Accepted Manuscript

A Linear Synthesis of Gemcitabine

Kylie Brown, Alex Weymouth-Wilson, Bruno Linclau

PII: S0008-6215(15)00012-9

DOI: 10.1016/j.carres.2015.01.001

Reference: CAR 6917

To appear in: *Carbohydrate Research*

Received Date: 21 October 2014

Revised Date: 3 January 2015

Accepted Date: 6 January 2015

Please cite this article as: Brown K, Weymouth-Wilson A, Linclau B, A Linear Synthesis of Gemcitabine, *Carbohydrate Research* (2015), doi: 10.1016/j.carres.2015.01.001.

A Linear Synthesis of Gemcitabine

Kylie Brown,[1,2] Alex Weymouth-Wilson,[2] Bruno Linclau*,[1]

[1] Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ.

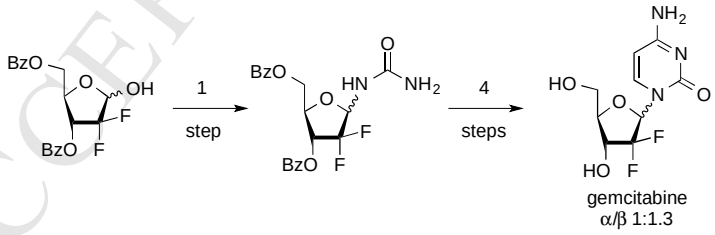
[2] Dextra Laboratories Ltd, The Science and Technology Centre, Earley Gate, Whiteknights Road, Reading RG6 6BZ, UK

bruno.linclau@soton.ac.uk

Abstract

Gemcitabine, 2'-deoxy-2',2'-difluorocytidine, is currently prescribed against a number of cancers. Here we report a linear synthesis of gemcitabine with a high-yielding direct conversion of 3,5-di-*O*-benzoyl-2-deoxy-2,2-difluororibose into the corresponding glycosyl urea as the key step, followed by conventional conversion to the cytosine base via the uracil derivative. The process proceeded with modest anomeric selectivity.

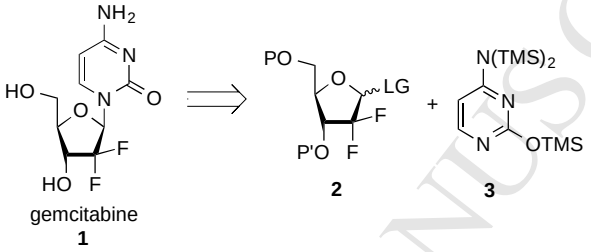
Graphical Abstract



Keywords: gemcitabine; glycosyl urea; fluorinated nucleoside; linear nucleoside introduction

1. Introduction

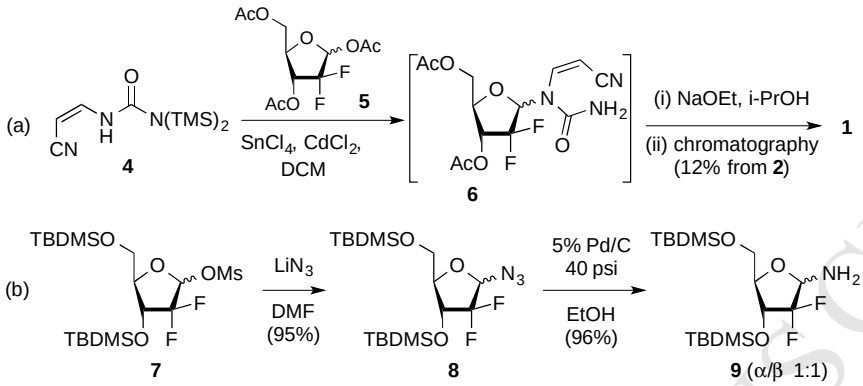
Gemcitabine **1** (Scheme 1), marketed by its discoverer Lilly as the HCl salt under the trade name of Gemzar, is a widely prescribed anticancer drug against pancreatic, ovarian and breast cancer. Before its recent patent expiry, it was a $1 bn dollar per annum drug, and while sales have now dropped below that figure, its market continues to grow. It is a fluorinated nucleoside[1-6] prodrug, undergoing intracellular phosphorylation to its active diphosphate and triphosphate form, which inhibits DNA synthesis leading to apoptosis.[7-11]



Scheme 1: Convergent nucleobase introduction in the synthesis of gemcitabine **1**.

The synthesis of gemcitabine[12] has been subject to tremendous and –spurred by the Lilly patent expiry– still continuing efforts.[13] In general, a convergent synthesis has been applied, in which a suitably protected and activated 2-deoxy-2,2-difluororibofuranose derivative **2** (Scheme 1) is combined with an activated cytosine base, eg **3**. Much optimisation has been devoted to an efficient 2-deoxy-2,2-difluororibofuranose synthesis, a feat which has now largely been achieved through use of specific alcohol protecting groups allowing efficient and high-yielding diastereomer separation by crystallisation(s). The nucleobase introduction reaction has also been subject to intensive research. This process is made difficult by the difluorination at the sugar 2-position, and typically leads to an anomeric mixture. Again, selective crystallisation procedures have been developed to obtain anomerically pure gemcitabine.[13]

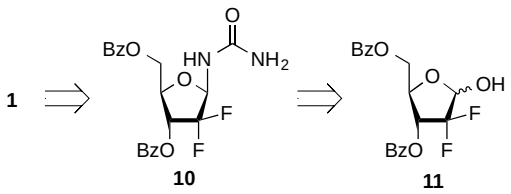
While alternative linear nucleoside syntheses are common in the carbanucleoside field,[14-16] this strategy has to our knowledge been applied only once for the gemcitabine synthesis (Scheme 2a).[17]



Scheme 2. Linear gemcitabine synthesis.

Here the sugar donor **5** was reacted with the *N*-2-cyanovinyl amide **4** to give the intermediate **6**, after which the cytosine synthesis was completed by base-induced cyclisation. An overall yield of 12% was reported. A linear synthesis is also possible – but not reported – from the corresponding amino-2-deoxy-2,2-difluoroglycoside **9**, which has been synthesised in two steps by a Lilly group.[18] (Scheme 2b). However, while **9** was obtained in high yield, no anomeric selectivity was obtained, and the anomeric ratio of **9** was independent of the anomeric ratio of the azide precursor **8**, suggesting an equilibration process.

Herein we describe our efforts leading to an alternative linear gemcitabine synthesis. In order to avoid the abovementioned anomerisation at the aminoglycoside stage, a nucleobase construction strategy starting from the corresponding glycosyl urea **10** was envisioned (Scheme 3). It was hoped that the possible crystallinity of this urea derivative would offer anomeric purification prospects.



Scheme 3.

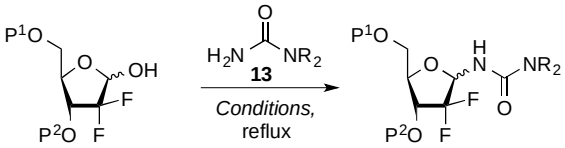
In addition, it was planned to introduce the urea by a direct condensation reaction with the known 3,5-di-*O*-benzoyl-2-deoxy-2,2-difluororibose **11**, thus bypassing the need for anomeric activation. This protected difluororibose derivative was first described by a Lilly group,[19] but obtained by us by reduction of the commercially available dibenzoylated difluororibonolactone.

2. Results and discussion

2.1 Urea condensation

The work commenced by investigating the direct urea introduction. The acid-catalysed condensation of aldehydes with urea has been reported in high yields,[20,21] and there were few reports available of this process with unprotected (and unfluorinated) carbohydrates.[22,23] In recent years, the reaction of unprotected and non-activated carbohydrates with ureas has received renewed attention, and has typically been achieved with an acid-catalyst.[24] Hence, it was first investigated whether the acid-catalysed condensation of a difluorinated sugar such as furanose **11** was possible with ureas. Initially, the 5-*O*-TBDPS lactol **12** was used as substrate (Table 1), and to avoid complication by the urea reacting at both ends, *N,N*-dimethyl urea **12** was used.

Table 1. Optimisation of the urea condensation reaction.



11 P^1=P^2= Bz
12 P^1=TBDPS P^2=Bn

a R = H
b R = Me

14 P^1=P^2= Bz
15 P^1=TBDPS P^2=Bn

Entry	Lactol	**13** (equiv)	Solvent	PTSA (equiv)	Drying agent	Time (h)	Work-up	Et_3N Column?	Yield[a] (%)
1	**12**	**b** (2)	toluene	0.05	MS	18	[b]	N	17
2	**12**	**b** (2)	toluene	0.05	-	5.5	[b]	N	22
3	**12**	**b** (1)	toluene	0.05	-	5	[b]	N	29
4	**12**	**b** (1)	toluene	0.05	MS	5.5	[b]	N	38
5	**11**	**b** (3.3)	toluene	1.1	MS	20	[c]	Y	65
6	**11**	**a** (3)	dioxane	1	MS	18	Aqueous	Y	40
7	**11**	**a** (5)	dioxane	1	MS	44	Aqueous	Y	59
8	**11**	**a** (3)	dioxane	1	Na_2SO_4	18	Aqueous	Y	61
9	**11**	**a** (3)	dioxane	1	Na_2SO_4	36	Column	Y	62
10	**11**	**a** (5)	dioxane	1	Na_2SO_4	72	Aqueous	Y	64
11	**11**	**a** (3)	dioxane	1	Na_2SO_4	66	[c]	Y	65
12	**11**	**a** (3)	dioxane	1	Na_2SO_4	36	[d]	N	79
13	**11**	**a** (3)	dioxane	1	-	36	[d]	N	76
14	**11**	**a** (5)	dioxane	1	Na_2SO_4	36	[d]	N	79
15[e]	**11**	**a** (3)	dioxane	1	Na_2SO_4	36	[d]	N	88

[a] Isolated yield. [b] Remove reaction solvent by evaporation *in vacuo*. [c] Filtration over silica. [d] Filtration over celite. [e] 2.3 mmol scale.

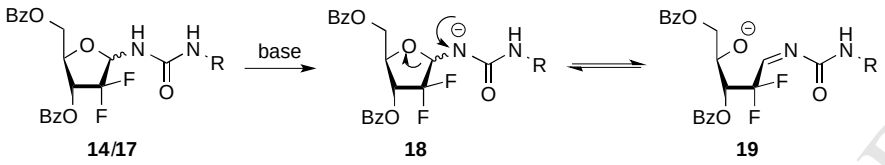
Reaction of **12** with **13b** using a variety of acid catalysts (TMSOTf, Ti(O*i*-Pr)$_4$, 4 N HCl, AcOH, TFA, *p*-TsOH) in a range of solvents (THF, DCM, toluene, trifluoroethanol, hexafluoroisopropanol) all gave no or only trace amounts of product **14** (not shown), except the combination of *p*-TsOH and toluene at reflux (entry 1). While, when using 2 equiv of urea the presence/absence of molecular sieves had no significant influence (entries 1,2), when 1 equiv of urea was used, the use of molecular sieves did lead to a higher yield (entries 3,4). Due to the possible lability of

the TBDPS ether in acidic medium, further optimisation experiments were carried out with the 3,5-di-*O*-benzoyl-2-deoxy-2,2-difluororibose **11**, obtained by reduction of the commercially available lactone. Indeed, reaction with **13b** now gave a much improved 65% yield (entry 5), although a further increase in amount of urea was used. At this stage, it was decided to investigate the non-methylated urea **13a**, which turned out to be unproblematic with regard to a sequential condensation process. The best solvent proved to be 1,4-dioxane (entry 6), giving a yield of 40%, with 39% recovered starting material, when using 1 equiv of *p*-TsOH. An increase in the relative amount of urea increased the yield to 59% (entry 7). As good conversion was seen by NMR analysis of crude material, losses during the workup due to difficult filtration of the molecular sieves were thought to be responsible for the reduced yield. However, a change of drying agent to Na_2SO_4 (entries 8–11), whilst facilitating the work-up, did not have a significant effect upon the yield, regardless of urea stoichiometry. Changing the workup to filtration over celite instead of silica gel did improve the yield (entry 12), though, surprisingly, the reaction without Na_2SO_4 worked equally well (entry 13). Increasing the number of urea equivalents to 5 had no further beneficial effect (entry 14). Finally, reaction on 2.3 mmol scale under the best conditions led to an excellent 88% yield (entry 15) of the glycosyl ureas in a 1:1.8 anomeric ratio. Unfortunately, assignment of the anomers could not be achieved.

