SUPPORTING INFORMATION

Synthesis, Binding and Antiviral Properties of Potent Core-Extended Naphthalene Diimide Targeting the HIV-1 Long Terminal Repeat Promoter G-quadruplexes

Rosalba Perrone,[†,§] Filippo Doria,[‡,§] Elena Butovskaya,[†,§] Ilaria Frasson,[†] Silvia Botti,[‡] Matteo Scalabrin,[†] Sara Lago,[†] Vincenzo Grande,[‡] Matteo Nadai,[†] Mauro Freccero,[*,‡] and Sara N. Richter[*,†]

[†]Department of Molecular Medicine, University of Padua, via Gabelli 63, 35121 Padua, Italy

[‡]Department of Chemistry, University of Pavia, V.le Taramelli 10, 27100 Pavia, Italy

Table of content

Table S1. Oligonucleotides used in this study

Application	Name	Sequence 5'-3'
FRET assay	LTR-III	*FAM*-TGGGAGGCGTGGCCTGGGCGGGACTGGGGT-*TAMRA*
	LTR-IV	*FAM*-TGGGCGGGACTGGGGAGTGGT-*TAMRA*
	F21T	*FAM*-GGGTTAGGGTTAGGGTTAGGG-*TAMRA*
	dsDNA	*FAM*-CTATAGCGCGCTATAG-*TAMRA*
CD and MS assays	LTR-III	GGGAGGCGTGGCCTGGGCGGGACTGGGG
	LTRIV	TGGGCGGGACTGGGGAGTGGT
	hTel	GGGTTAGGGTTAGGGTTAGGG
	c-kit2	CGGGCGGGCGCGAGGGAGGGG
	c-myc	TGGGGAGGGTGGGGAGGGTGGGGAAGG
Taq polymerase stop assay	*Taq* Primer	GGCAAAAAGCAGCTGCTTATATGCAG
	No G4	TTGTCGTTAAAGTCTGACTGCGAGCTCTCAGATCCTGCATATAAGCAGCTGCTTTTTGCC
	hTel	TTTTTGGGTTAGGGTTAGGGTTAGGGTTTTTCTGCATATAAGCAGCTGCTTTTTGCC
	LTR III	TTTTTGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGTTTTTCTGCATATAAGCAGCTGCTTTTTGCC
	LTR IV	TTTTTGGGCGGGACTGGGGAGTGGTTTTTCTGCATATAAGCAGCTGCTTTTTGCC

FAM: 6-carboxyfluorescein,*TAMRA*: 6-carboxy-tetramethylrhodamine

Table S2. C-exNDIs ΔT_m values obtained in the FRET melting assay. Equimolar amounts of the compounds were added to each FAM (6-carboxyfluorescein) 5'-end- and TAMRA (6-carboxy-tetramethylrhodamine) 3'-end-labelled oligonucleotide (0.25 μM) folded in the lithium cacodylate buffer supplemented with potassium (100 mM).

c-exNDIs	LTR-III	LTR-IV	hTel	dsDNA
2	10.5 ± 0.1	11.7 ± 0.6	6.0 ± 0.1	1.5 ± 0.1
3	7.0 ± 0.7	10.3 ± 1.1	3.5 ± 0.7	1.6 ± 0.4
4	7.0 ± 0.7	12.0 ± 1.7	5.3 ± 1.1	1.6 ± 0.5
5	2.5 ± 0.1	5.0 ± 0.1	2.0 ± 0.1	1.1 ± 0.6
6	5.0 ± 0.7	6.3 ± 0.6	4.2 ± 0.3	1.8 ± 0.6
7	3.7 ± 1.0	4.7 ± 0.6	3.0 ± 0.1	1.5 ± 0.1
8	1.5 ± 0.1	3.0 ± 0.1	1.5 ± 0.7	1.1 ± 0.6
9	1.2 ± 1.1	3.0 ± 1.0	0.5 ± 0.7	0.5 ± 0.1
10	12.4 ± 0.1	16.0 ± 1.7	8.5 ± 0.6	4.1 ± 0.6
11	5.5 ± 0.1	11.7 ± 1.2	3.8 ± 0.3	1.5 ± 0.6

Table S2. Cartesian coordinates of the **c-exNDI** core (Scheme 1, R_2=CH_3, R_3=Y=H) optimised at B3LYP/6-31+G(d,p) level of theory, in gas phase, using Gaussian 09, Revision B.01 software package.

