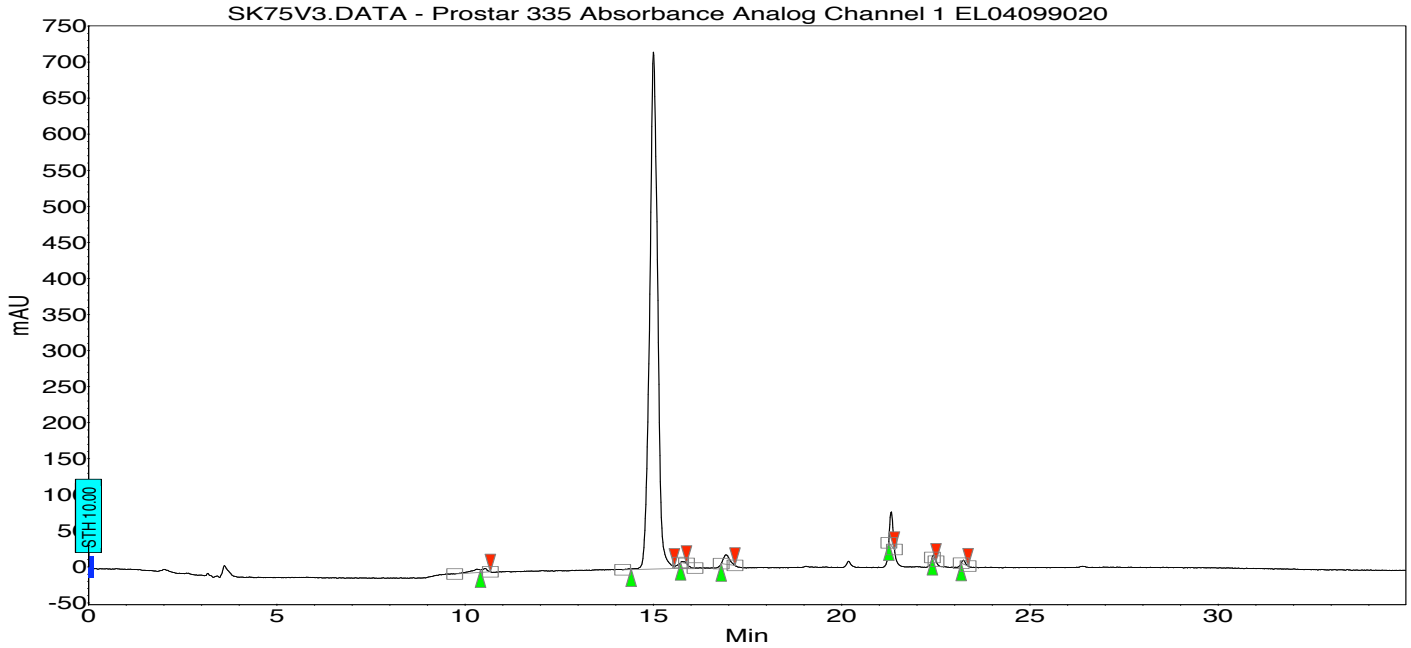

```
NAME                CM-SK75Vir
EXPNO                        4
PROCNO                       1
Date_                 20140809
Time                     12.24
INSTRUM             Avance500
PROBHD    5 mm QNP 1H/13
PULPROG                pendant
TD                       65536
SOLVENT                  CDC13
NS                         512
DS                           4
SWH                  29761.904 Hz
FIDRES                0.454131 Hz
AQ                    1.1010548 sec
RG                        3250
DW                       16.800 usec
DE                       12.00 usec
TE                       298.1 K
CNST2               145.0000000
D1                    2.00000000 sec
D4                    0.00172414 sec
D12                   0.00002000 sec
D15                   0.00431034 sec
D20                   0.00345000 sec
TD0                          4

======== CHANNEL f1 ========
NUC1                        13C
P1                        7.20 usec
P2                       14.40 usec
PL1                      -2.00 dB
PL1W               101.27846527 W
SFO1                125.7703643 MHz

======== CHANNEL f2 ========
CPDPRG2                waltz16
NUC2                        1H
P3                       11.50 usec
P4                       23.00 usec
PCPD2                    80.00 usec
PL2                      -2.00 dB
PL12                     14.85 dB
PL2W               14.33185768 W
PL12W               0.29600734 W
SFO2                500.1320005 MHz
SI                        32768
SF                  125.7577890 MHz
WDW                          EM
SSB                          0
LB                       1.00 Hz
GB                           0
PC                       1.40
```

HPLC analysis spectrum of ProTide (5a)

Chromatogram : SK75V3_channel1

System : HPLC_Analytical
Method : 5FU-Blanka ACN 30+5min ok
User : User1
Acquisition time: 8/22/2014 4:43:10 PM
DAte: 8/22/2014
Time: 5:26:41 PM

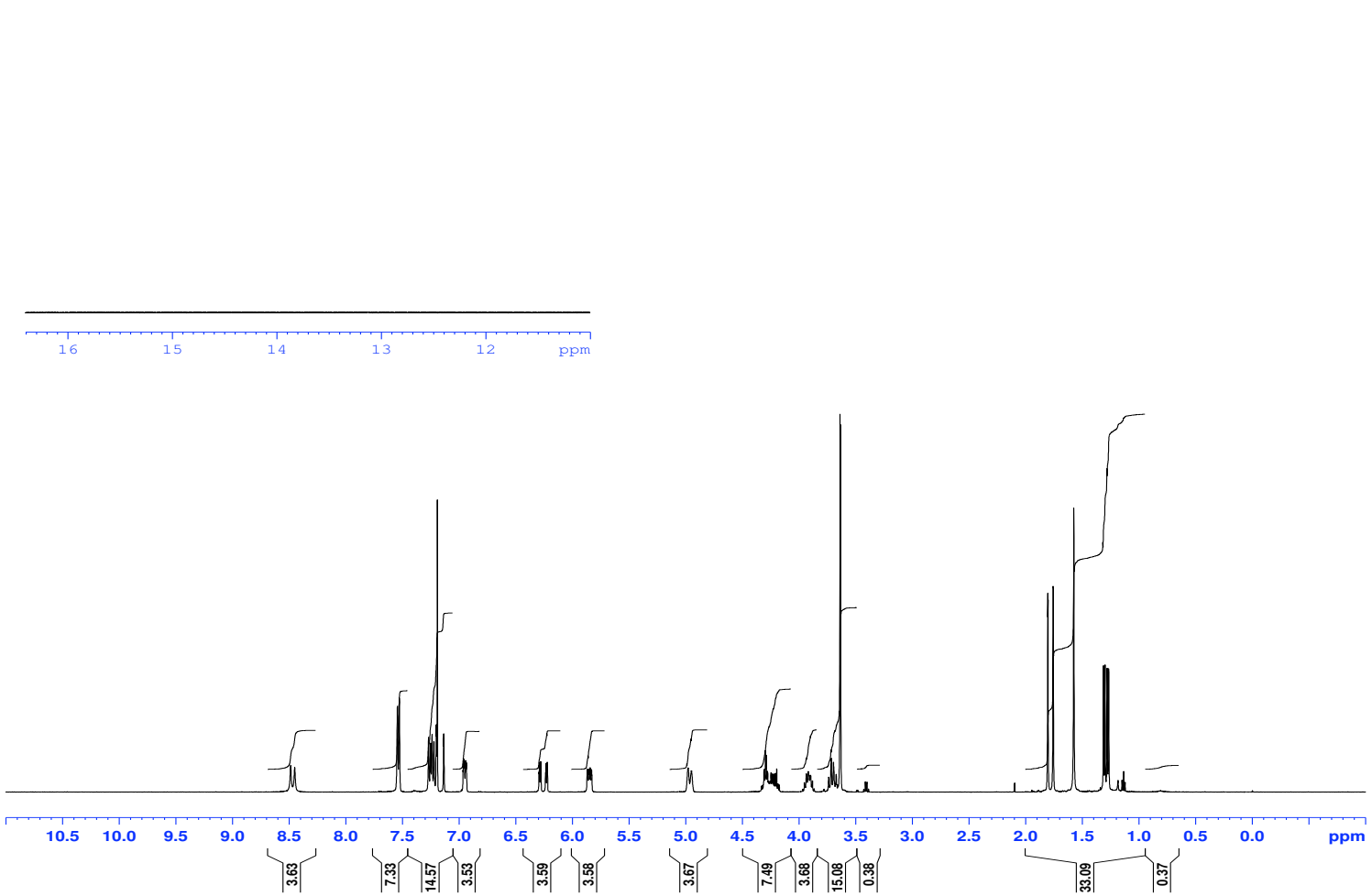
Acquired : 8/22/2014 4:43:10 PM
Processed : 8/22/2014 5:26:08 PM
Printed : 8/22/2014 5:26:41 PM
Method: 5FU-Blanka ACN 30+5min ok



Peak results :