2.2 Formation of the cytosine nucleobase

Next, acylation of the urea group with acyl chloride **16**, synthesised by the method of Tietze,[25] to give the nucleobase cyclisation precursor was investigated. Interestingly, this glycosyl urea acylation does not have a precedent for *O*-nucleosides, but has been demonstrated with carbanucleosides using pyridine with DMAP.[26,27] However, reaction of **14a** with **16** under those conditions (Table 2, entry 1), only afforded 30% of the desired product **17**. Unexpectedly, significant anomerisation was observed,

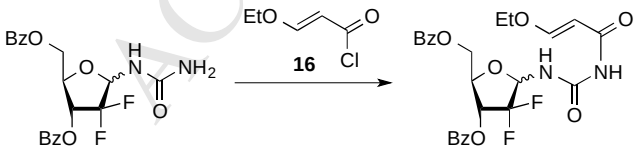
which led to a decrease in anomeric ratio (1:2.4 to 1:1). This anomerisation may be due to urea deprotonation by pyridine (Scheme 4), facilitated by the electron withdrawing fluorines at the 2-position, to give the conjugate base **18**. Anomerisation can then be envisaged by ring opening to **19** followed by non-selective ring closure.



Scheme 4. Possible anomerisation mechanism

Thus, the basic pyridine solvent was replaced by DCM, but even with 1 equiv of DMAP, none of the desired product was observed after 3.5 d (entry 2). Pleasingly, refluxing in acetonitrile with 2 equiv of **16** (entry 3) did yield the product in 45% yield, with a further improvement to 71% seen when the amount of **16** was increased to 4 equiv (entry 4). These conditions also resulted in less anomerisation (1:2.4 to 1:1.5). The same conditions with 1,4-dioxane as solvent gave a slightly lower yield, but more anomerisation was observed (entry 5). The least anomerisation was seen when $ZnCl_2$ was employed, with a 1:1.9 anomeric ratio seen in the isolated product (entry 6), but a lower yield was obtained.

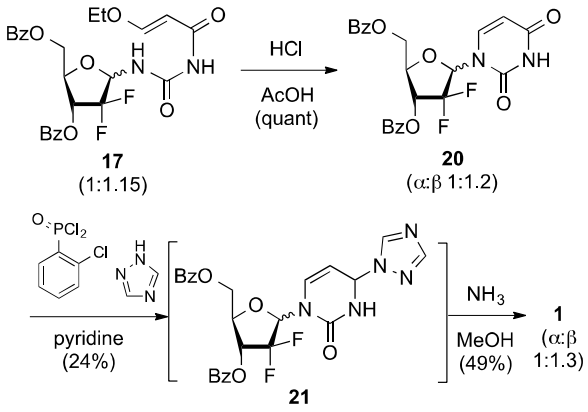
Table 2. Optimisation of the urea acylation reaction.



Entry	Equiv **16**	Reagents	Solvent	Temp (℃)	Time	Yield[a]	Anomeric Ratio[b]
1	1.5	0.25 equiv DMAP	Pyridine	Reflux	40 h	30%	1:1
2	1	1 equiv DMAP	DCM	RT	3.5 d	0%	-

3	2	None	MeCN	Reflux	18 h	45%	1:1.5
4	4	None	MeCN	Reflux	18 h	71%	1:1.5
5	4	None	Dioxane	Reflux	18 h	61%	1:1.15
6	3	0.25 equiv $ZnCl_2$	$CHCl_3$	Reflux	24 h	45%	1:1.9

[a] isolated yield. [b] Determined by NMR analysis on the crude reaction mixture.



Scheme 5. Conversion to gemcitabine **1** via the corresponding uridine **20**.

With the acyl carbamate **17** in hand, cyclisation to the protected difluorouridine **20** was studied (Scheme 5) using methodology originally developed by Shaw.[28] Among the various acidic protocols screened, HCl/AcOH emerged as the most efficient, leading to quantitative conversion to **20** in a 1:1.2 α/β mixture. Next, conversion of uridine to cytidine was effected using described methodology[29,30] via the triazole intermediate **21** by treatment with 2-chlorophenyl phosphorodichloridate and 1,2,4-triazole in pyridine, followed by reaction with ammonia. The ammonia treatment also caused concomitant benzoate deprotection, which then furnished gemcitabine **1** as a 1:1.3 mixture of α:β anomers.

3. Conclusion

The linear synthesis of gemcitabine was achieved in 5 steps from a 2-deoxy-2,2-difluororibose substrate with moderate anomeric selectivity. The demonstration of a high-yielding direct glycosyl urea formation through reaction of a reducing 2-deoxy-2,2-difluorinated sugar derivative with urea will be of interest for other applications.

4. Experimental section

4.1 Synthesis of the urea derivative 14a. Lactol **11** (870 mg, 2.30 mmol), urea (3 equiv, 6.90 mmol, 414 mg), PTSA (1 equiv, 2.30 mmol, 437 mg) and Na_2SO_4 (2 equiv, 4.60 mmol, 653 mg) were stirred in 1,4-dioxane (6.5 mL) at reflux for 36 h. The reaction mixture was cooled to RT, diluted with DCM and filtered through celite. The solvents were reduced *in vacuo* to yield a crude residue, which was purified by column chromatography on silica gel (10:90→40:60 acetone/petrol) to yield the desired glycosylurea **14a** as a white foam, as a 1:1.8 anomeric mixture (853 mg, 88%). **IR** (film) 3366 (br), 2925 (w), 1723 (m), 1678 (m), 1266 (s) cm^{-1}. **LRMS** (ESI^+) *m/z* 443.2 $(M+Na)^+$ (100). **HRMS** (ESI^+) for $C_{20}H_{18}F_2N_2O_6$ $(M+Na)^+$ calcd 443.1025, found 443.1018. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.04 (4 H, m, 4×CH_{Ar}), 7.61–7.33 (6 H, m, 6×CH_{Ar}), 6.69 ($^2/_3$ H, d, *J* = 10.0 Hz, NH_{Major}), 6.54 ($^1/_3$ H, d, *J* = 8.8 Hz, NH_{Minor}), 5.94 ($^1/_3$ H, q, *J* = 8.2 Hz, C<u>H</u>NH_{Minor}), 5.78 ($^2/_3$ H, ddd, *J* = 14.4, 10.1, 5.4 Hz, C<u>H</u>NH_{Major}), 5.67 ($^1/_3$ H, ddd, *J* = 11.5, 8.1, 6.6 Hz, $CHOBz_{Minor}$), 5.46 ($^2/_3$ H, m, $CHOBz_{Major}$), 5.34 ($^2/_3$ H, br.s, NH_{2Minor}), 5.32 ($^4/_3$ H, br.s, NH_{2Major}), 4.65–4.51 ($^7/_3$ H, m, C<u>H</u>CH_{2Minor}, CH_2), 4.36 ($^2/_3$ H, q, *J* = 4.3 Hz, C<u>H</u>CH_{2Major}) ppm. **^{13}C NMR + DEPT** (100 MHz, $CDCl_3$) δ 166.3 (C=O), 165.0 ($C{=}O_{Major}$), 164.9 ($C{=}O_{Minor}$), 157.8 ($C{=}O_{Minor}$), 157.7 ($C{=}O_{Major}$), 134.0 (CH_{Ar}), 133.9 (CH_{Ar}), 133.3 (CH_{Ar}), 130.0 (CH_{Ar}), 129.7 (CH_{Ar}), 129.2 (C_{Ar}), 128.6 (CH_{Ar}), 128.5 (CH_{Ar}), 128.43 (CH_{Ar}), 128.38 (CH_{Ar}), 128.1 (C_{Ar}), 121.4 (dd, *J* = 262.4, 257.6 Hz, CF_{2Minor}), 120.4 (dd, *J* = 266.3, 253.7 Hz, CF_{2Major}), 81.9 (dd, *J* = 37.9, 19.4 Hz, CHN_{Minor}), 81.2 (dd, *J* = 34.0, 20.4 Hz, CHN_{Major}), 77.2 (d, *J* = 5.8 Hz, <u>C</u>HCH_{2Minor}), 76.0 (<u>C</u>HCH_{2Major}), 72.3 (dd, *J* = 31.1, 17.5 Hz, $CHOBz_{Minor}$), 72.0 (dd, *J* = 36.0, 16.5 Hz, $CHOBz_{Major}$), 63.4 (CH_{2Minor}), 63.3 (CH_{2Major}) ppm (some CH_{Ar} overlap).

Major Anomer: **^{19}F NMR** (282 MHz, $CDCl_3$) δ –117.9 (1 F, dt, *J* = 246.6, 14.7 Hz, C<u>F</u>F), –122.9 (1 F, dd, *J* = 247.4, 5.2 Hz, CF<u>F</u>) ppm.

Minor Anomer: **^{19}F NMR** (282 MHz, $CDCl_3$) δ –113.2 (1 F, d, *J* = 243.1 Hz, C<u>F</u>F), –125.1 (1 F, dt, *J* = 244.0, 7.8 Hz, CF<u>F</u>) ppm.

4.2 Synthesis of the *N*,*N*-dimethyl urea derivative 14b. To lactol **11** (35 mg, 0.09 mmol) in toluene (1 mL) was added 3 Å molecular sieves (53 mg), *N*,*N*-dimethylurea (3 equiv, 0.30 mmol, 27 mg) and PTSA (1 equiv, 0.09 mmol, 19 mg). The reaction mixture was stirred at reflux for 20 h then cooled to RT and diluted with DCM before filtration through a plug of silica (pre-treated with Et_3N), $NaHCO_3$ and $NaSO_4$. The solvents were reduced *in vacuo* to yield a crude residue which was purified by column chromatography on silica gel (15:85→25:75 acetone/hexane + 0.5% Et_3N) to yield the desired glycosylurea **14b** as an oil, as a 1:2.7 anomeric mixture (27 mg, 65%). **IR** (film) 3307 (br), 2923 (w), 2853 (w), 1724 (s), 1659 (m), 1522 (m), 1266 (s), 1095 (s) cm^{-1}. **LRMS** (ESI^+) *m/z* 919.5 $(2M+Na)^+$ (100), 471.1 $(M+Na)^+$ (71). **HRMS** (ESI^+) for $C_{22}H_{22}F_2N_2O_6$ $(M+Na)^+$ calcd 471.1338, found 471.1342. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.09–8.03 (4 H, m, 4×CH_{Ar}), 7.65–7.39 (6 H, m, 6×CH_{Ar}), 6.12 (0.25 H, q, *J* = 8.2 Hz, CHN_{Minor}), 5.89 (0.75 H, ddd, *J* = 14.9, 9.9, 5.2 Hz, CHN_{Major}), 5.72 (0.25 H, ddd, *J* = 10.1, 8.3, 5.8 Hz, $CHOBz_{Minor}$), 5.51 (0.75 H, dd, *J* = 15.6, 5.0 Hz, $CHOBz_{Major}$), 5.41 (0.25 H, d, *J* = 9.5 Hz, NH_{Minor}), 5.28 (0.75 H, d, *J* = 9.8 Hz, NH_{Major}), 4.67–4.56 (2.25 H, m, CH_2OBz, C<u>H</u>CH_{2Minor}), 4.40 (0.75 H, q, *J* = 4.4 Hz, C<u>H</u>CH_{2Major}), 2.98 (6 H, br. s, $N(CH_3)_2$) ppm. **^{13}C NMR + DEPT** (100 MHz, $CDCl_3$) δ 166.0 (PhC=O), 164.9 (PhC=O), 156.0 (NHC=O), 134.02 (CH_{Ar}), 133.98 (CH_{Ar}), 133.3 (CH_{Ar}), 130.0 (CH_{Ar}), 129.9 (CH_{Ar}), 129.8 (CH_{Ar}), 129.4 (CH_{Ar}), 129.3 (CH_{Ar}), 128.6 (CH_{Ar}), 128.4 (CH_{Ar}), 128.3 (CH_{Ar}), 128.2 (2×C_{Ar}), 82.1 (dd, *J* = 36.9, 18.5 Hz, CHN_{Minor}), 81.6 (dd, *J* = 34.0, 19.4 Hz, CHN_{Major}), 77.1 (<u>C</u>HCH_{2Minor}), 76.2 (t, *J* = 2.9 Hz, <u>C</u>HCH_{2Major}), 72.6 (dd, *J* = 31.1, 17.5 Hz, $CHOBz_{Minor}$), 72.0 (dd, *J* = 36.0, 17.5 Hz, $CHOBz_{Major}$), 63.8 (CH_{2Minor}), 63.3 (CH_{2Major}), 36.3 ($N(CH_3)_{2Major}$), 36.2 ($N(CH_3)_{2Minor}$) ppm (some CH_{Ar} overlap, CF_2 not visible).