Center Number	Atomic Number	Atomic Type	Coordinates (Angstroms) X	Y	Z
1	6	0	-0.217963	-2.890097	-0.000065
2	6	0	-0.210573	-1.429193	-0.000147
3	6	0	-1.450603	-0.710875	-0.000178
4	6	0	-2.679294	-1.396967	-0.000121
5	6	0	-2.701691	-2.872757	-0.000130
6	6	0	1.002582	-0.725637	-0.000164
7	6	0	-1.450606	0.710875	-0.000252
8	6	0	-0.210578	1.429200	-0.000288
9	6	0	1.002583	0.725650	-0.000206
10	6	0	-0.217978	2.890104	-0.000228
11	6	0	-2.701709	2.872749	-0.000221
12	6	0	-2.679301	1.396959	-0.000203
13	6	0	-3.893181	0.694469	-0.000050
14	6	0	-3.893178	-0.694485	-0.000060
15	1	0	-4.818044	1.260047	0.000021
16	8	0	-3.735670	3.532525	0.000780
17	8	0	0.829008	3.567832	0.000517
18	8	0	-3.735646	-3.532542	-0.000082
19	8	0	0.829030	-3.567819	-0.000101
20	7	0	-1.458036	-3.528911	0.000183
21	7	0	-1.458059	3.528909	-0.000463
22	6	0	-1.503535	-4.998420	0.000408
23	6	0	-1.503570	4.998420	-0.000160
24	1	0	-2.044109	5.346589	-0.882700
25	1	0	-0.481647	5.367348	-0.004398
26	1	0	-2.042154	-5.346457	0.884197
27	1	0	-0.481601	-5.367338	0.002629
28	1	0	-2.036402	5.346903	0.887019
29	1	0	-2.038288	-5.347048	-0.885535
30	1	0	2.141728	-2.381023	0.000014
31	1	0	2.141735	2.381033	-0.000273
32	6	0	3.440486	0.703687	-0.000125
33	6	0	3.440486	-0.703676	-0.000075
34	6	0	4.650709	1.403085	-0.000087
35	1	0	4.639964	2.489027	-0.000115
36	6	0	4.650706	-1.403079	-0.000011
37	1	0	4.639958	-2.489020	-0.000004
38	6	0	5.856171	-0.699850	0.000007
39	1	0	6.794033	-1.245446	0.000026
40	6	0	5.856173	0.699852	-0.000019
41	1	0	6.794035	1.245449	0.000036
42	1	0	-4.818040	-1.260064	-0.000019
43	7	0	2.211187	-1.356421	-0.000088
44	7	0	2.211189	1.356433	-0.000195

Gaussian 09, Revision B.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A.

Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2010.

Figure S1. Size comparison between a) the **NDI** and b) **c-exNDI** cores. The size of the NDI, as well as the G4 size [reported in c) and d)], have been evaluated from the crystal structure of the intramolecular human telomeric DNA G-quadruplex bound by the di-substituted NDI BMSG-SH-3 (3SC8.pdb, from Reference: Collie, G.W., Promontorio, R., Hampel, S.M., Micco, M., Neidle, S., Parkinson, G.N. Structural basis for telomeric G-quadruplex targeting by naphthalene diimide ligands. *J. Am. Chem. Soc.* **2012**, *134*, 2723-2731). The size of the **c-exNDI** core was evaluated from DFT calculations in gas phase, at B3LYP/6-31+G(d,p) level of theory. Overlap of the c) **NDI** and d) **c-exNDI** cores to the intramolecular human telomeric DNA G-quartet, for an estimation of the ligand-quartet overlap extension.