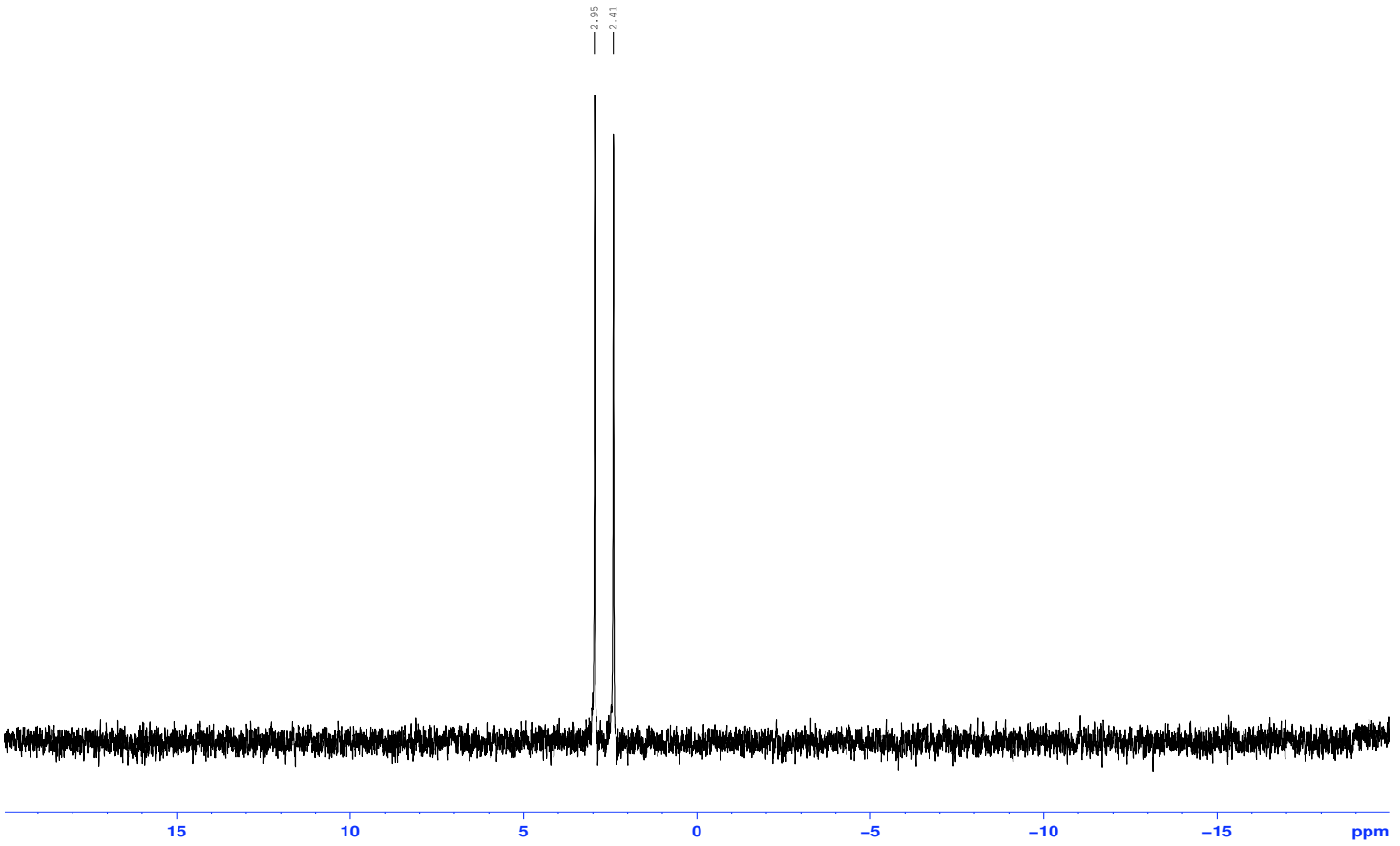
Index	Name	Time [Min]	Quantity [% Area]	Height [mAU]	Area [mAU.Min]	Area % [%]
1	UNKNOWN	10.52	0.44	5.2	0.8	0.437
2	UNKNOWN	15.00	95.44	716.6	174.4	95.442
3	UNKNOWN	15.84	0.12	1.8	0.2	0.117
4	UNKNOWN	16.93	1.31	13.0	2.4	1.315
5	UNKNOWN	21.32	2.25	46.8	4.1	2.254
6	UNKNOWN	22.45	0.17	4.9	0.3	0.173
7	UNKNOWN	23.24	0.26	5.4	0.5	0.261
Total			100.00	793.7	182.8	100.000

^{1}H NMR spectrum of ProTide (5b)




```
NAME            CM-SK71Vir
EXPNO                   11
PROCNO                   1
Date_             20140810
Time                 10.40
INSTRUM         Avance500
PROBHD    5 mm QNP 1H/13
PULPROG               zg30
TD                   65536
SOLVENT              CDCl3
NS                      16
DS                       2
SWH              10330.578 Hz
FIDRES           0.157632 Hz
AQ               3.1719923 sec
RG                     645
DW                  48.400 use
DE                   6.50 use
TE                  298.1 K
D1               1.00000000 sec
TD0                      1

======== CHANNEL f1 =======
NUC1                    1H
P1                   11.50 use
PL1                  -1.00 dB
PL1W          11.38419914 W
SFO1          500.1330885 MHz
SI                   32768
SF            500.1300466 MHz
WDW                     EM
SSB                      0
LB                0.30 Hz
GB                       0
PC                    1.00
```

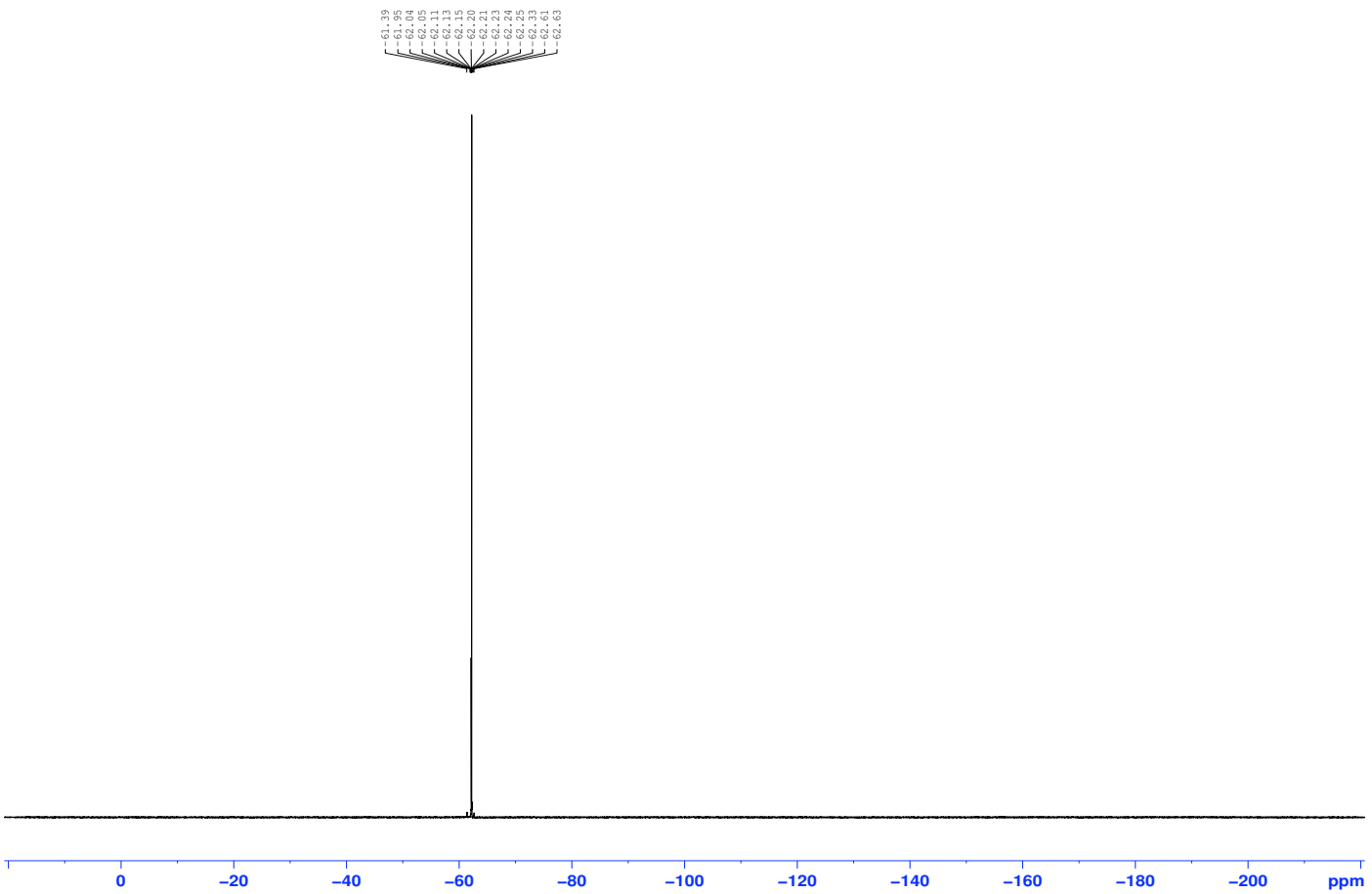
^{31}P NMR spectrum of ProTide (5b)



CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

NAME	CM-SK71Vir	
EXPNO	9	
PROCNO	1	
Date_	20140721	
Time	13.19	
INSTRUM	Avance500	
PROBHD	5 mm QNP 1H/13	
PULPROG	zgpg30	
TD	16384	
SOLVENT	CDCl3	
NS	16	
DS	4	
SWH	8081.896	Hz
FIDRES	0.493280	Hz
AQ	1.0136734	sec
RG	3250	
DW	61.867	use
DE	12.00	use
TE	298.1	K
D1	1.65999997	sec
D11	0.03000000	sec
TD0	1	
======== CHANNEL f1 =======		
NUC1	31P	
P1	10.40	use
PL1	4.00	dB
PL1W	24.34000015	W
SFO1	202.4563350	MHz
======== CHANNEL f2 =======		
CPDPRG2	waltz16	
NUC2	1H	
PCPD2	80.00	use
PL2	−2.00	dB
PL12	14.85	dB
PL13	120.00	dB
PL2W	14.33185768	W
PL12W	0.29600734	W
PL13W	0.00000000	W
SFO2	500.1320005	MHz
SI	32768	
SF	202.4563350	MHz
WDW	EM	
SSB	0	
LB	1.00	Hz
GB	0	
PC	1.40	