Major anomer: **^{19}F NMR** (282 MHz, $CDCl_3$) δ –118.4 (1 F, dt, *J* = 247.1, 15.0 Hz, C<u>F</u>F), –123.3 (1 F, dd, *J* = 247.1, 5.4 Hz, CF<u>F</u>) ppm.

Minor anomer: **^{19}F NMR** (282 MHz, $CDCl_3$) δ –114.9 (1 F, dt, *J* = 243.9, 9.7 Hz, C<u>F</u>F), –126.4 (1 F, dt, *J* = 242.8, 8.6 Hz, CF<u>F</u>) ppm.

4.3 Synthesis of the *N*,*N*-dimethyl urea derivative 15b. To lactol **12** (102 mg, 0.20 mmol) in toluene (0.8 mL) was added *N*,*N*-dimethylurea (1 equiv, 0.16 mmol, 18 mg), a catalytic amount of PTSA and molecular sieves (90 mg). The reaction mixture was stirred at reflux for 5.5 h, cooled to RT and the solvents removed *in vacuo* to yield a crude residue. This residue was first purified by column chromatography on silica gel (10:90 acetone/hexane) followed by HPLC (10:90 acetone/hexane) to yield **15b** as a colourless oil as a 1:2.9 anomeric mixture (44 mg, 38%). **IR** (film) 3070 (w), 2930 (m), 2857 (m), 1472 (m), 1456 (m) cm^{-1}.

Major anomer: **^{1}H NMR** (400 MHz, acetone-d_6) δ 7.76–7.67 (4 H, m, 4×ArH), 7.49–7.31 (11 H, m, 11×ArH), 4.87 (1 H, d, *J* = 11.8 Hz, OC<u>H</u>HPh), 4.66 (1 H, d, *J* = 11.6 Hz, OCH<u>H</u>Ph), 4.63 (1 H, m, CHN), 4.23 (1 H, ddd, *J* = 12.3, 11.2, 8.3 Hz, CHOBn), 3.97 (1 H, m, C<u>H</u>CH_2OSi), 3.84 (1 H, dd, *J* = 11.6, 3.5 Hz, C<u>H</u>HOSi), 3.75 (1 H, m, CH<u>H</u>OSi), 2.47 (6 H, s, $N(CH_3)_2$), 1.01 (9 H, s, $C(CH_3)_3$) ppm (NH not visible). **^{13}C NMR + DEPT** (100 MHz, acetone-d_6) δ 182.2 (C=O), 138.6 (C_{Ar}), 136.5 (2×CH_{Ar}), 134.1 (2×C_{Ar}), 130.84 (2×CH_{Ar}), 130.76 (2×CH_{Ar}), 129.3 (2×CH_{Ar}), 128.8 (5×CH_{Ar}), 128.7 (2×CH_{Ar}), 125.1 (dd, *J* = 259.5, 255.1 Hz, CF_2), 96.6 (dd, *J* = 31.8, 24.3 Hz, CHN), 81.7 (<u>C</u>HCH_2), 77.7 (m, <u>C</u>HOBn), 73.3 (CH_2), 63.3 (CH_2), 41.0 ($N(CH_3)_2$), 27.3 ($C(\underline{C}H_3)_3$), 19.9 ($\underline{C}(CH_3)_3$) ppm. **^{19}F NMR** (376 MHz, acetone-d_6) δ –112.3 (1 F, ddd, *J* = 240.3, 14.3, 11.0 Hz, C<u>F</u>F), –113.0 (1 F, dt, *J* = 240.3, 11.8 Hz, CF<u>F</u>) ppm.

Minor anomer: **^{1}H NMR** (400 MHz, acetone-d_6) δ 7.76–7.67 (4 H, m , 4×ArH), 7.49–7.31 (11 H, m, 11×ArH), 4.86 (1 H, d, *J* = 11.9 Hz, OC<u>H</u>HPh), 4.74 (1 H, t, *J* = 11.1 Hz, CHN), 4.65 (1 H, d, *J* = 11.7 Hz, OCH<u>H</u>Ph), 4.39 (1 H, ddd, *J* = 12.2, 9.2,

7.1 Hz, CHOBn), 4.09 (1 H, dtd, *J* = 6.9, 3.4, 1.3 Hz, C<u>H</u>CH_2OSi), 3.86 (1 H, m, C<u>H</u>HOSi), 3.76 (1 H, m, CH<u>H</u>OSi) 2.44 (6 H, s, $N(CH_3)_2$), 1.02 (9 H, s, $C(CH_3)_3$) ppm (NH not observed). **^{13}C NMR + DEPT** (100 MHz, acetone-d_6) δ 182.2 (C=O), 138.6 (C_{Ar}), 136.5 (2×CH_{Ar}), 134.0 (2×C_{Ar}), 130.9 (2×CH_{Ar}), 130.8 (2×CH_{Ar}), 129.4 (2×CH_{Ar}), 128.8 (5×CH_{Ar}), 128.7 (2×CH_{Ar}), 96.1 (dd, *J* = 36.4, 18.0 Hz, CHN), 81.6 (<u>C</u>HCH_2), 77.9 (m, CHOBn), 73.4 (CH_2), 64.4 (CH_2), 41.4 ($N(CH_3)_2$), 27.3 ($C(\underline{C}H_3)_3$), 19.9 ($\underline{C}(CH_3)_3$) ppm (CF_2 not observed). **^{19}F NMR** (376 MHz, acetone-d_6) δ –105.4 (1 F, dt, *J* = 240.8, 11.8 Hz, C<u>F</u>F), –126.4 (1 F, dt, *J* = 240.3, 11.0 Hz, CF<u>F</u>) ppm.

4.4 Synthesis of 17. To glycosylurea **14a** (93 mg, 0.22 mmol) stirring at reflux in MeCN (0.6 mL) was added dropwise a solution of acyl chloride **16** (4 equiv, 0.88 mmol, 119 mg) in MeCN (0.4 mL). The reaction mixture was stirred at reflux for 18 h, then quenched with water (0.1 mL). The solvents were reduced *in vacuo* to yield a crude residue which was taken up in DCM (15 mL), washed with water (5 mL), then brine and dried over anhydrous $MgSO_4$. The solvents were reduced *in vacuo* to yield a crude residue which was purified by flash chromatography on silica gel (10:90→40:60 acetone/petrol) to yield the desired product **17** as a colourless oil as a 1:1.5 mixture of anomers (82 mg, 71%). **IR** (film) 3246 (br), 2981 (w), 1717 (s), 1680 (s), 1603 (m), 1537 (s), 1492 (w), 1452 (m), 1378 (w), 1316 (m), 1247 (s), 1176 (m), 1093 (s) cm^{-1}. **LRMS** (ESI^+) *m/z* 1059.8 $(2M+Na)^+$ (50), 541.3 $(M+Na)^+$ (100). **HRMS** (ESI^+) for $C_{25}H_{24}F_2N_2O_8$ $(M+Na)^+$ Calcd: 541.1393; Found: 541.1411. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 10.01 (0.5 H, d, *J* = 9.2 Hz, 0.5×CHN<u>H</u>), 9.79 (0.5 H, d, *J* = 9.5 Hz, 0.5×CHN<u>H</u>), 9.65 (0.5 H, s, 0.5×O=CN<u>H</u>C=O), 9.28 (0.5 H, s, 0.5×O=CN<u>H</u>C=O), 8.22–8.20 (1 H, m, CH_{Ar}), 8.08–8.02 (3 H, m, 3×CH_{Ar}), 7.69 (1 H, d, *J* = 11.9 Hz, H<u>C</u>=CHOEt), 7.65–7.40 (6 H, m, 6×CH_{Ar}), 6.08 (0.5 H, t, *J* = 9.3 Hz, 0.5×CHN), 5.90 (0.5 H, ddd, *J* = 12.9, 9.5, 6.0 Hz, 0.5×CHN), 5.70–5.67 (0.5 H, m, 0.5×CHOBn), 5.57 (0.5 H, ddd, *J* = 14.7, 5.1, 1.0 Hz, 0.5×CHOBn), 5.37 (0.5 H, d, *J* = 12.1 Hz,

0.5×CHOEt), 5.34 (0.5 H, d, *J* = 12.0 Hz, 0.5×CHOEt), 4.69–4.58 (2.5 H, m, CHC$\underline{H}_2$, 0.5×C$\underline{H}$CH$_2$), 4.43 (0.5 H, q, *J* = 4.5 Hz, 0.5×C$\underline{H}$CH$_2$), 3.98 (2 H, q, *J* = 7.1 Hz, C$\underline{H}_2$CH$_3$), 1.34 (1.5 H, t, *J* = 7.0 Hz, 0.5×CH$_3$), 1.33 (1.5 H, t, *J* = 7.1 Hz, 0.5×CH$_3$) ppm. **^{13}C NMR + DEPT** (100 MHz, $CDCl_3$) δ 168.3 (0.5×$\underline{C}$=OCH=CH), 168.1 (0.5×$\underline{C}$=OCH=CH), 166.07 (0.5×Ph$\underline{C}$OCH$_2$), 166.05 (0.5×Ph$\underline{C}$OCH$_2$), 164.9 (0.5×Ph$\underline{C}$OCH), 164.7 (0.5×Ph$\underline{C}$OCH), 164.0 (0.5×CH=$\underline{C}$HOEt), 163.6 (0.5×CH=$\underline{C}$HOEt), 155.1 (0.5×NC=ON), 154.9 (0.5×NC=ON), 134.0 (CH$_{Ar}$), 133.3 (CH$_{Ar}$), 133.2 (CH$_{Ar}$), 130.4 (CH$_{Ar}$), 130.0 (CH$_{Ar}$), 129.8 (CH$_{Ar}$), 128.6 (CH$_{Ar}$), 128.5 (CH$_{Ar}$), 128.4 (CH$_{Ar}$), 129.33 (0.5×C$_{Ar}$), 129.30 (0.5×C$_{Ar}$), 128.2 (0.5×C$_{Ar}$), 128.0 (0.5×C$_{Ar}$), 97.5 (H$\underline{C}$=CHOEt), 81.2 (dd, *J* = 39.5, 22.0 Hz, 0.5×CHN), 80.5 (dd, *J* = 36.6, 22.0 Hz, 0.5×CHN), 79.3 (0.5×$\underline{C}$HCH$_2$), 76.7 (0.5×$\underline{C}$HCH$_2$), 72.2 (dd, *J* = 33.7, 17.6 Hz, 0.5×CHOBz), 72.0 (dd, *J* = 35.1, 17.6 Hz, 0.5×CHOBz), 68.0 (0.5×$\underline{C}$H$_2$CH$_3$), 67.7 (0.5×$\underline{C}$H$_2$CH$_3$), 63.3 (0.5×CH$\underline{C}$H$_2$), 63.2 (0.5×CH$\underline{C}$H$_2$), 14.43 (0.5×CH$_3$), 14.36 (0.5×CH$_3$) ppm (some CH$_{Ar}$ overlap, CF$_2$ not visible). **^{19}F NMR** (282 MHz, $CDCl_3$) δ Anomer 1: –109.5 (1 F, ddd, *J* = 251.4, 16.1, 6.4 Hz, C$\underline{F}$F), –125.4 (1 F, d, *J* = 252.5 Hz, CF$\underline{F}$) ppm; Anomer 2: –123.6 (1 F, dd, *J* = 241.8, 9.7 Hz, C$\underline{F}$F), –125.5 (1 F, ddd, *J* = 241.8, 10.7, 6.4 Hz, CF$\underline{F}$) ppm. (Data assigned from an earlier experiment which gave a 1:1 mixture of anomers).

4.5 Synthesis of the uridine 20. Acyl carbamate **17** (136 mg, 0.26 mmol) was stirred in 1:10 HCl/AcOH (2.75 mL) at RT in a stoppered flask for 24 h. The solvents were reduced *in vacuo* to yield the desired protected difluorouracil **20** as a yellow oil as a 1:1.2 α/β mixture of anomers (125 mg, quantitative yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.91 (1 H, br. s, NH), 8.11–8.02 (4 H, m, CH$_{Ar}$), 7.69–7.39 (7 H, m, 6×CH$_{Ar}$, NC$\underline{H}$=CH), 6.55 (0.45 H, t, *J* = 7.5 Hz, CHN$_\alpha$), 6.40 (0.55 H, dd, *J* = 11.8, 6.4 Hz, CHN$_\beta$), 5.82 (0.45 H, d, *J* = 8.1 Hz, NCH=C$\underline{H}_\alpha$), 5.79 (0.45 H, m, CHOBz$_\alpha$), 5.68 (0.55 H, d, *J* = 8.3 Hz, NCH=C$\underline{H}_\beta$), 5.65 (0.55 H, m, CHOBz$_\beta$), 4.89–4.81 (1 H, m, CHC$\underline{H}$H), 4.73–4.59 (2 H, m, C$\underline{H}$O, CHCH$\underline{H}$) ppm. **^{19}F NMR** (282 MHz,

$CDCl_3$) δ –109.2 (0.45 F, d, *J* = 249.3 Hz, C<u>F</u>F), –115.5 (0.55 F, dt, *J* = 247.4, 12.9 Hz, CF<u>F</u>), –120.3 (0.5 F, d, *J* = 246.6 Hz, C<u>F</u>F), –122.0 (0.5 F, d, *J* = 250.9 Hz, CF<u>F</u>) ppm.