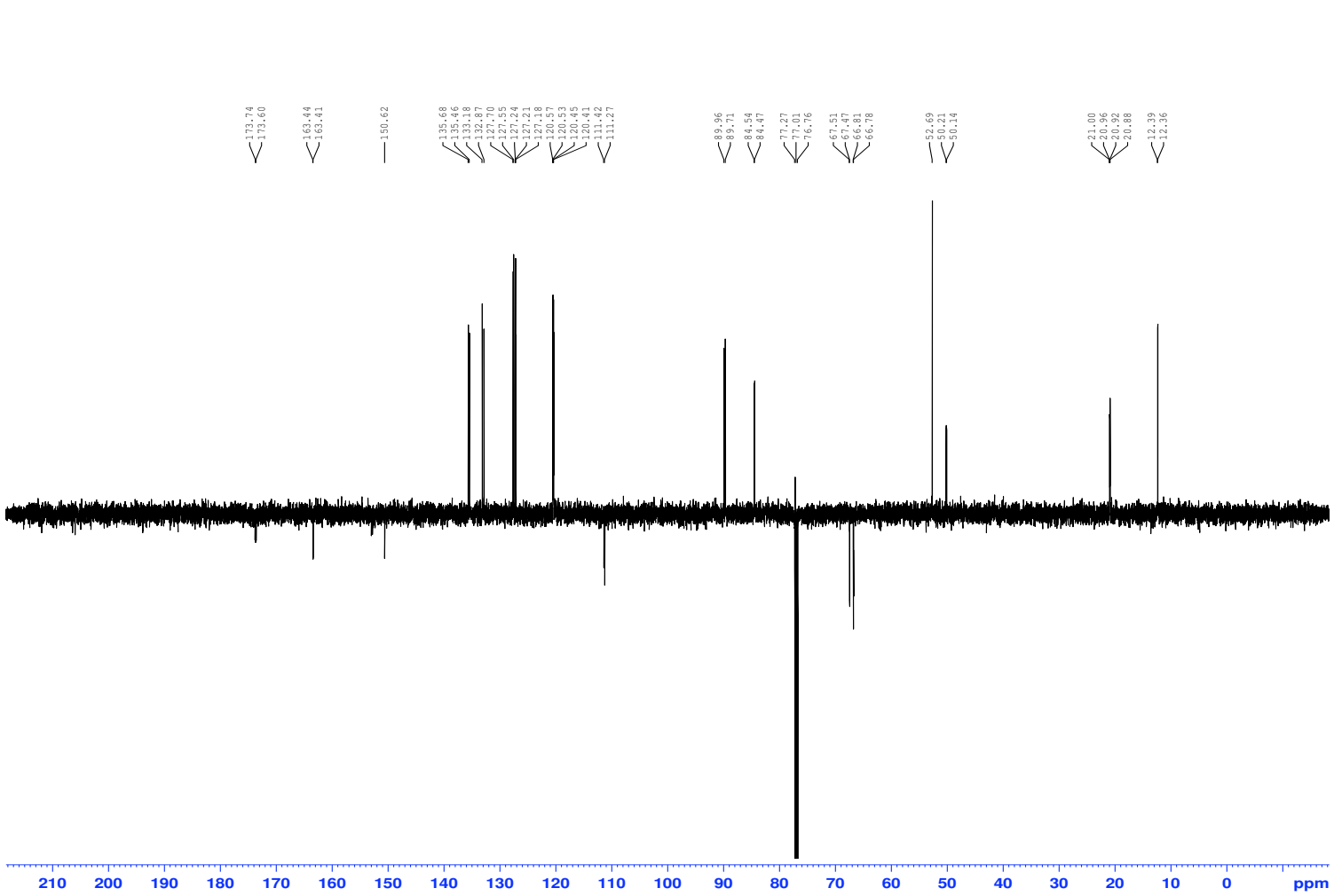
^{19}F NMR spectrum of ProTide (5b)



CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

NAME	CM-SK71Vir	
EXPNO	5	
PROCNO	1	
Date_	20140717	
Time	11.49	
INSTRUM	Avance500	
PROBHD	5 mm QNP 1H/13	
PULPROG	zgfhigqn	
TD	131072	
SOLVENT	CDC13	
NS	16	
DS	4	
SWH	113636.367	Hz
FIDRES	0.866977	Hz
AQ	0.5767668	sec
RG	2300	
DW	4.400	use
DE	6.00	use
TE	298.1	K
D1	1.00000000	sec
D11	0.03000000	sec
D12	0.00002000	sec
TD0	1	
======== CHANNEL f1 =======		
NUC1	19F	
P1	18.60	use
PL1	-1.50	dB
PL1W	11.14113998	W
SFO1	470.5453180	MHz
======== CHANNEL f2 =======		
CPDPRG2	waltz16	
NUC2	1H	
PCPD2	80.00	use
PL2	-2.00	dB
PL12	14.85	dB
PL2W	14.33185768	W
PL12W	0.29600734	W
SFO2	500.1320005	MHz
SI	65536	
SF	470.5923770	MHz
WDW	EM	
SSB	0	
LB	0.30	Hz
GB	0	
PC	1.40	

^{13}C NMR spectrum of ProTide (5b)



CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

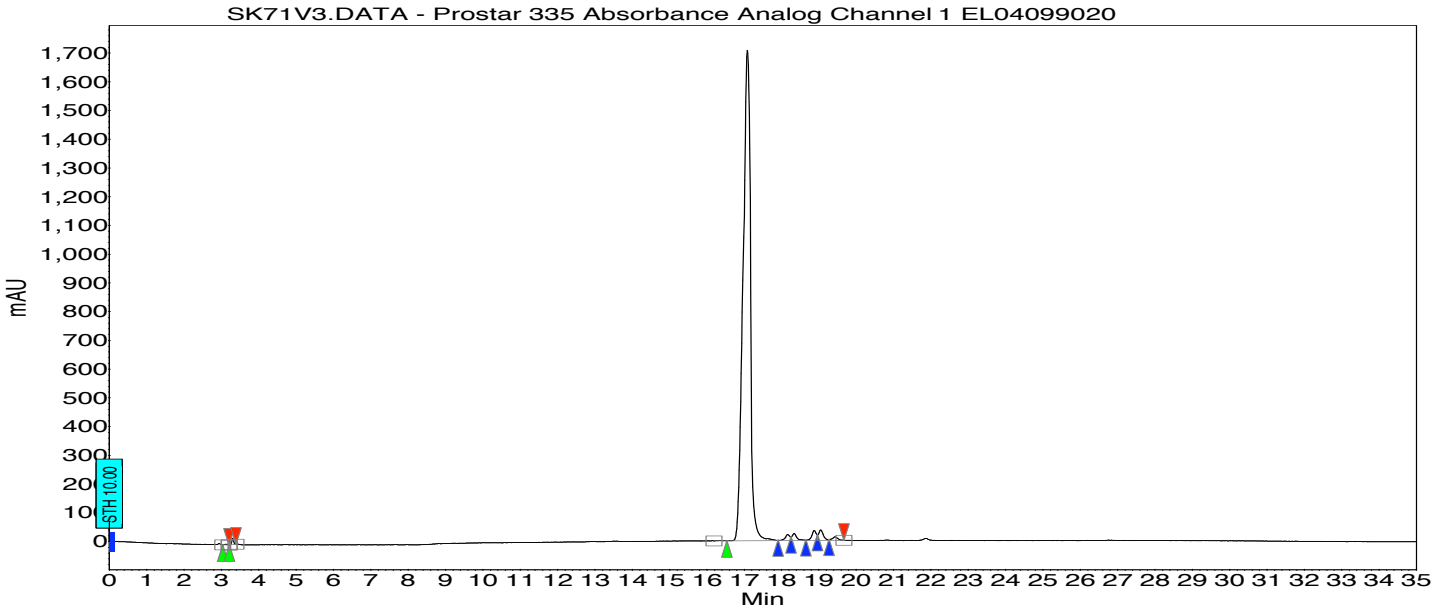
NAME	CM-SK71V
EXPNO	1
PROCNO	1
Date_	20140822
Time	21.46
INSTRUM	Avance500
PROBHD	5 mm QNP 1H/13
PULPROG	pendant
TD	65536
SOLVENT	CDCl3
NS	1152
DS	4
SWH	29761.904 Hz
FIDRES	0.454131 Hz
AQ	1.1010548 sec
RG	3250
DW	16.800 usec
DE	12.00 usec
TE	298.1 K
CNST2	145.0000000
D1	2.00000000 sec
D4	0.00172414 sec
D12	0.00002000 sec
D15	0.00431034 sec
D20	0.00345000 sec
TD0	9
======== CHANNEL f1 ========	
NUC1	13C
P1	7.20 usec
P2	14.40 usec
PL1	-2.00 dB
PL1W	101.27846527 W
SFO1	125.7703643 MHz
======== CHANNEL f2 ========	
CPDPRG2	waltz16
NUC2	1H
P3	11.50 usec
P4	23.00 usec
PCPD2	80.00 usec
PL2	-2.00 dB
PL12	14.85 dB
PL2W	14.33185768 W
PL12W	0.29600734 W
SFO2	500.1320005 MHz
SI	32768
SF	125.7577890 MHz
WDW	EM
SSB	0
LB	1.00 Hz
GB	0
PC	1.40