4.6 Synthesis of gemcitabine 1. To crude protected difluorouridine **20** (378 mg, 0.75 mmol) in pyridine (9 mL) was added 1,2,4-triazole (3 equiv, 2.26 mmol, 156 mg) and 2-chlorophenyl phosphorodichloridate (6 equiv, 4.51 mmol, 0.74 mL). The reaction mixture was stirred at RT for 5 d. The solvents were reduced *in vacuo* to yield a crude residue which was taken up in DCM and washed with sat. $NaHCO_3$ (aq), then brine and dried over anhydrous Na_2SO_4. The solvents were reduced *in vacuo* to yield a crude oil which was passed through a short silica column (10:90 acetone/petrol) to yield the impure triazole **21** (94 mg, ~24%). This triazole (76.5 mg, 0.17 mol) was stirred in 7 N NH_3 in MeOH (5 mL) at RT for 36 h. The solvents were reduced *in vacuo* and the resultant residue evaporated onto silica gel for purification by column chromatography (10:90–20:80 MeOH/DCM) to yield the desired difluorocytidine **1** as a 1:1.3 α/β mixture of anomers (18.7 mg, 49%). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 7.69 (0.6 H, d, *J* = 7.6 Hz, NC<u>H</u>=CH_β), 7.52 (0.4 H, dd, *J* = 7.5, 1.6 Hz, NC<u>H</u>=CH_α), 7.37–7.30 (2 H, m, NH_2), 6.30–6.26 (0.8 H, m, CHN_α, $CHO\underline{H}_\alpha$), 6.22 (0.6 H, d, *J* = 6.6 Hz, $CHO\underline{H}_\beta$), 6.13 (0.6 H, t, *J* = 8.3 Hz, CHN_β), 5.79 (0.4 H, d, *J* = 7.5 Hz, NCH=C<u>H</u>$_\alpha$), 5.78 (0.6 H, d, *J* = 7.6 Hz, NCH=C<u>H</u>$_\beta$), 5.19 (0.6 H, t, *J* = 5.4 Hz, $CH_2O\underline{H}_\beta$), 5.07 (0.4 H, t, *J* = 5.7 Hz, $CH_2O\underline{H}_\alpha$), 4.34 (0.4 H, m, C<u>H</u>OH_α), 4.18–4.07 (1.2 H, m, C<u>H</u>OH_β, C<u>H</u>$CH_{2(\beta)}$), 3.81–3.74 (1 H, m, C<u>H</u>$CH_{2\alpha}$, CHC<u>H</u>H_β), 3.64–3.59 (1 H, m, CHC<u>H</u>H), 3.53 (0.4 H, m, CHCH<u>H</u>$_\alpha$) ppm. **^{13}C NMR + DEPT** (100 MHz, DMSO-d_6) δ 165.7 (0.5×NCON), 165.6 (0.5×NCON), 154.8 (0.5×N=C), 154.7 (0.5×N=C), 141.3 (0.5×NC=<u>C</u>), 140.8 (0.5×NC=<u>C</u>), 123.1 (t, *J* = 252.6 Hz, CF_2), 94.6 (0.5×N<u>C</u>=C), 94.4 (0.5×N<u>C</u>=C), 83.9–83.2 (m, CHN), 80.4 (<u>C</u>HCH_2), 69.6 (dd, *J* = 26.3, 17.6 Hz, 0.5×CHOH), 68.7 (t, *J* = 22.0 Hz, CHOH), 60.0 (0.5×CH_2), 59.0

(0.5×CH_2) ppm. **^{19}F NMR** (282 MHz, DMSO-d_6) δ –114.3 (0.4 F, d, *J* = 232.8 Hz, $C\underline{F}F_\alpha$), –115.5—117.3 (1.2 F, m, $CF_{2(\beta)}$), –124.3 (0.4 F, d, *J* = 234.5 Hz, $CF\underline{F}_\alpha$) ppm.

Data consistent with the literature.[12]

Acknowledgements

We thank Dextra Laboratories and the EPSRC for a CASE award (EP/P503841/1) and a Knowledge Transfer Secondment (EP/H500243/1). We are grateful for EPSRC Core Capability Funding (EP/K039466/1).

1. Wojtowicz-Rajchel, H. *J. Fluorine Chem.* **2012**, *143*, 11–48.
2. Qiu, X.-L.; Xu, X.-H.; Qing, F.-L. *Tetrahedron* **2010**, *66*, 789–843.
3. Liu, P.; Sharon, A.; Chu, C. K. *J. Fluorine Chem.* **2008**, *129*, 743–766.
4. Watts, J. K.; Damha, M. J. *Can. J. Chem.* **2008**, *86*, 641–656.
5. Meng, W.-D.; Qing, F.-L. *Curr. Top. Med. Chem.* **2006**, *6*, 1499–1528.
6. Pankiewicz, K. W. *Carbohydr. Res.* **2000**, *327*, 87–105.
7. Sneader, W. In *Drug Discovery: a history*; Wiley, New York, 2005; p259.
8. Sweetman, S. C., Ed. *Martindale 2011. The complete drug reference*, 37th ed.; Royal Pharmaceutical Society: London, 2011.
9. Silvestris, N .; Cinieri, S.; La Torre, I.; Pezzella, G.; Numico, G.; Orlando, L.; Lorusso, V. *The Breast* **2008**, *17*, 220–226.
10. Vulfovich, M.; Rocha-Lima, C.; *Expert Rev. Anticancer Ther.* **2008**, *8*, 993–1002.
11. Lorusso, D.; Di Stefano, A.; Fanfani, F.; Scambia, G. *Ann. Oncol.* **2006**, *17*, V188–V194.
12. First report: Hertel, L. W.; Kroin, J. S.; Misner, J. W.; Tustin, J. M. *J. Org. Chem.* **1988**, *53*, 2406–2409.
13. Review: Brown, K.; Dixey, M.; Weymouth-Wilson, A.; Linclau B. *Carbohydr. Res.* **2014**, *387*, 59–73
14. Crimmins, M. T. *Tetrahedron* **1998**, *54*, 9229–9272.
15. Leung, L. M. H.; Gibson,V.; Linclau, B. *J. Org. Chem*. **2008**, *73*, 9197–9206.
16. Ludek, O. R.; Marquez, V. E. *Synthesis* **2007**, 3451–3460.
17. Gong, C. US0003963, 2006.
18. Hertel, L. W.; Jones, C. D.; Kroin, J. S.; Mabry, T. E. US5594155, 1997.
19. Chou, T. S.; Heath, P. C.; Patterson, L. E.; Poteet, L. M.; Lakin, R. E.; Hunt, A. H. *Synthesis* **1992**, 565–570.
20. Muñoz-Muñiz, O. Juaristi, E. *ARKIVOC* **2003**, *11*, 16–26.
21. Armstrong, A.; Edmonds, I. D.; Swarbrick, M. E.; Treweeke, N. R. *Tetrahedron* **2005**, *61*, 8423–8442.
22. Schoorl, M. N. *Rec. Trav. Chim. Pay-Bas* **1903**, *22*, 31–77.
23. Benn, M. H.; Jones, A. S. *J. Chem. Soc. (Resumed)* **1960**, 3837–3841.
24. Review: McKay, M. J.; Nguyen, H. M. *Carbohydr. Res.* **2014**, *385*, 18–44.
25. Tietze, L. F.; Schneider, C.; Pretor, M. *Synthesis* **1993**, 1079–1080.
26. Cookson, R. C., Dudfield, P. J. *J. Chem. Soc. Perkin I* **1986**, 399–404.
27. Balzarini, J.; Baumgartner, H.; Bodenteich, M.; De Clercq, E.; Griengl, H. *J. Med. Chem.* **1989**, *32*, 1861–1865.
28. Shaw, G.; Warrener, R. N. *J. Chem. Soc.* **1958**, 157–161.
29. Sung, W. L. *J. Org. Chem.* **1982**, *47*, 3623–3628.

30. Divakar, K. J.; Reese, C. B. *J. Chem. Soc. Perkin I* **1982**, 1171–1176.

* A linear synthesis of gemcitabine has been achieved
* The key step is a urea introduction to an unactivated 2,2-difluorinated furanose
* Only moderate anomeric selectivity was obtained

A Linear Synthesis of Gemcitabine

Kylie Brown,[1,2] Alex Weymouth-Wilson,[2] Bruno Linclau*,[1]

[1] Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ.
[2] Dextra Laboratories Ltd, The Science and Technology Centre, Earley Gate, Whiteknights Road, Reading RG6 6BZ, UK

bruno.linclau@soton.ac.uk

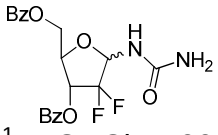
SUPPORTING INFORMATION

Table of Contents:

General information:

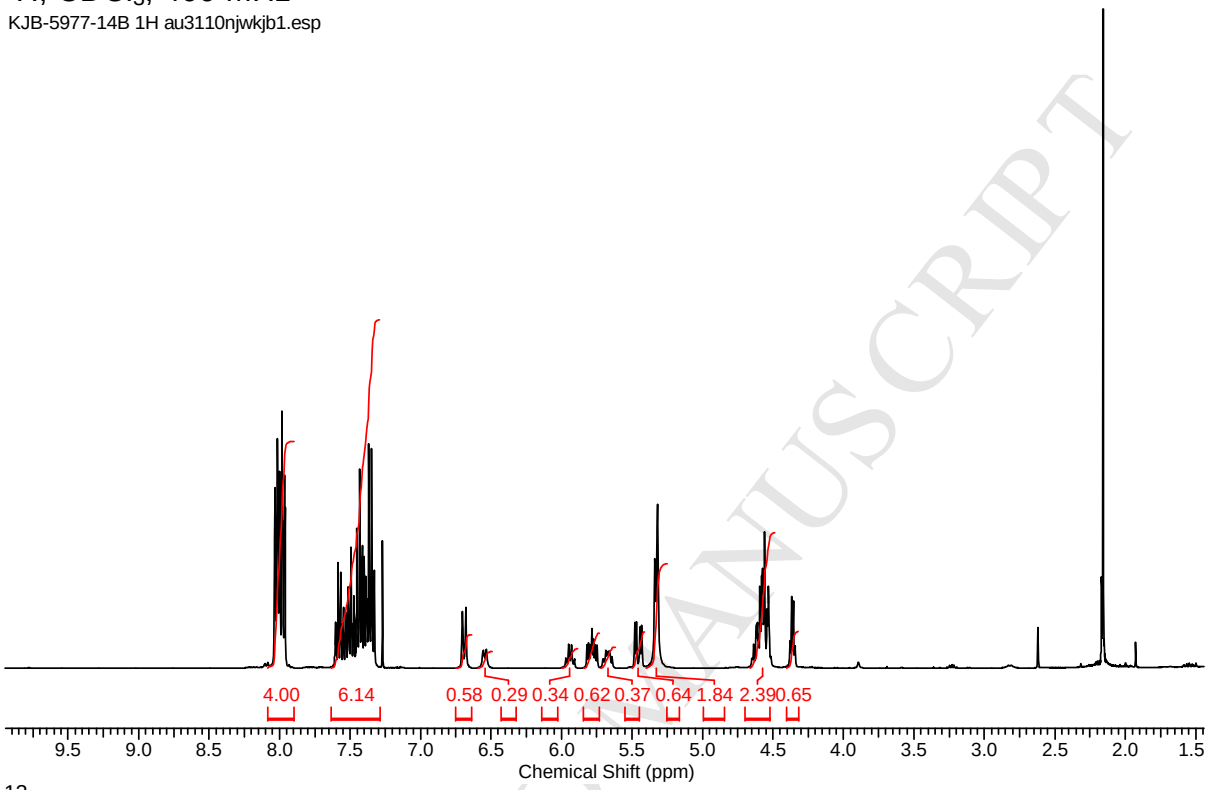
^{1}H and ^{13}C NMR spectra were recorded at room temperature on a Bruker DPX400 or AV300 spectrometer as indicated. Chemical shifts are quoted in ppm relative to residual solvent peaks as appropriate. ^{19}F NMR spectra were recorded on a Bruker AV300 spectrometer or a Bruker DPX400 spectrometer and are referenced to C_6F_6 and $CFCl_3$ respectively. Assignments were assisted by COSY and HMQC experiments. LRMS were recorded on a Thermoquest Trace GCMS Quadrapole system. HRMS were recorded on a VG Analytical 70-250-SE, normal geometry, double focussing. Infrared spectra were recorded as neat films on a Thermo Matson Fourier Transform spectrometer. Melting points were recorded on a Gallencamp Melting Point Apparatus and are uncorrected. Column chromatography was performed on silica gel (60 Å, 35-70 microns). Preparative HPLC was carried out using a Biorad Biosil D 90-10, 250 × 22 mm column eluting at 20 mL min^{-1}, connected to a Kontron 475 refractive index detector. Reactions were monitored by TLC (Merck). Reaction solvents were dried before use as follows: THF and Et2O were distilled from sodium/benzophenone ketyl; DCM, pyridine and Et_3N were distilled from CaH_2; toluene was distilled from Na; MeCN was dried over 3 Å molecular sieves. Reaction vessels were flame dried under vacuum and cooled under N_2 prior to use and all experiments were carried out under a N_2 atmosphere. All other reagents were purchased from commercial sources and used without further purification.