HPLC analysis spectrum of ProTide 5b

Chromatogram : SK71V3_channel1

System : HPLC_Analytical
Method : 5FU-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

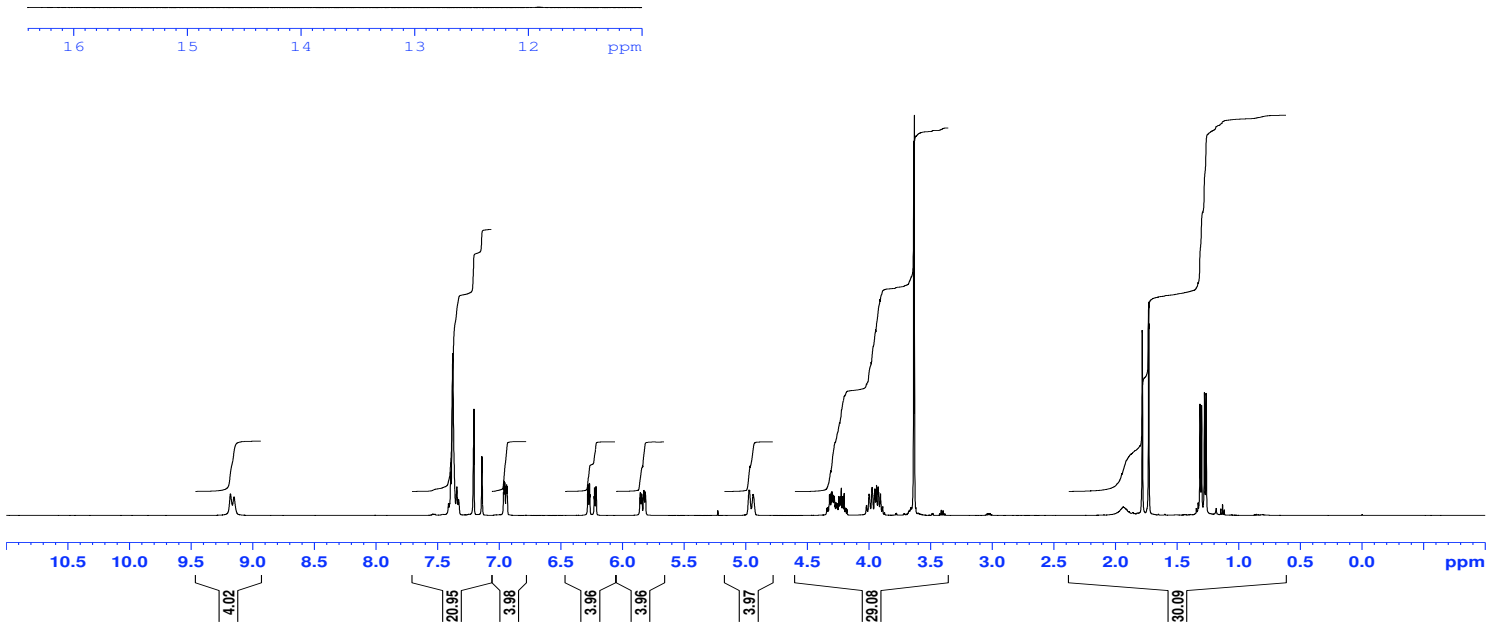
Acquired : 8/22/2014 11:30:01 AM
Processed : 8/22/2014 12:14:03 PM
Printed : 8/22/2014 12:16:05 PM
Method: 5FU-Blanka ACN 30+5min ok



Peak results :

Index	Name	Time [Min]	Quantity [% Area]	Height [mAU]	Area [mAU.Min]	Area % [%]
1	UNKNOWN	3.15	0.05	2.2	0.2	0.050
8	UNKNOWN	3.29	0.30	21.9	1.2	0.297
2	UNKNOWN	17.08	95.15	1706.9	394.5	95.146
3	UNKNOWN	18.16	0.74	20.9	3.1	0.741
4	UNKNOWN	18.33	0.85	24.1	3.5	0.852
5	UNKNOWN	18.88	1.14	34.3	4.7	1.145
6	UNKNOWN	19.05	1.29	36.3	5.4	1.292
7	UNKNOWN	19.43	0.48	12.1	2.0	0.478
Total			100.00	1858.8	414.7	100.000

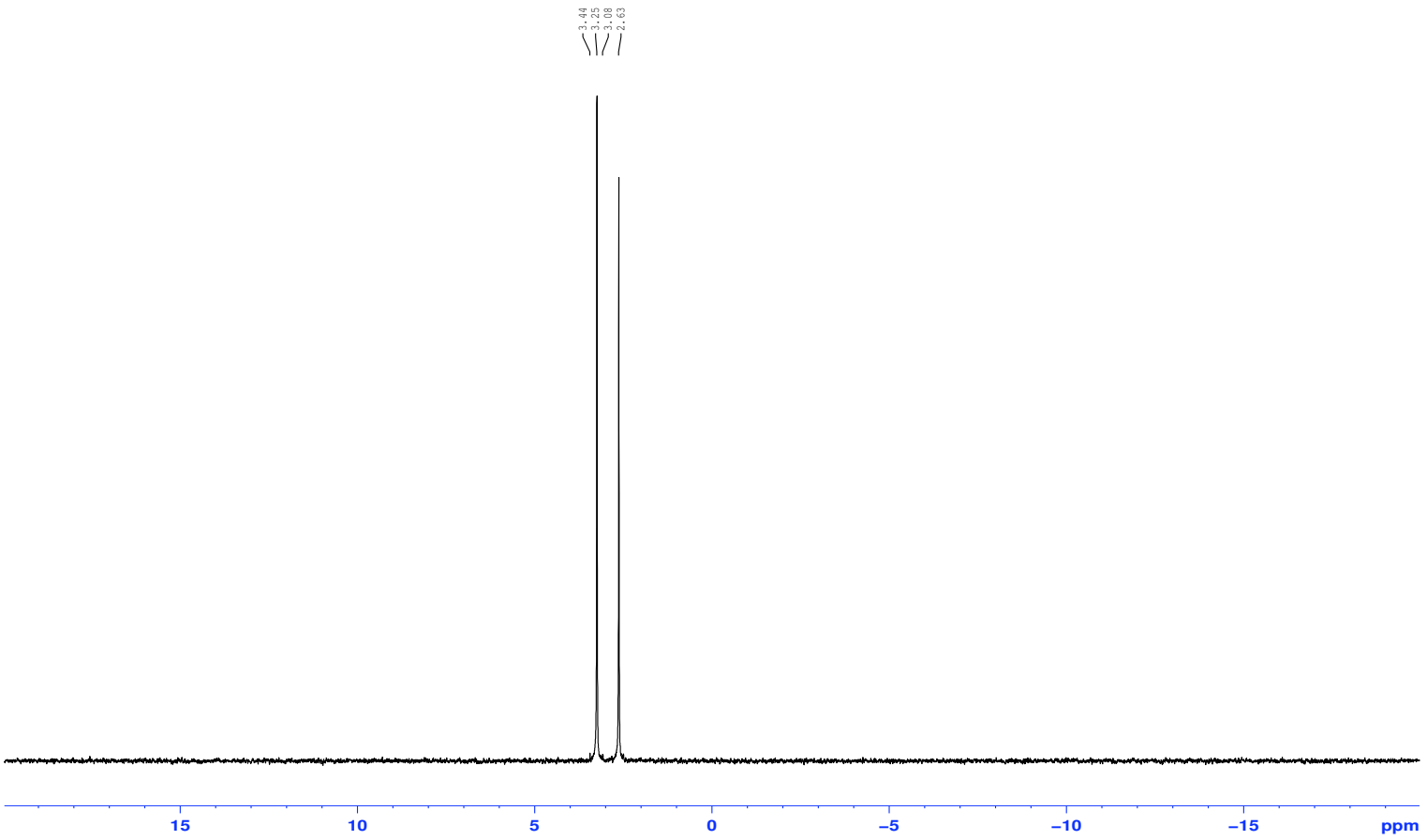
^{1}H NMR spectrum of ProTide (5c)



CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

NAME	CM-SK73Vir
EXPNO	10
PROCNO	1
Date_	20140810
Time	11.01
INSTRUM	Avance500
PROBHD	5 mm QNP 1H/13
PULPROG	zg30
TD	65536
SOLVENT	CDCl3
NS	16
DS	2
SWH	10330.578 Hz
FIDRES	0.157632 Hz
AQ	3.1719923 sec
RG	181
DW	48.400 use
DE	6.50 use
TE	298.1 K
D1	1.00000000 sec
TD0	1
======== CHANNEL f1 =======	
NUC1	1H
P1	11.50 use
PL1	-1.00 dB
PL1W	11.38419914 W
SFO1	500.1330885 MHz
SI	32768
SF	500.1300415 MHz
WDW	EM
SSB	0
LB	0.30 Hz
GB	0
PC	1.00