1) Urea 14a



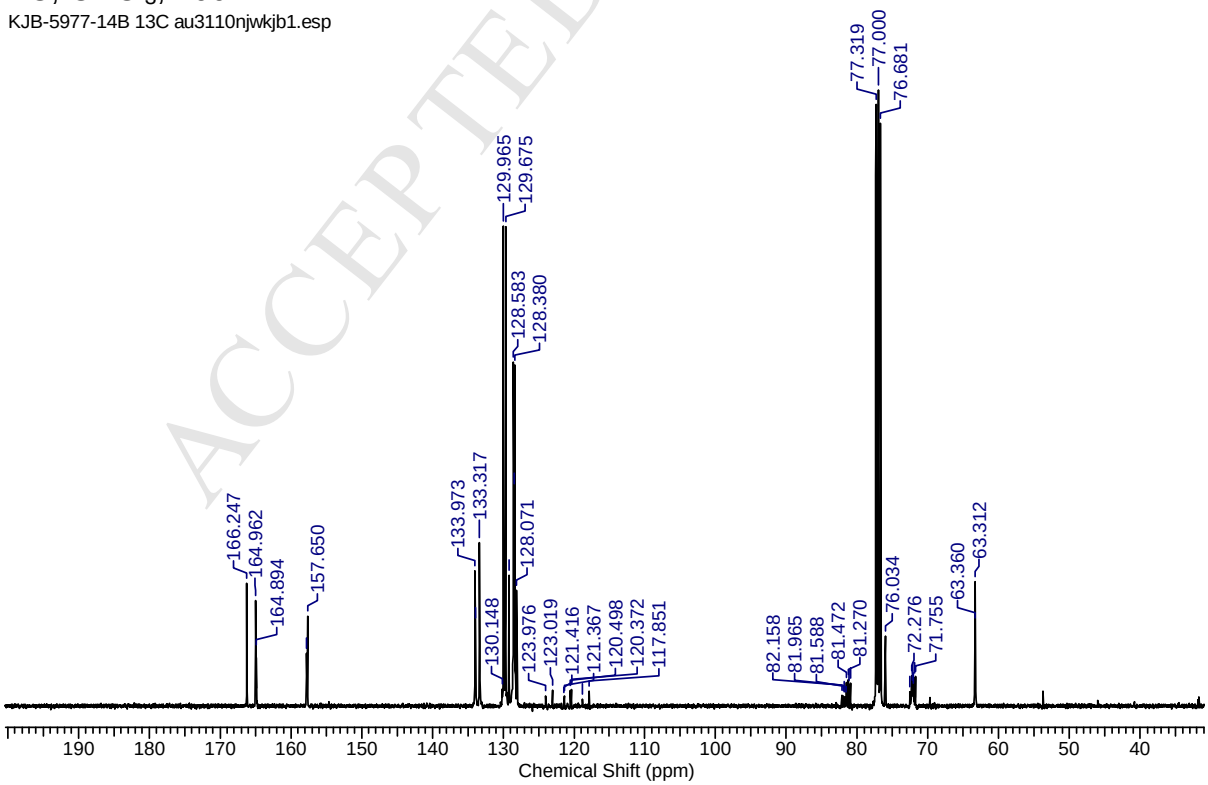
^{1}H, $CDCl_3$, 400 MHz

KJB-5977-14B 1H au3110njwkjb1.esp



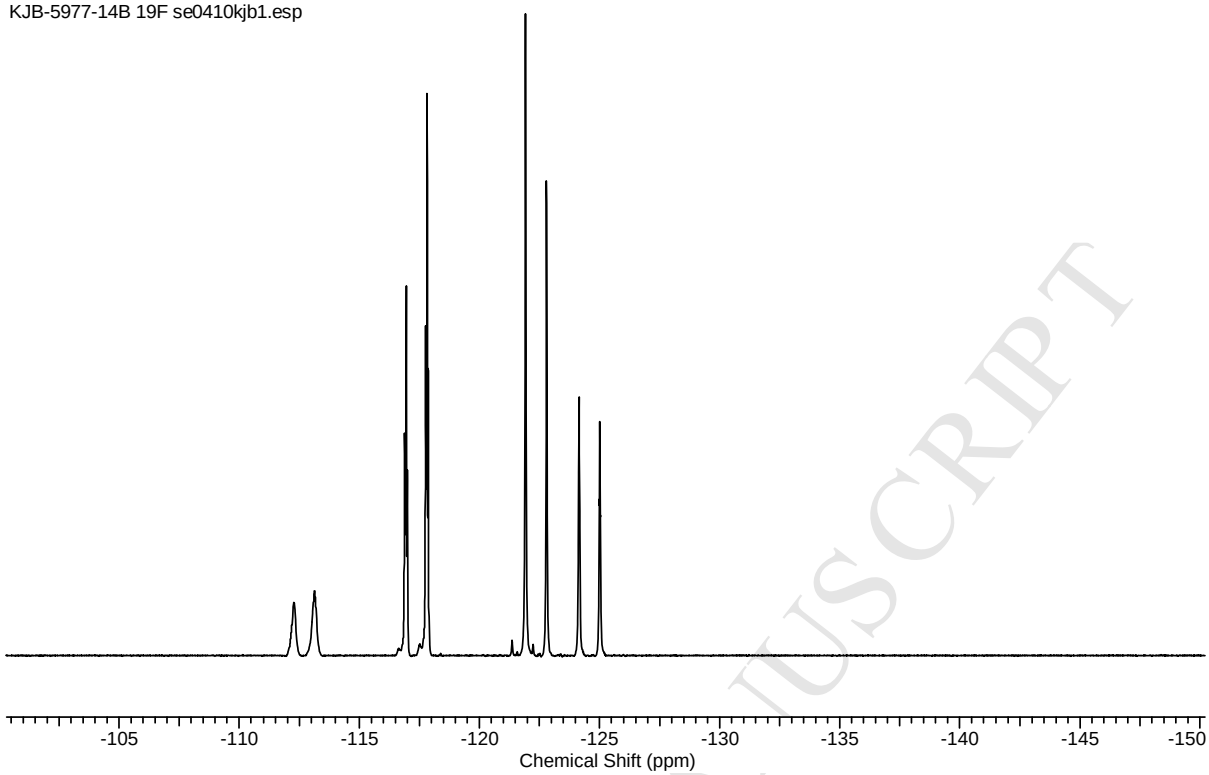
^{13}C, $CDCl_3$, 100 MHz

KJB-5977-14B 13C au3110njwkjb1.esp

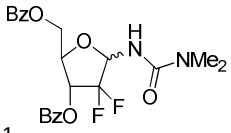


^{19}F, $CDCl_3$, 282 MHz

KJB-5977-14B 19F se0410kjb1.esp

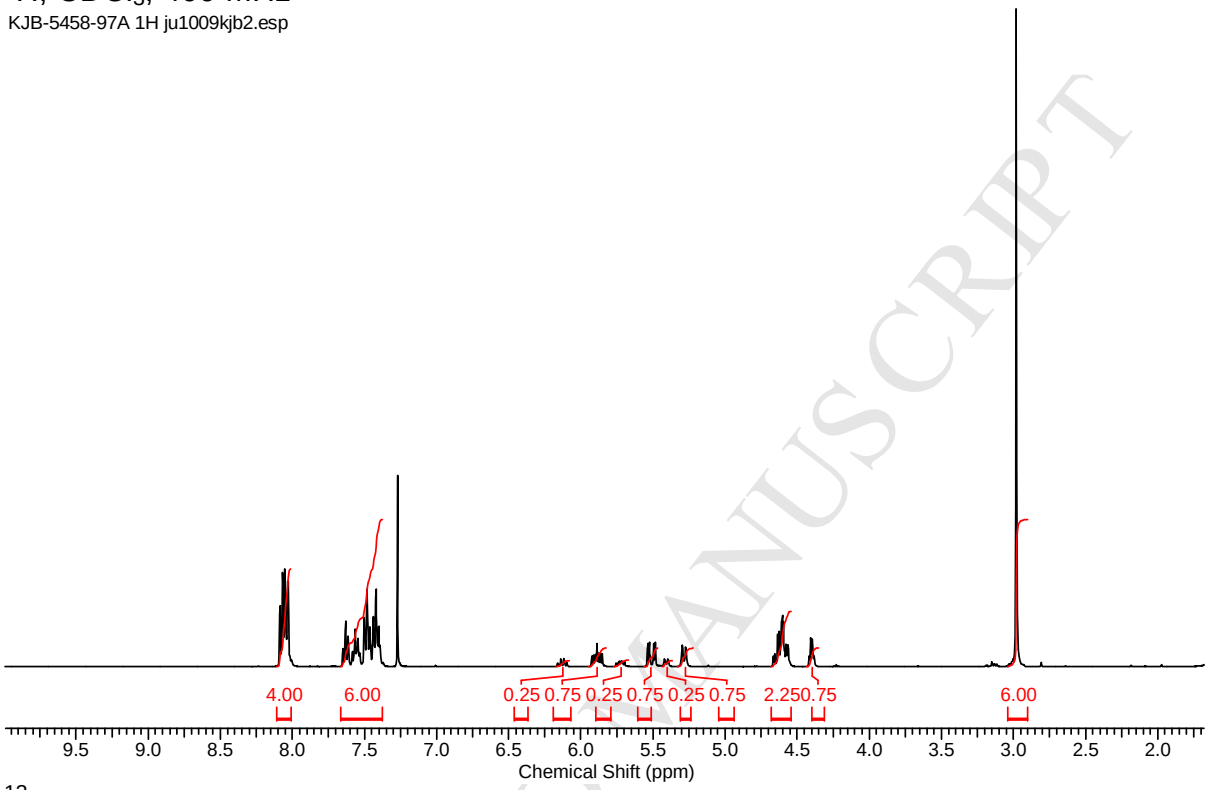


2) Urea 14b



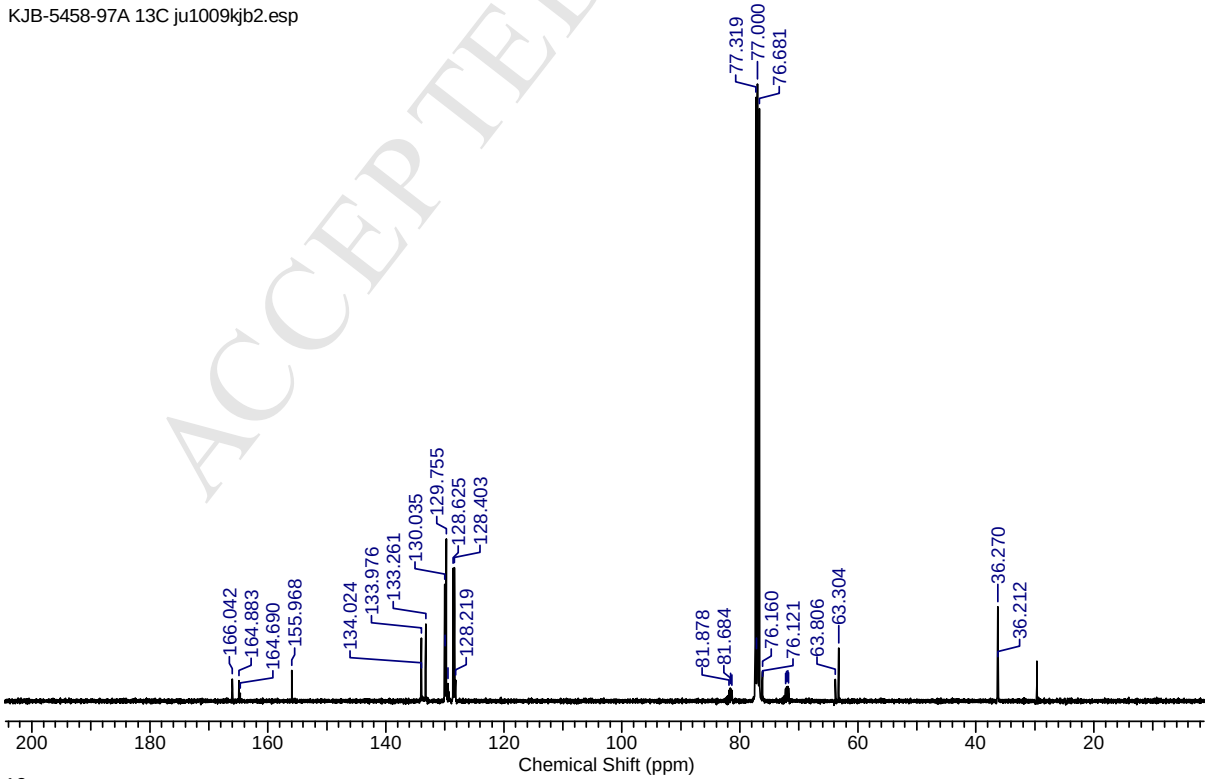
^{1}H, $CDCl_3$, 400 MHz

KJB-5458-97A 1H ju1009kjb2.esp

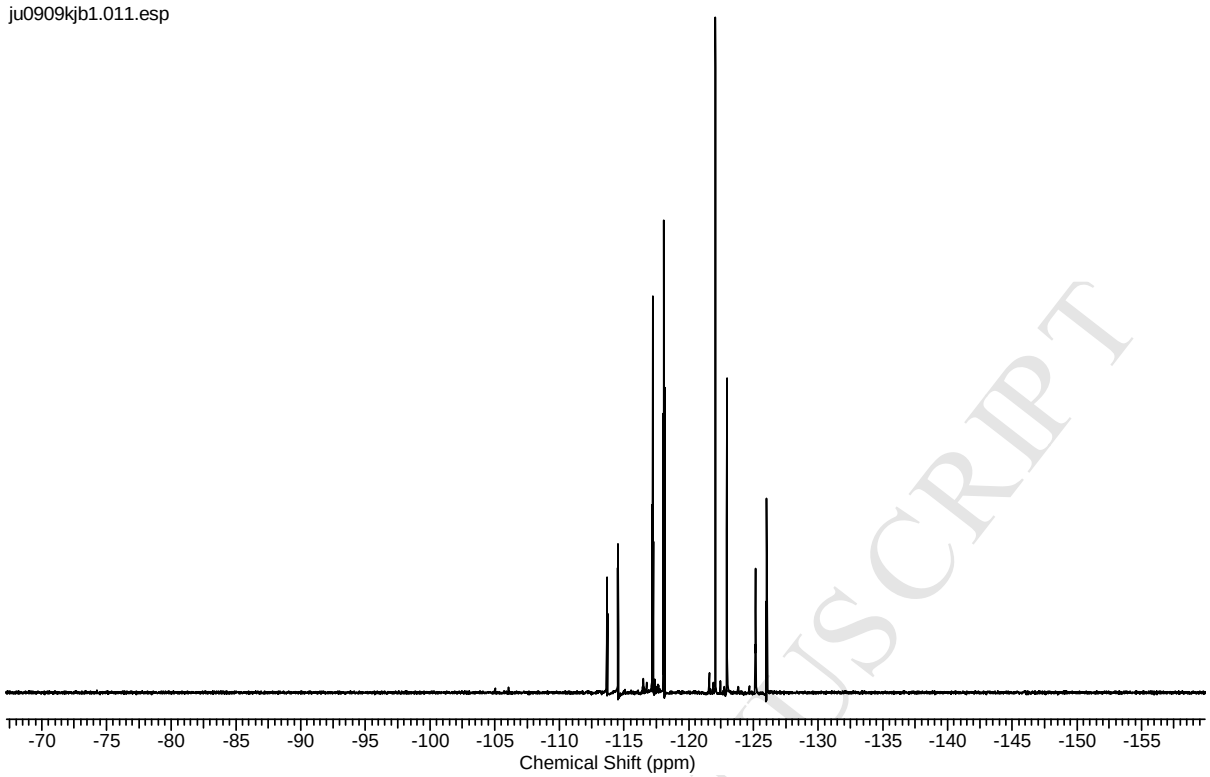


^{13}C, $CDCl_3$, 100 MHz

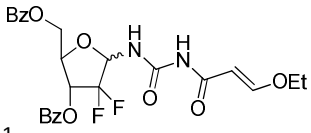
KJB-5458-97A 13C ju1009kjb2.esp



^{19}F, $CDCl_3$, 282 MHz

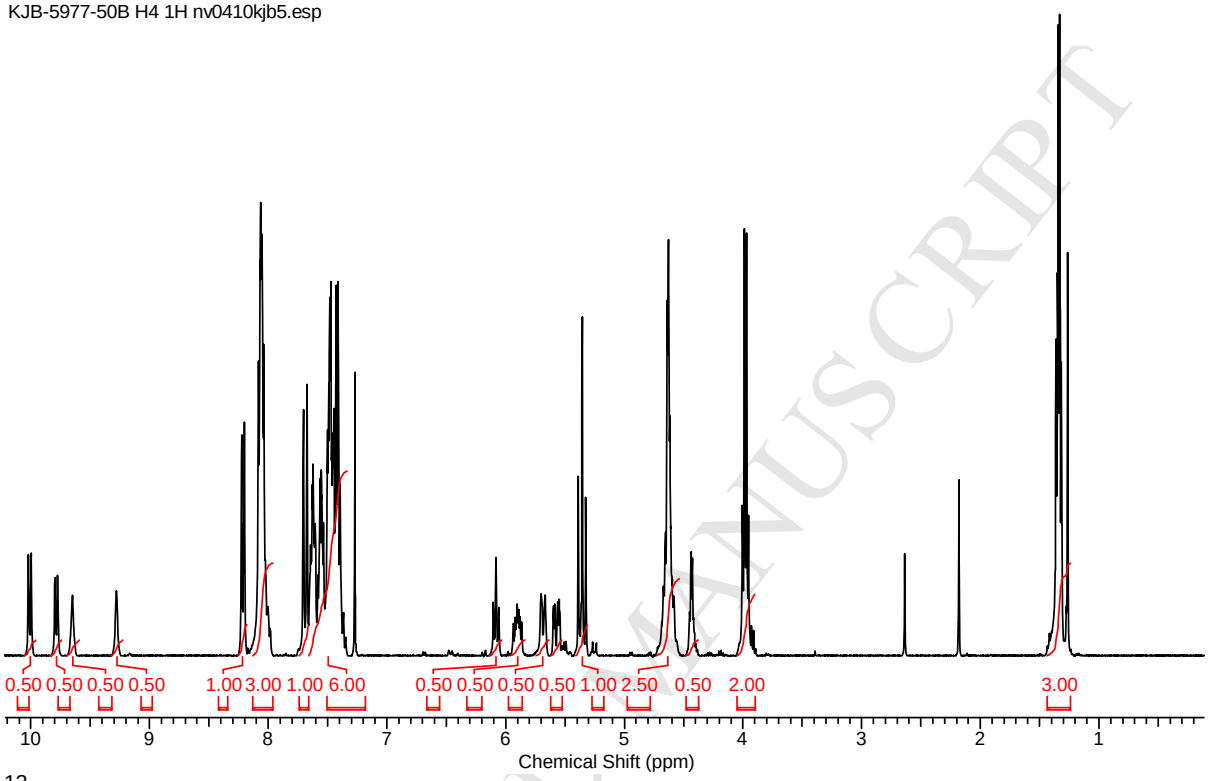


3) Acyl carbamate 17



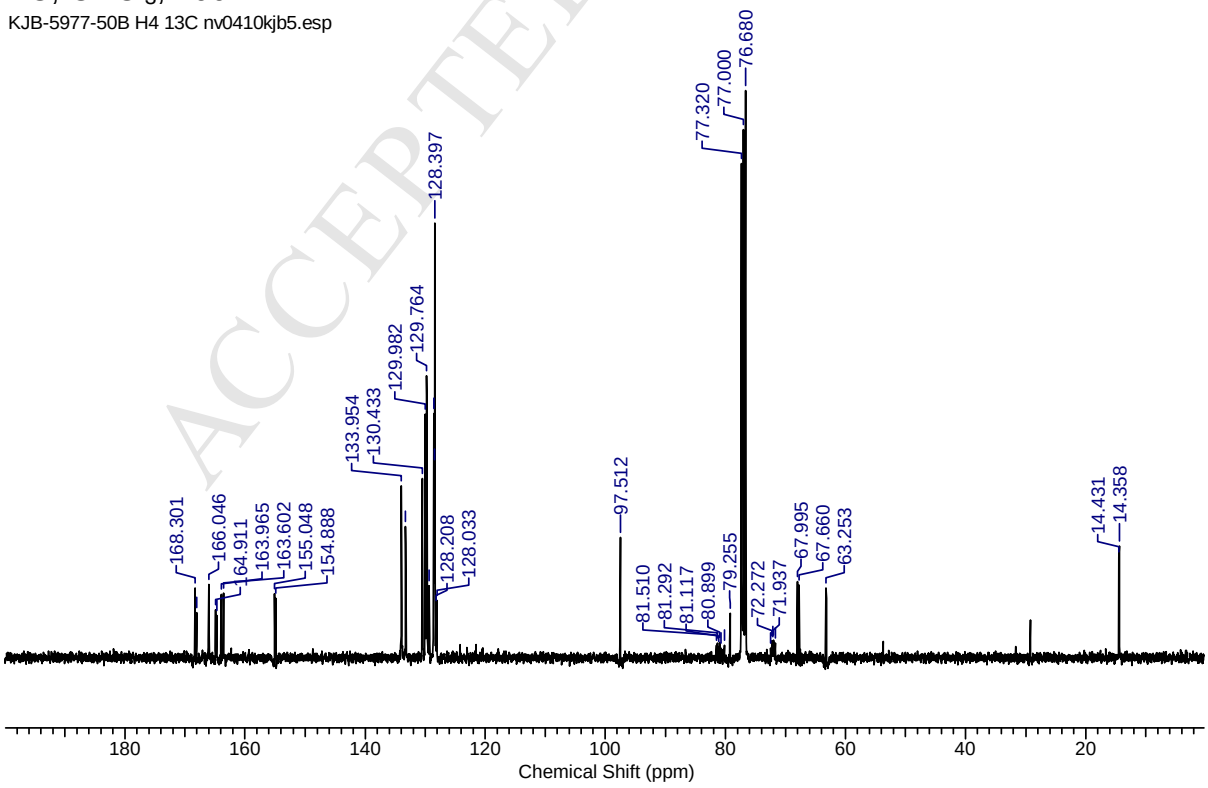
^{1}H, $CDCl_3$, 400 MHz

KJB-5977-50B H4 1H nv0410kjb5.esp



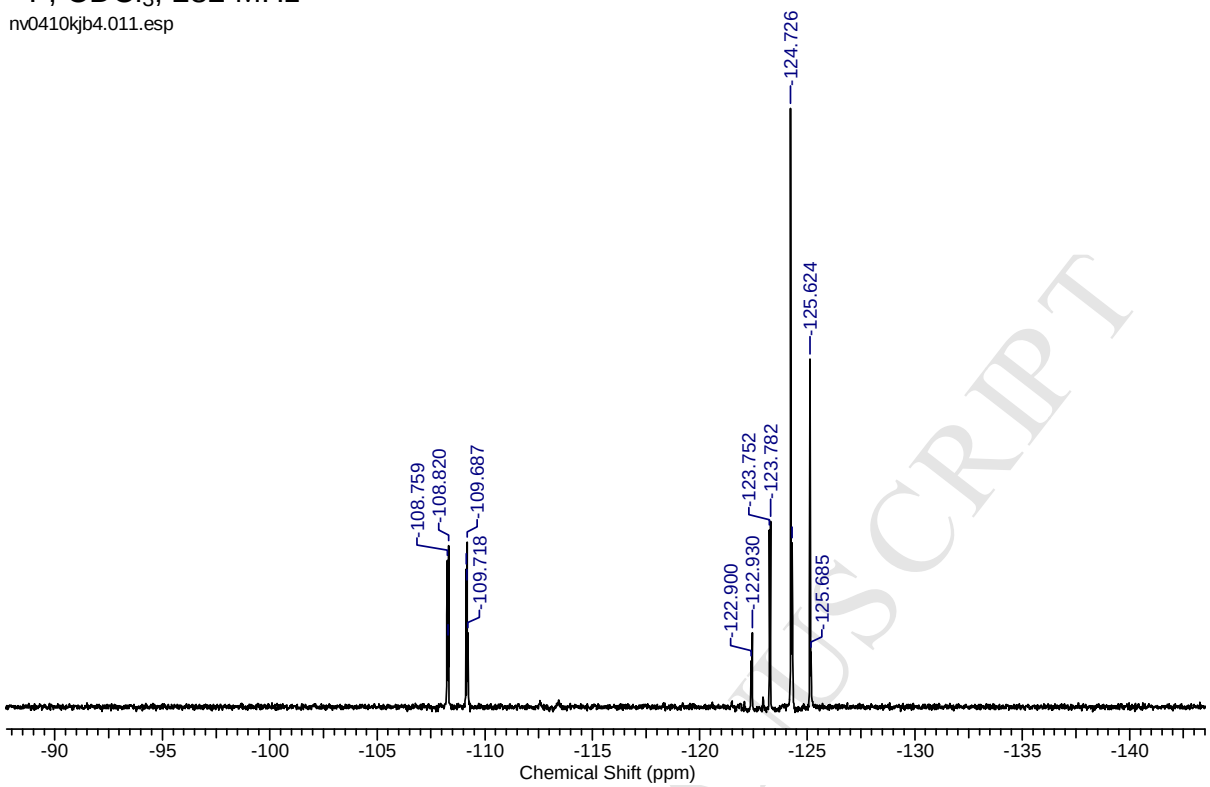
^{13}C, $CDCl_3$, 100 MHz