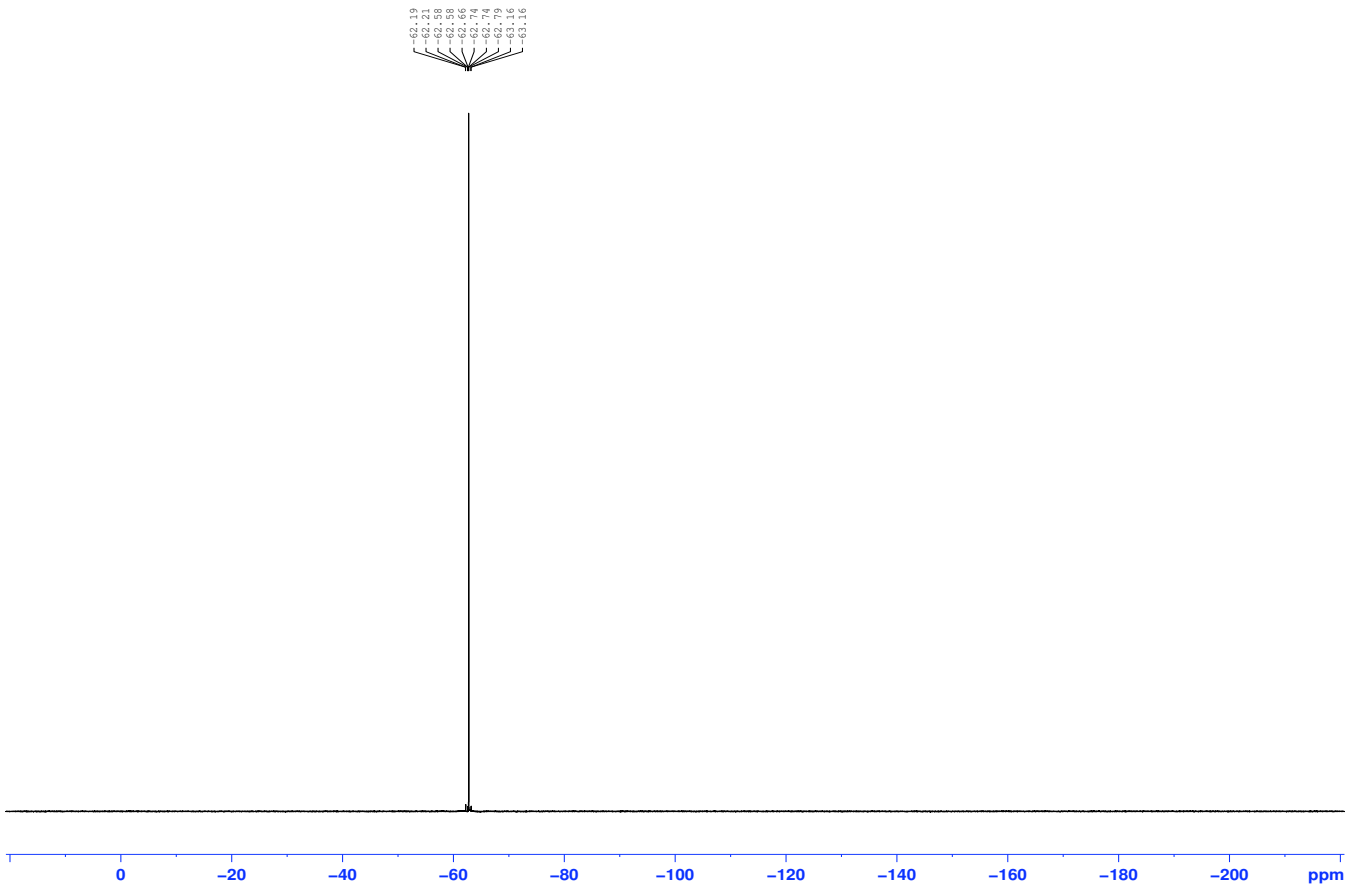
^{31}P NMR spectrum of ProTide (5c)



CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

NAME	CM-SK73Vir
EXPNO	5
PROCNO	1
Date_	20140721
Time	11.27
INSTRUM	Avance500
PROBHD	5 mm QNP 1H/13
PULPROG	zgpg30
TD	16384
SOLVENT	CDCl3
NS	16
DS	4
SWH	8081.896 Hz
FIDRES	0.493280 Hz
AQ	1.0136734 sec
RG	3250
DW	61.867 use
DE	12.00 use
TE	298.1 K
D1	1.65999997 sec
D11	0.03000000 sec
TD0	1
======== CHANNEL f1 =======	
NUC1	31P
P1	10.40 use
PL1	4.00 dB
PL1W	24.34000015 W
SFO1	202.4563350 MHz
======== CHANNEL f2 =======	
CPDPRG2	waltz16
NUC2	1H
PCPD2	80.00 use
PL2	-2.00 dB
PL12	14.85 dB
PL13	120.00 dB
PL2W	14.33185768 W
PL12W	0.29600734 W
PL13W	0.00000000 W
SFO2	500.1320005 MHz
SI	32768
SF	202.4563350 MHz
WDW	EM
SSB	0
LB	1.00 Hz
GB	0
PC	1.40

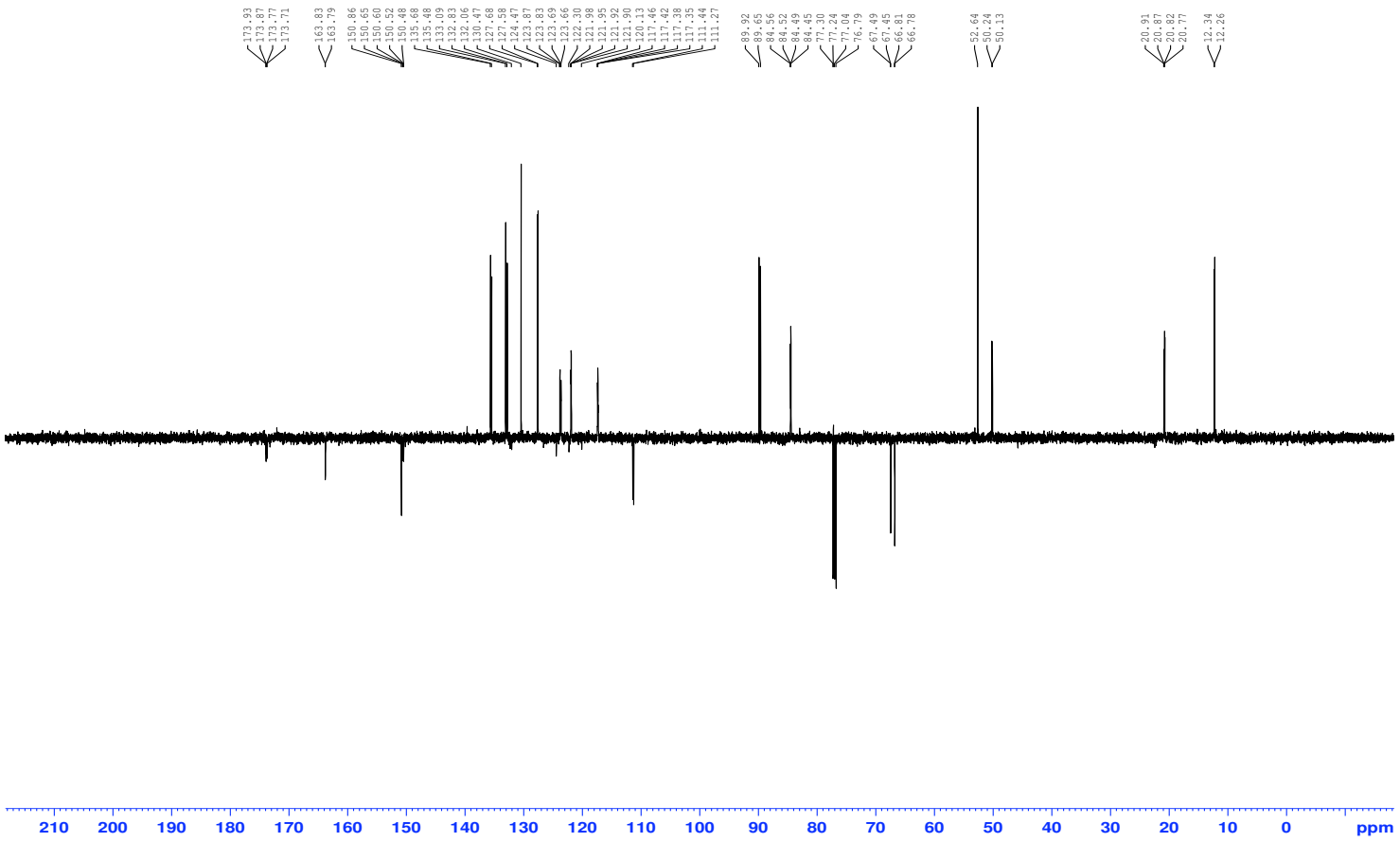
^{19}F NMR spectrum of ProTide (5c)



CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

NAME	CM-SK73Vir	
EXPNO	6	
PROCNO	1	
Date_	20140721	
Time	11.43	
INSTRUM	Avance500	
PROBHD	5 mm QNP 1H/13	
PULPROG	zgfhiggn	
TD	131072	
SOLVENT	CDCl3	
NS	16	
DS	4	
SWH	113636.367	Hz
FIDRES	0.866977	Hz
AQ	0.5767668	sec
RG	2580	
DW	4.400	use
DE	6.00	use
TE	298.1	K
D1	1.00000000	sec
D11	0.03000000	sec
D12	0.00002000	sec
TD0	1	
======== CHANNEL f1 =======		
NUC1	19F	
P1	18.60	use
PL1	-1.50	dB
PL1W	11.14113998	W
SFO1	470.5453180	MHz
======== CHANNEL f2 =======		
CPDPRG2	waltz16	
NUC2	1H	
PCPD2	80.00	use
PL2	-2.00	dB
PL12	14.85	dB
PL2W	14.33185768	W
PL12W	0.29600734	W
SFO2	500.1320005	MHz
SI	65536	
SF	470.5923770	MHz
WDW	EM	
SSB	0	
LB	0.30	Hz
GB	0	
PC	1.40	