KJB-5977-50B H4 13C nv0410kjb5.esp

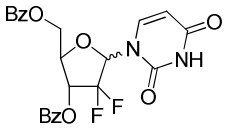


^{19}F, $CDCl_3$, 282 MHz

nv0410kjb4.011.esp

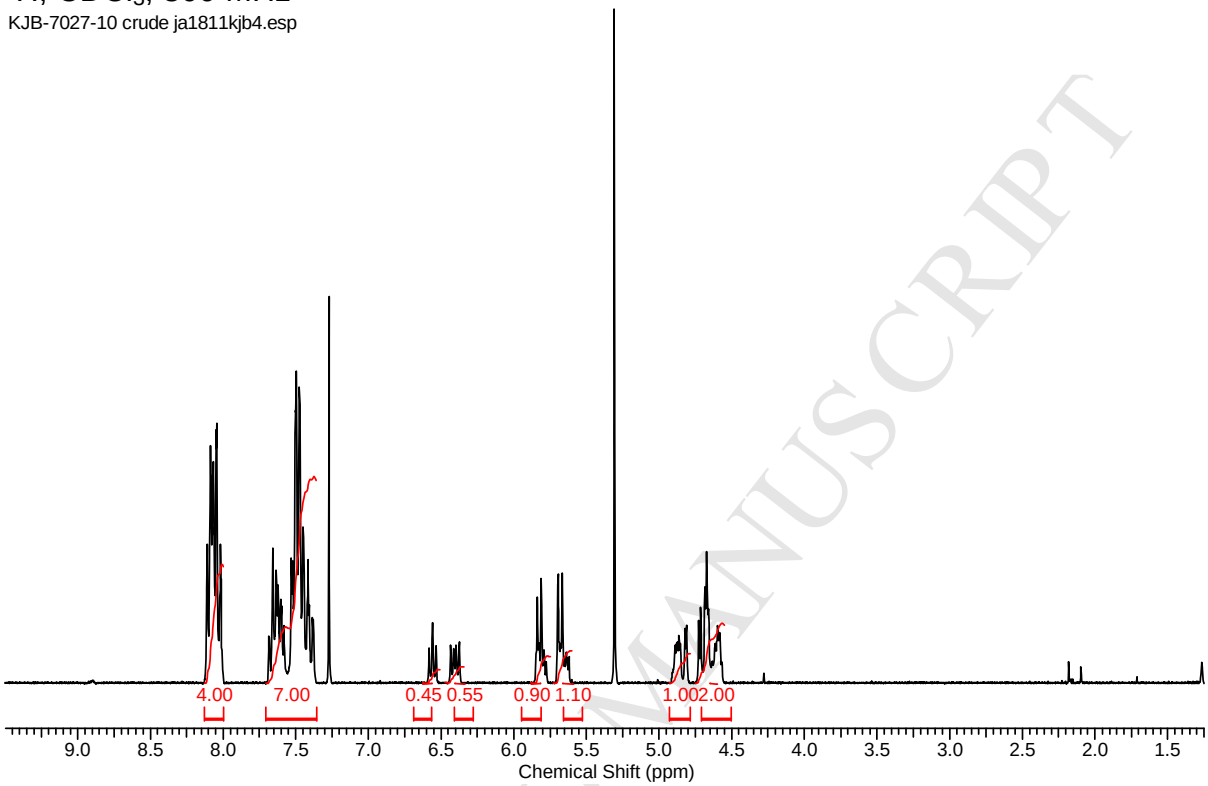


4) Uridine 20, including comparison with commercial sample



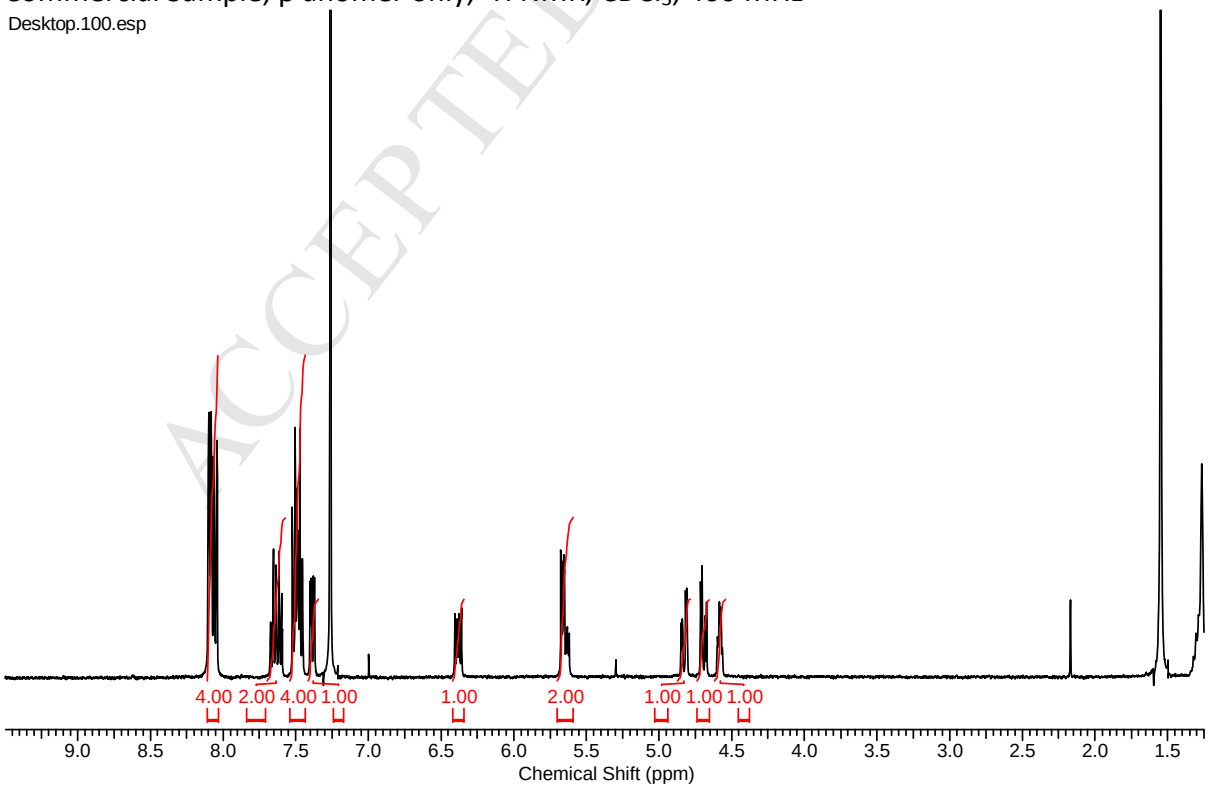
^{1}H, $CDCl_3$, 300 MHz

KJB-7027-10 crude ja1811kjb4.esp



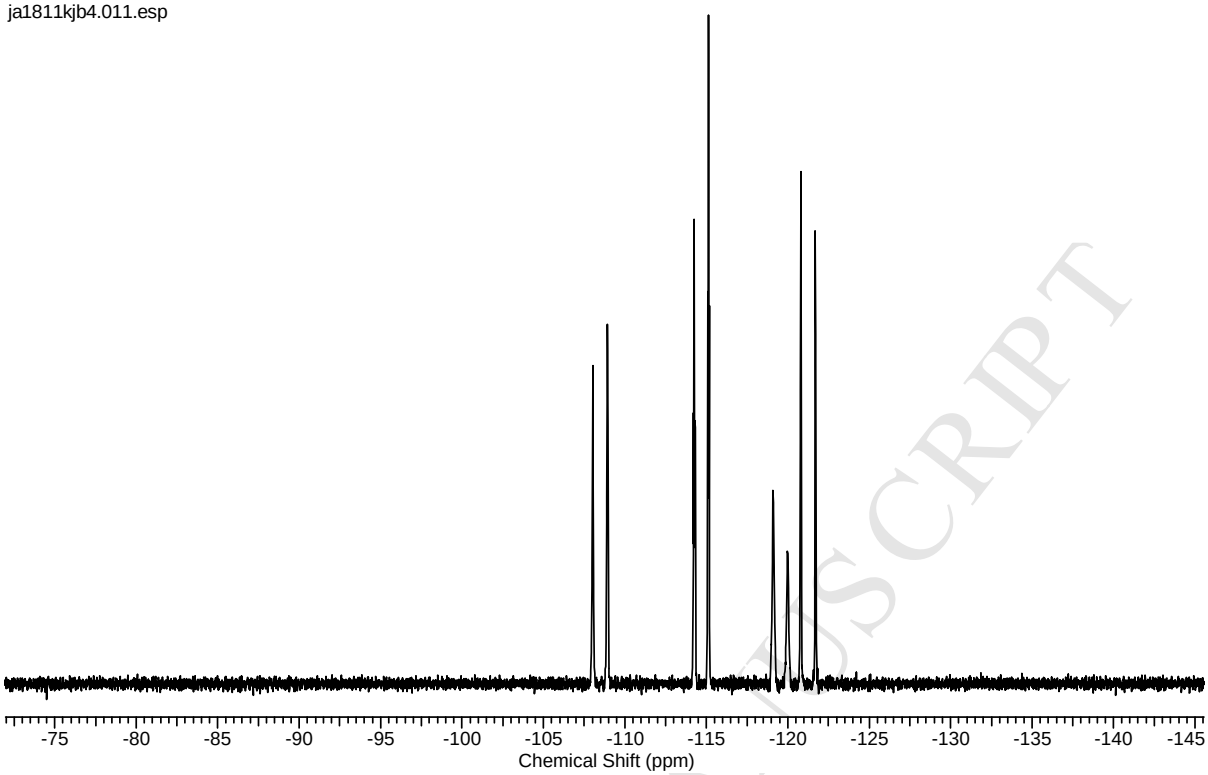
Commercial Sample, β anomer only, ^{1}H NMR, $CDCl_3$, 400 MHz