^{13}C NMR spectrum of ProTide (5c)



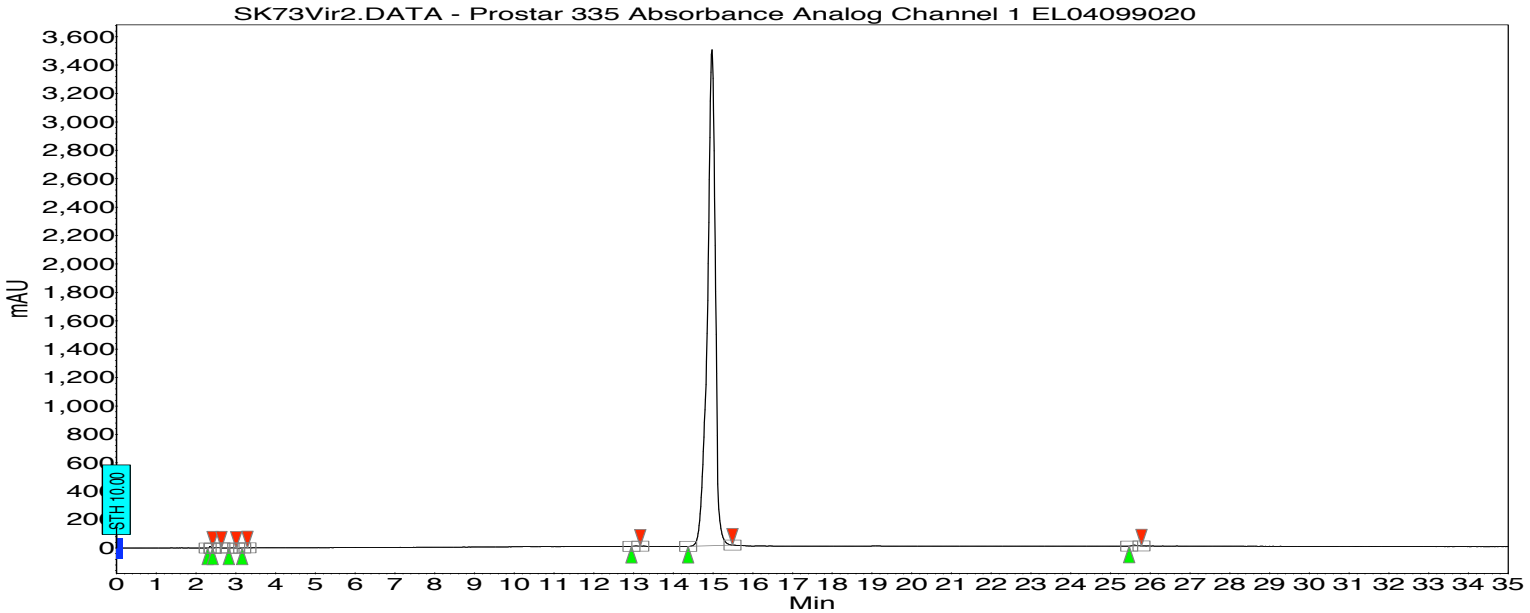
NAME	CM-SK73Vir
EXPNO	9
PROCNO	1
Date_	20140810
Time	10.53
INSTRUM	Avance500
PROBHD	5 mm QNP 1H/13
PULPROG	pendant
TD	65536
SOLVENT	CDC13
NS	256
DS	4
SWH	29761.904 Hz
FIDRES	0.454131 Hz
AQ	1.1010548 sec
RG	3250
DW	16.800 usec
DE	12.00 usec
TE	298.1 K
CNST2	145.0000000
D1	2.00000000 sec
D4	0.00172414 sec
D12	0.00002000 sec
D15	0.00431034 sec
D20	0.00345000 sec
TD0	2
======== CHANNEL f1 ========	
NUC1	13C
P1	7.20 usec
P2	14.40 usec
PL1	-2.00 dB
PL1W	101.27846527 W
SFO1	125.7703643 MHz
======== CHANNEL f2 ========	
CPDPRG2	waltz16
NUC2	1H
P3	11.50 usec
P4	23.00 usec
PCPD2	80.00 usec
PL2	-2.00 dB
PL12	14.85 dB
PL2W	14.33185768 W
PL12W	0.29600734 W
SFO2	500.1320005 MHz
SI	32768
SF	125.7577890 MHz
WDW	EM
SSB	0
LB	1.00 Hz
GB	0
PC	1.40

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

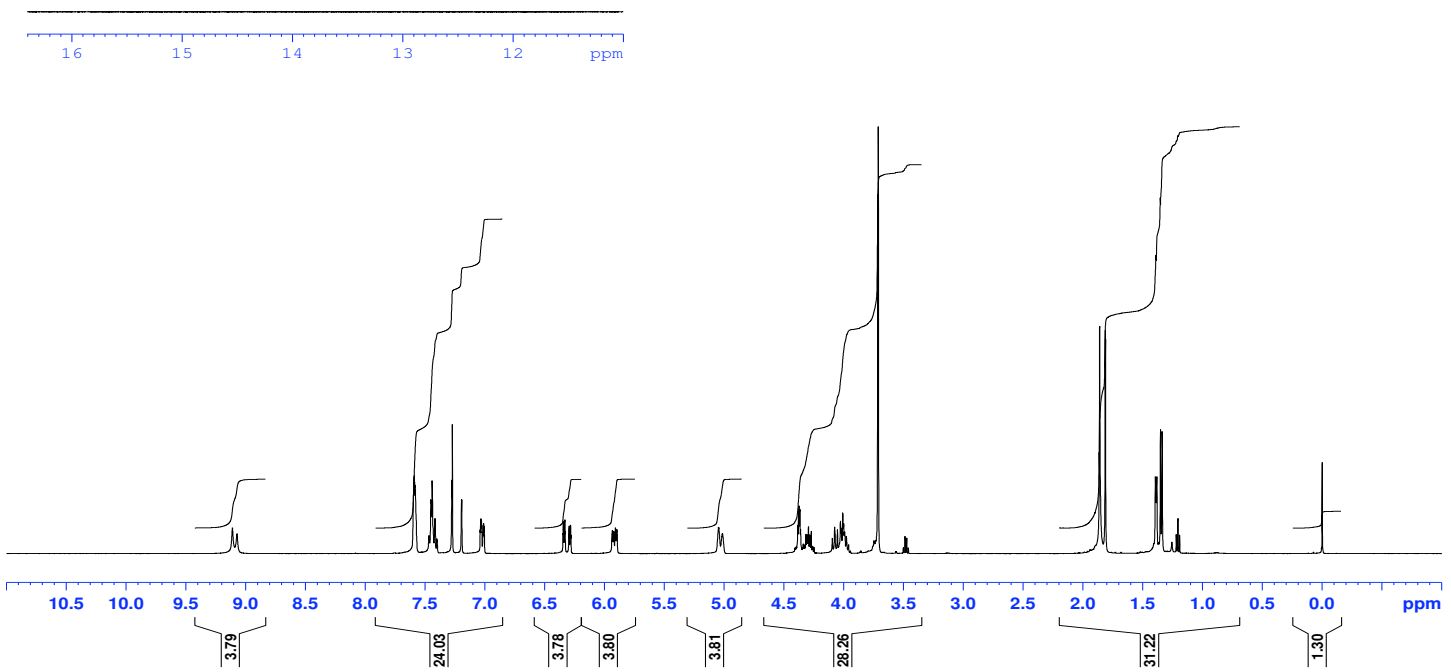
Acquired : 7/21/2014 11:11:44 AM
Processed : 7/21/2014 11:48:53 AM
Printed : 7/21/2014 11:56:32 AM
Method: 5FU-Blanka ACN 30+5min ok



Peak results :

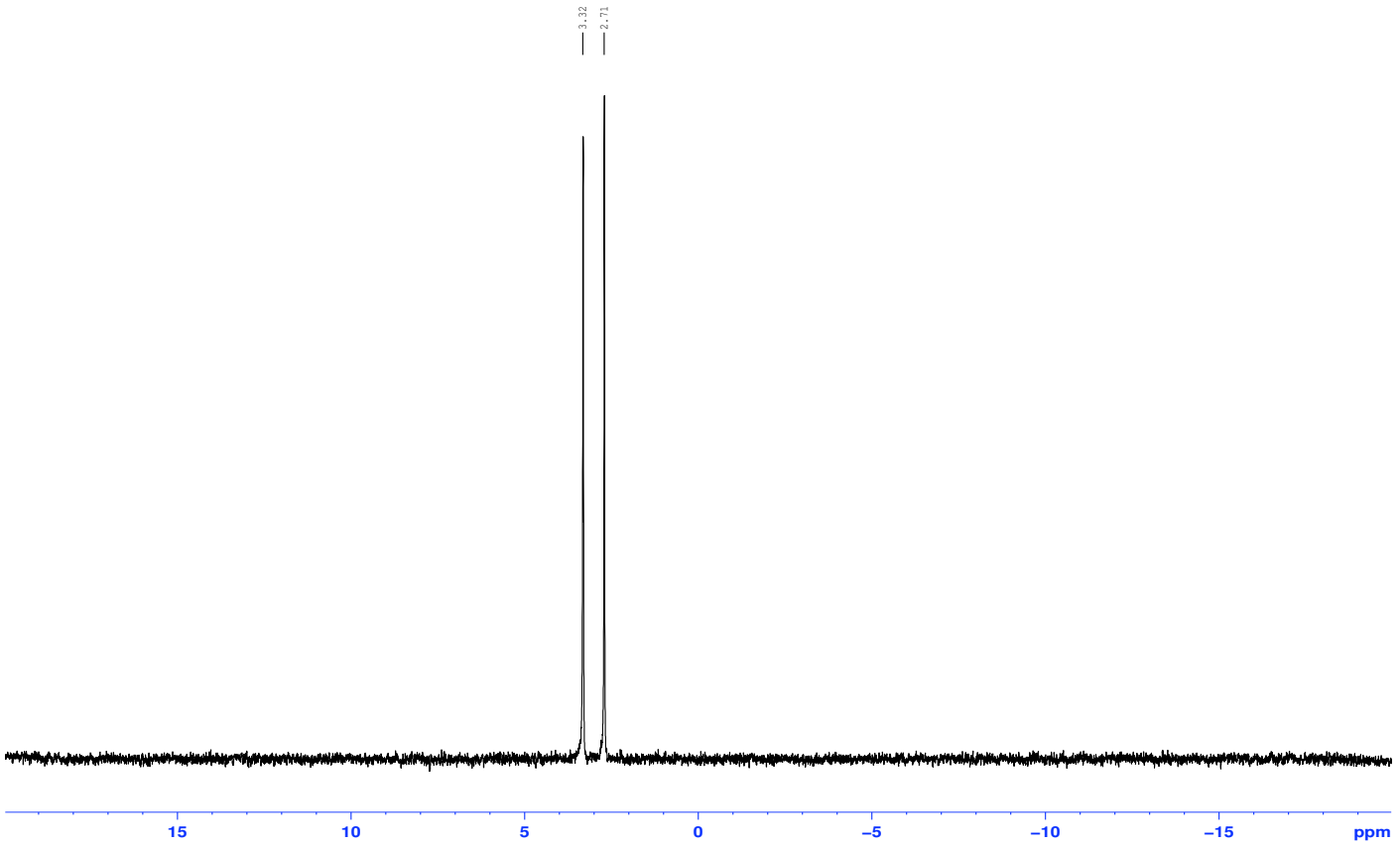
Index	Name	Time [Min]	Quantity [% Area]	Height [mAU]	Area [mAU.Min]	Area % [%]
1	UNKNOWN	2.36	0.07	11.6	0.5	0.066
2	UNKNOWN	2.55	0.05	3.0	0.4	0.050
3	UNKNOWN	2.91	0.02	1.9	0.1	0.018
4	UNKNOWN	3.20	0.01	0.9	0.1	0.009
5	UNKNOWN	13.07	0.03	2.0	0.2	0.028
6	UNKNOWN	14.97	99.72	3491.3	810.1	99.721
7	UNKNOWN	25.60	0.11	6.3	0.9	0.107
Total			100.00	3517.0	812.3	100.000

^{1}H NMR spectrum of ProTide (5d)




```
NAME            CM-SK57Vir
EXPNO                   16
PROCNO                   1
Date_             20140715
Time                 21.50
INSTRUM          Avance500
PROBHD    5 mm QNP 1H/13
PULPROG               zg30
TD                   65536
SOLVENT              CDC13
NS                      16
DS                       2
SWH              10330.578 Hz
FIDRES            0.157632 Hz
AQ               3.1719923 sec
RG                     228
DW                  48.400 use
DE                   6.50 use
TE                  298.1 K
D1               1.00000000 sec
TD0                      1

======== CHANNEL f1 =======
NUC1                    1H
P1                   11.50 use
PL1                  -1.00 dB
PL1W          11.38419914 W
SFO1         500.1330885 MHz
SI                   32768
SF           500.1300070 MHz
WDW                     EM
SSB                      0
LB                    0.30 Hz
GB                       0
PC                    1.00
```

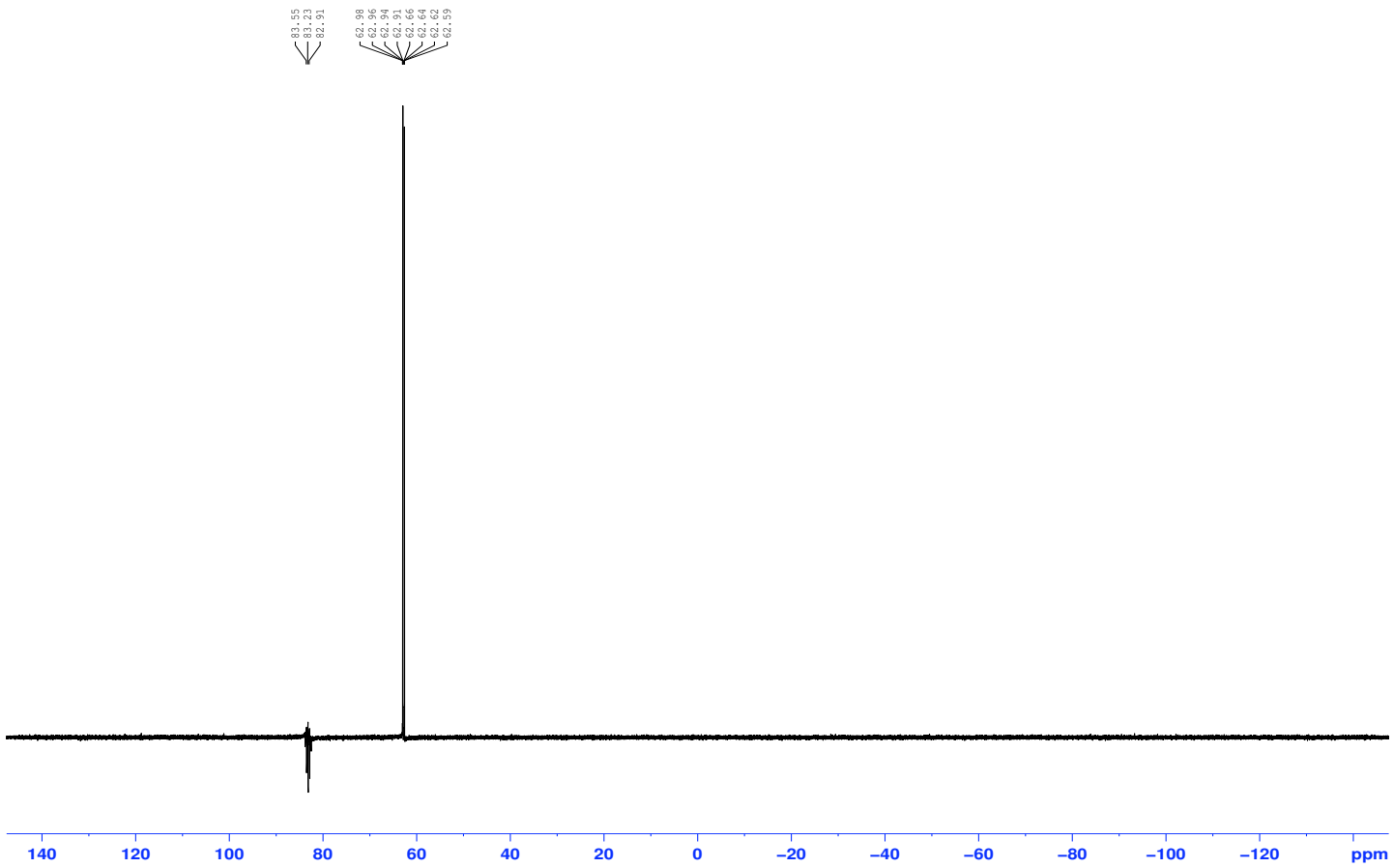
^{31}P NMR spectrum of ProTide (5d)



CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

NAME	CM-SK57Vir
EXPNO	12
PROCNO	1
Date_	20140703
Time	13.13
INSTRUM	Avance500
PROBHD	5 mm QNP 1H/13
PULPROG	zgpg30
TD	16384
SOLVENT	CDCl3
NS	16
DS	4
SWH	8081.896 Hz
FIDRES	0.493280 Hz
AQ	1.0136734 sec
RG	3250
DW	61.867 use
DE	12.00 use
TE	298.1 K
D1	1.65999997 sec
D11	0.03000000 sec
TD0	1
======== CHANNEL f1 =======	
NUC1	31P
P1	10.40 use
PL1	4.00 dB
PL1W	24.34000015 W
SFO1	202.4563350 MHz
======== CHANNEL f2 =======	
CPDPRG2	waltz16
NUC2	1H
PCPD2	80.00 use
PL2	-2.00 dB
PL12	14.85 dB
PL13	120.00 dB
PL2W	14.33185768 W
PL12W	0.29600734 W
PL13W	0.00000000 W
SFO2	500.1320005 MHz
SI	32768
SF	202.4563350 MHz
WDW	EM
SSB	0
LB	1.00 Hz
GB	0
PC	1.40