Desktop.100.esp



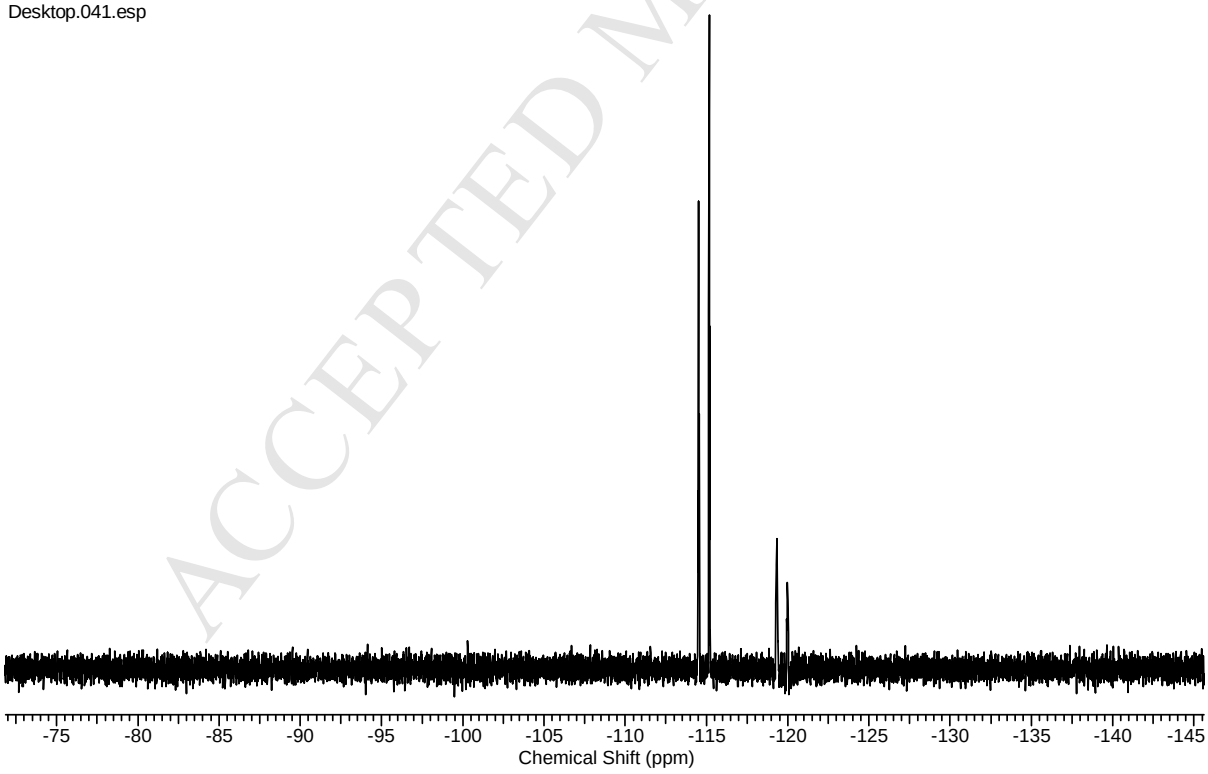
^{19}F, $CDCl_3$, 282 MHz

ja1811kjb4.011.esp

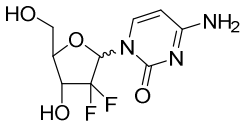


Commercial Sample, β anomer only, ^{1}F NMR, $CDCl_3$, 376 MHz

Desktop.041.esp

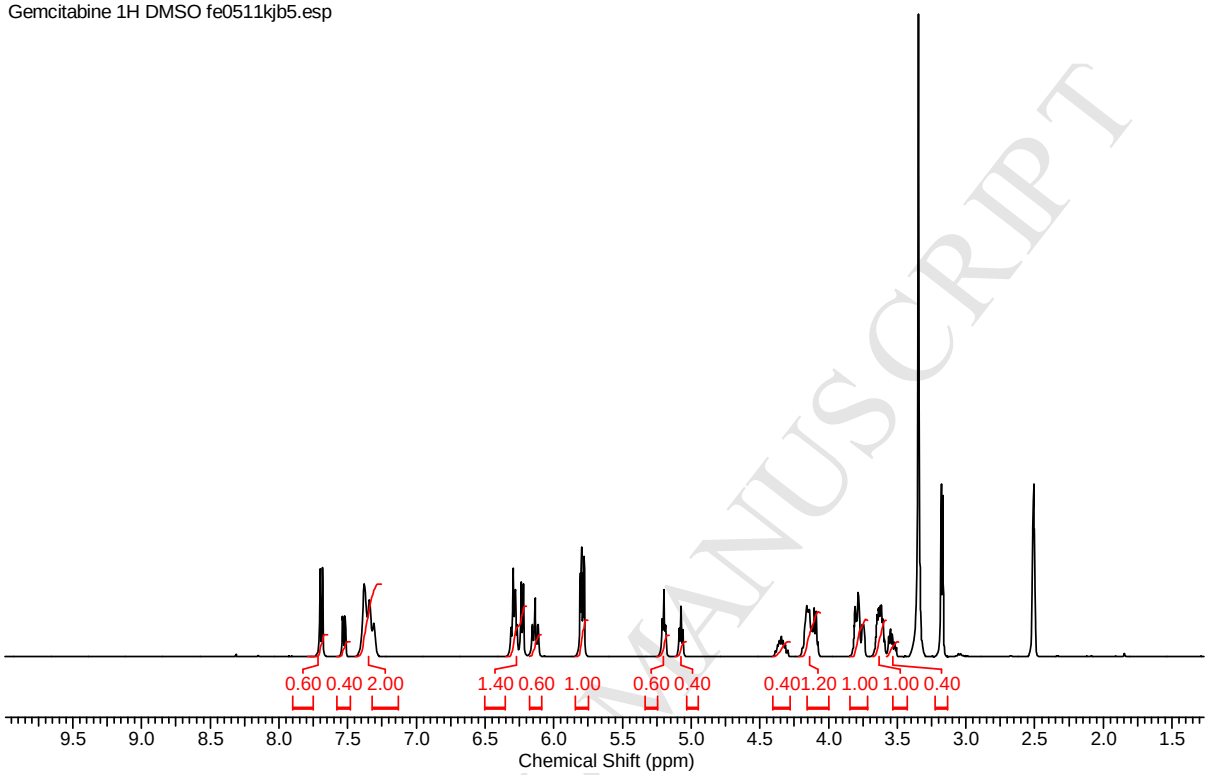


5) Gemcitabine 1, including comparison with commercial sample



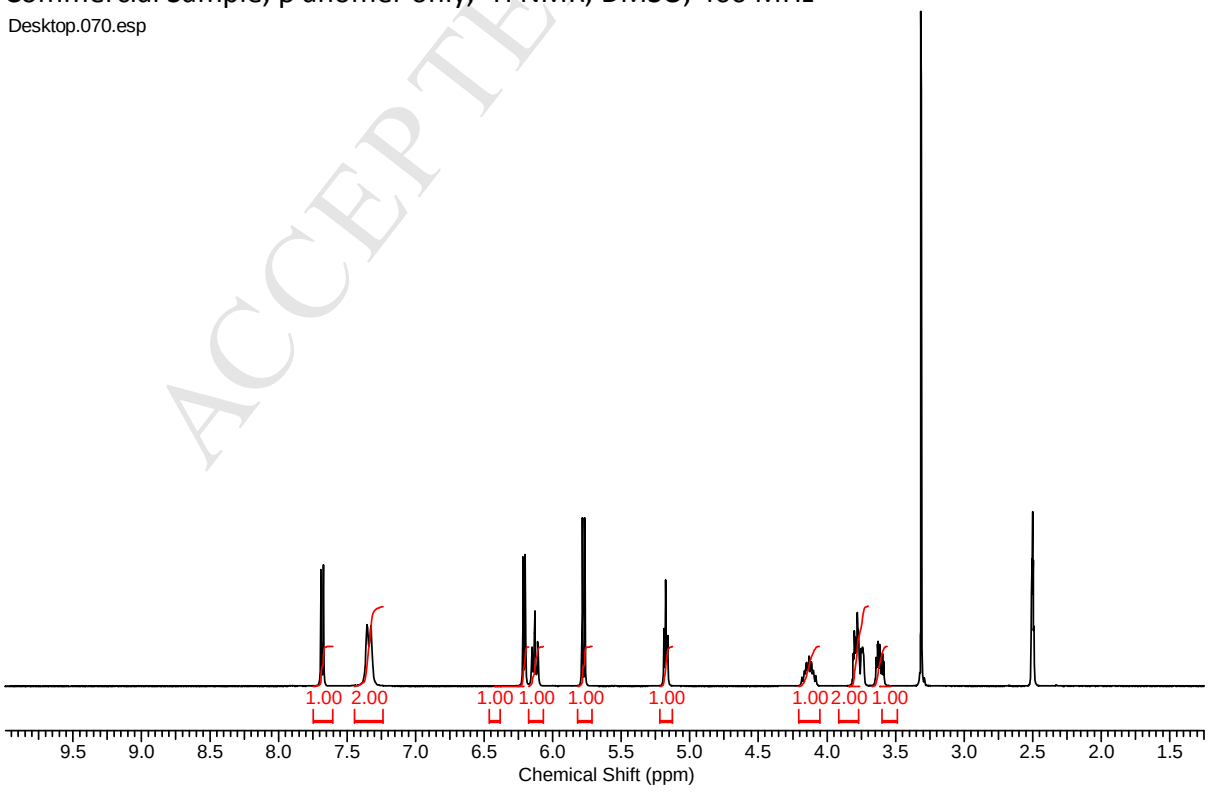
^{1}H, DMSO, 400 MHz

Gemcitabine 1H DMSO fe0511kjb5.esp

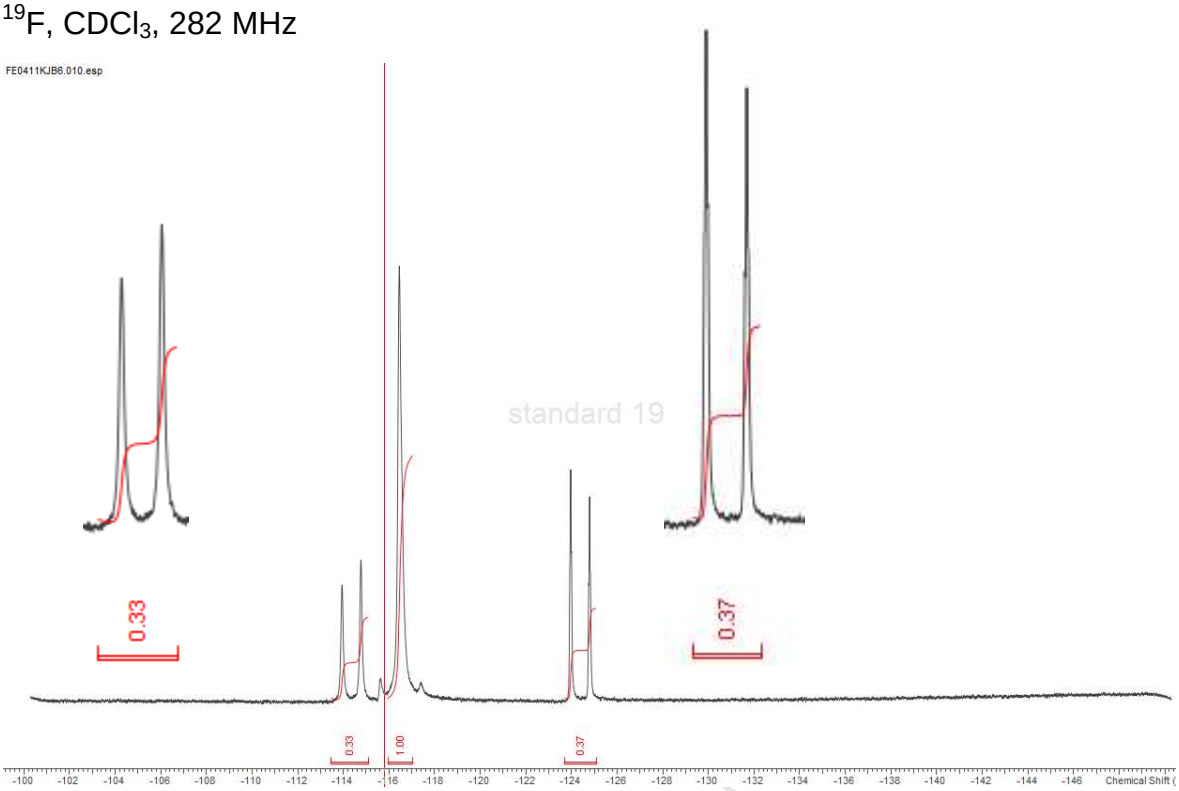


Commercial Sample, β anomer only, ^{1}H NMR, DMSO, 400 MHz

Desktop.070.esp



^{19}F, $CDCl_3$, 282 MHz



Commercial Sample, β anomer only, ^{1}F NMR, DMSO, 376 MHz