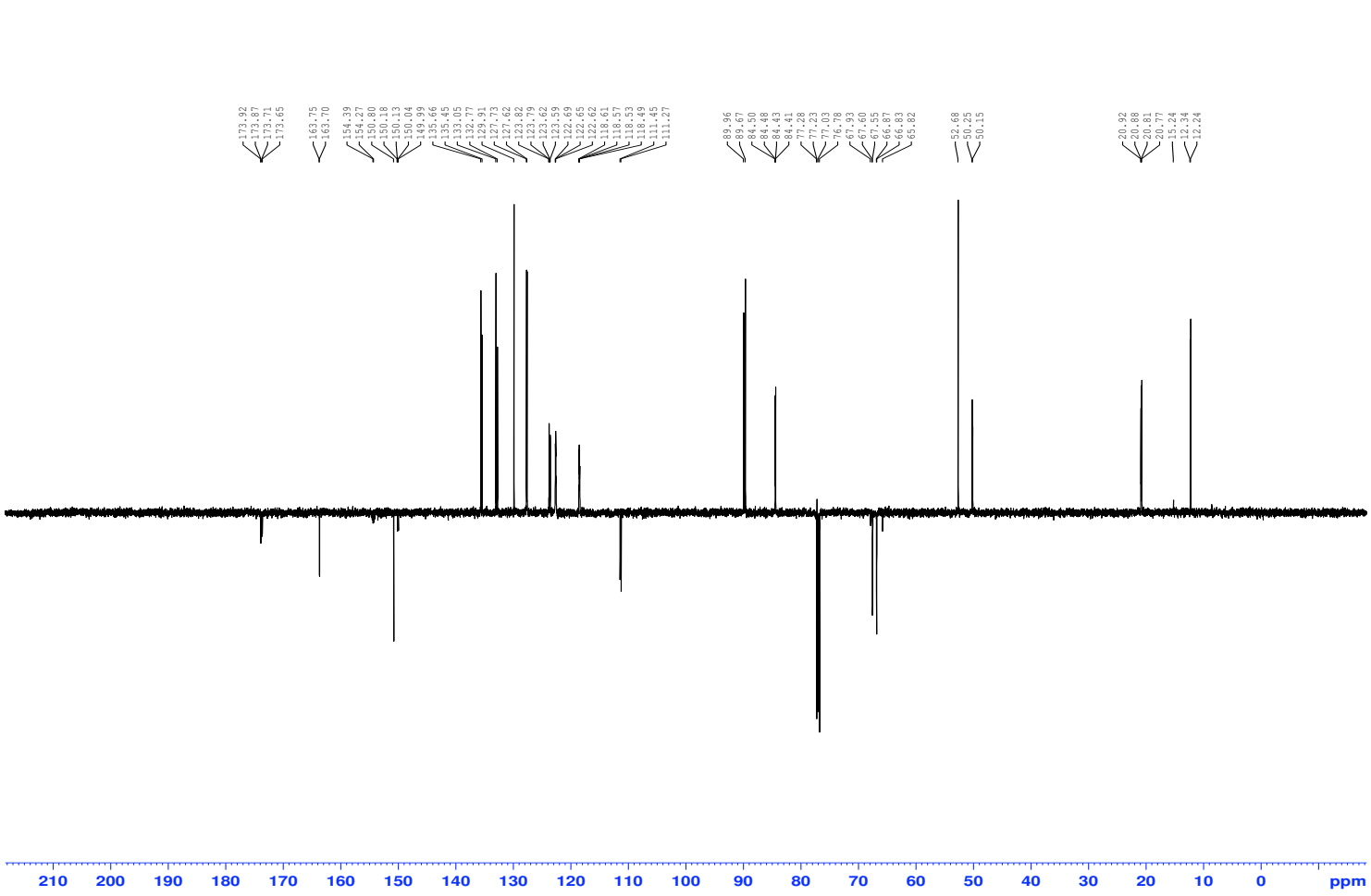
^{19}F NMR spectrum of ProTide (5d)



CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

NAME	CM-SK57Vir
EXPNO	13
PROCNO	1
Date_	20140703
Time	13.14
INSTRUM	Avance500
PROBHD	5 mm QNP 1H/13
PULPROG	zgfhigqn
TD	131072
SOLVENT	CDCl3
NS	16
DS	4
SWH	138888.891 Hz
FIDRES	1.059638 Hz
AQ	0.4719092 sec
RG	2300
DW	3.600 use
DE	6.00 use
TE	298.1 K
D1	1.00000000 sec
D11	0.03000000 sec
D12	0.00002000 sec
TD0	1
======== CHANNEL f1 =======	
NUC1	19F
P1	18.60 use
PL1	−1.50 dB
PL1W	11.14113998 W
SFO1	470.5923770 MHz
======== CHANNEL f2 =======	
CPDPRG2	waltz16
NUC2	1H
PCPD2	80.00 use
PL2	−2.00 dB
PL12	14.85 dB
PL2W	14.33185768 W
PL12W	0.29600734 W
SFO2	500.1320005 MHz
SI	65536
SF	470.5923770 MHz
WDW	EM
SSB	0
LB	0.30 Hz
GB	0
PC	1.40

^{13}C NMR spectrum of ProTide (5d)



CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

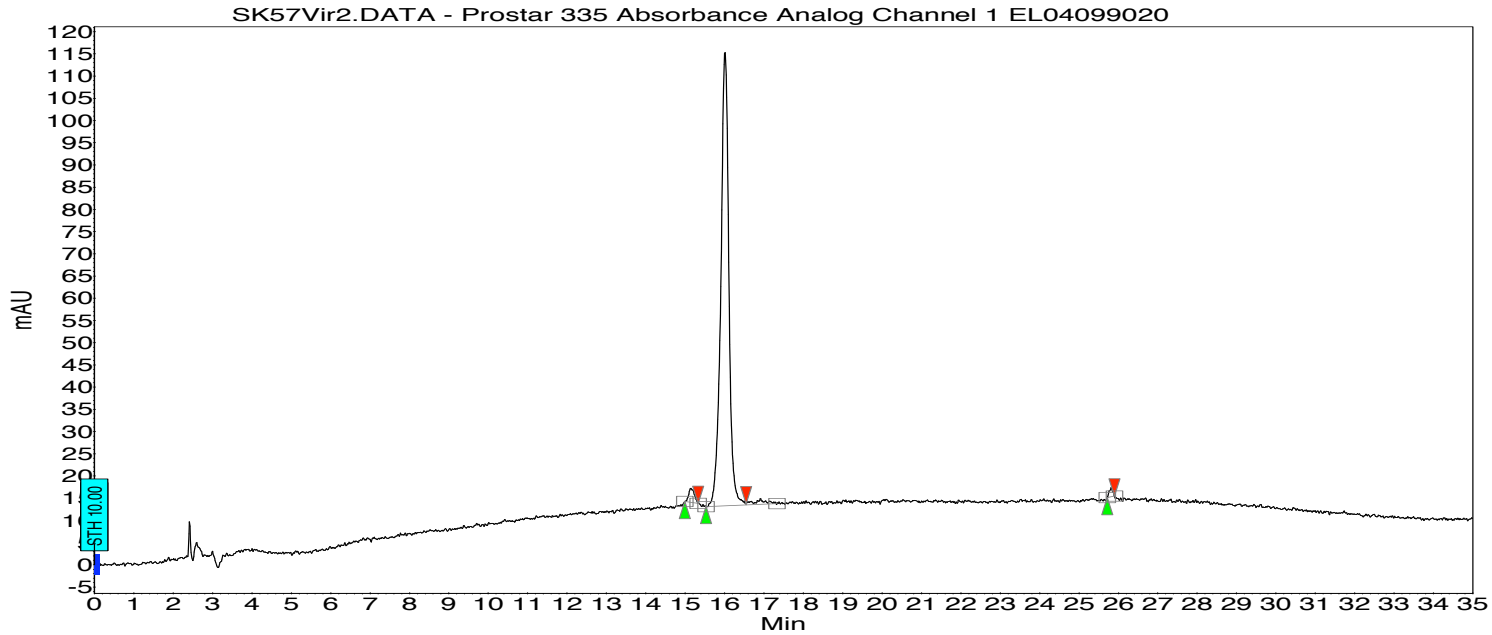
NAME	CM-SK57Vir
EXPNO	15
PROCNO	1
Date_	20140715
Time	21.02
INSTRUM	Avance500
PROBHD	5 mm QNP 1H/13
PULPROG	pendant
TD	65536
SOLVENT	CDCl3
NS	1024
DS	4
SWH	29761.904 Hz
FIDRES	0.454131 Hz
AQ	1.1010548 sec
RG	3250
DW	16.800 usec
DE	12.00 usec
TE	298.1 K
CNST2	145.0000000
D1	2.00000000 sec
D4	0.00172414 sec
D12	0.00002000 sec
D15	0.00431034 sec
D20	0.00345000 sec
TD0	8
======== CHANNEL f1 ========	
NUC1	13C
P1	7.20 usec
P2	14.40 usec
PL1	-2.00 dB
PL1W	101.27846527 W
SFO1	125.7703643 MHz
======== CHANNEL f2 ========	
CPDPRG2	waltz16
NUC2	1H
P3	11.50 usec
P4	23.00 usec
PCPD2	80.00 usec
PL2	-2.00 dB
PL12	14.85 dB
PL2W	14.33185768 W
PL12W	0.29600734 W
SFO2	500.1320005 MHz
SI	32768
SF	125.7577890 MHz
WDW	EM
SSB	0
LB	1.00 Hz
GB	0
PC	1.40

HPLC analysis spectrum of ProTide 5d

Chromatogram : SK57Vir2_channel1

System : HPLC_Analytical
Method : 5FU-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: 5FU-Blanka ACN 30+5min ok



Peak results :

Index	Name	Time [Min]	Quantity [% Area]	Height [mAU]	Area [mAU.Min]	Area % [%]
3	UNKNOWN	15.16	2.06	2.9	0.5	2.055
1	UNKNOWN	16.01	97.15	102.0	24.2	97.145
2	UNKNOWN	25.81	0.80	2.1	0.2	0.800
Total			100.00	107.1	24.9	100.000