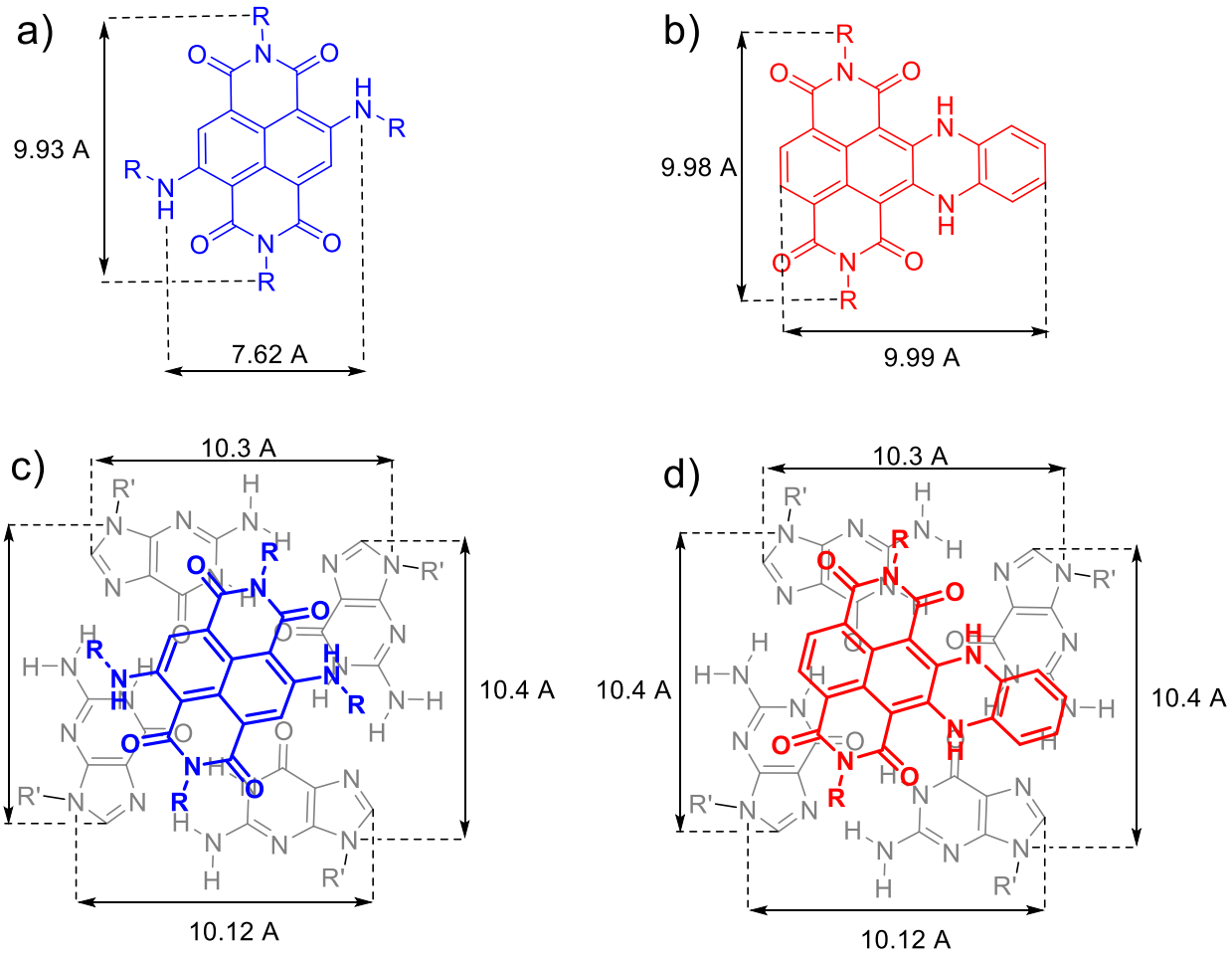
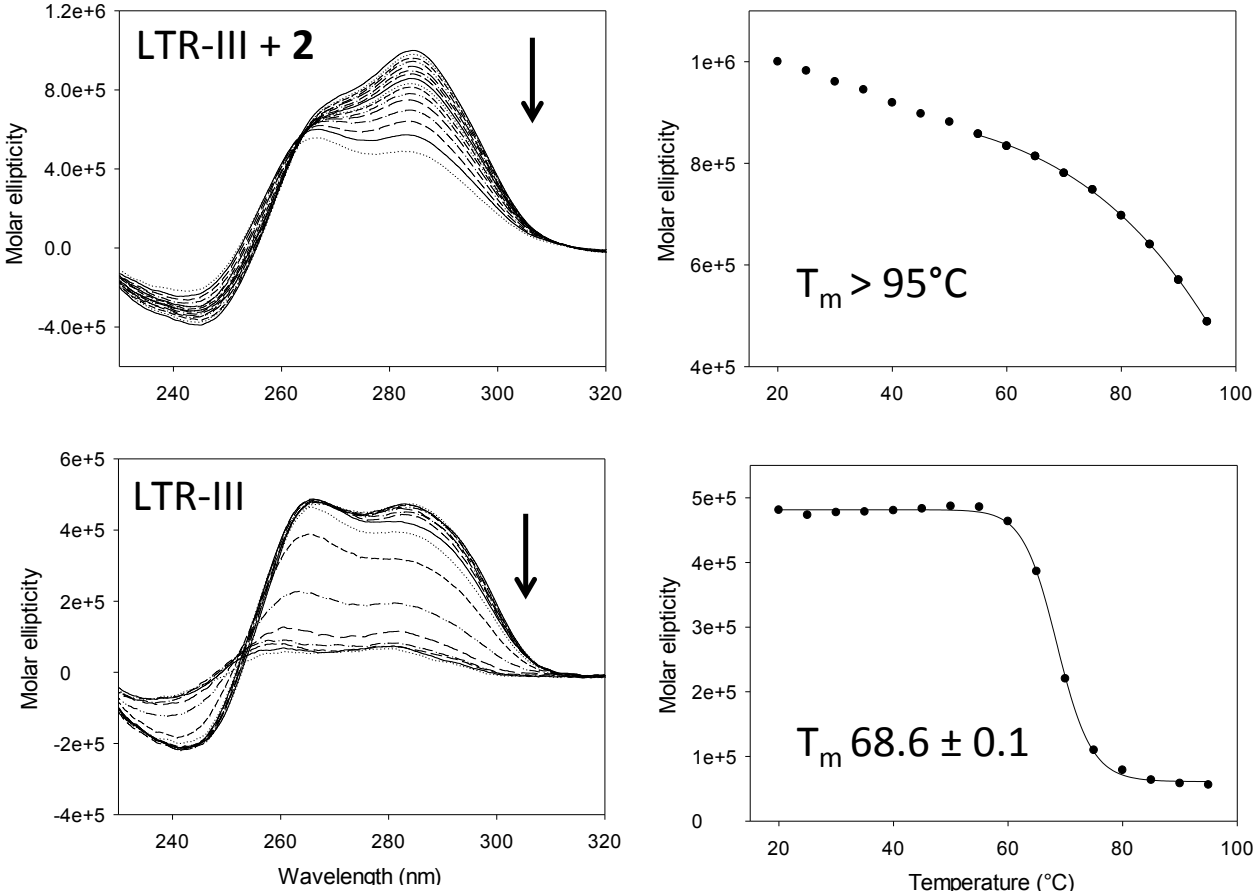
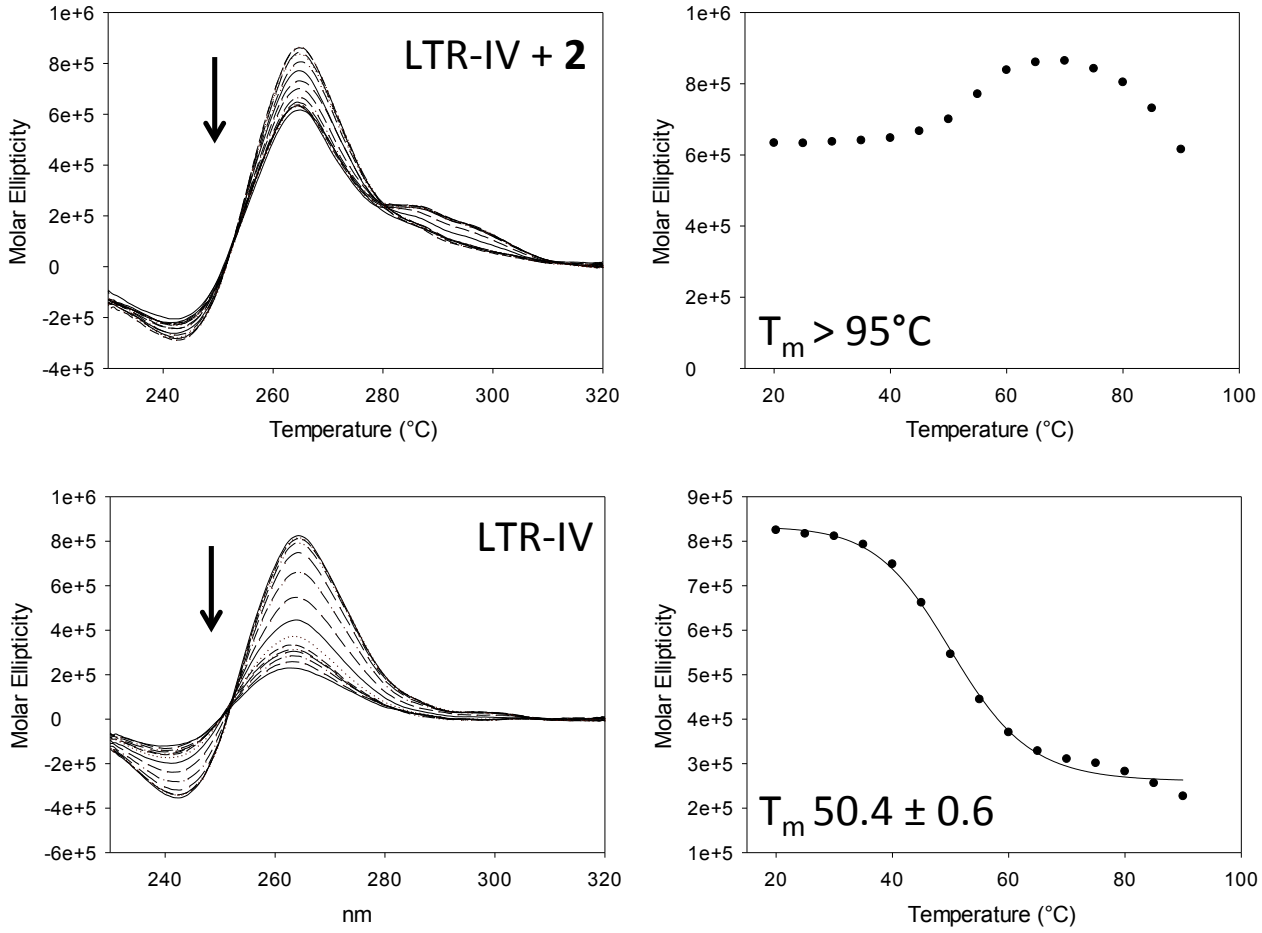


Figure S2. CD analysis of LTR-III (A), LTR-IV (B) and hTel (C) G4 oligonucleotides in the absence (lower panels) and in the presence of compound **2** (upper panels). The left panels show CD spectra variation in function of the temperature; arrows indicate the spectral change from low to high temperatures. The right panels show the molar ellipticity at the peak wavelength as a function of the temperature.

A



B


LTR-IV + 2
Molar Ellipticity
Temperature (°C)
T_m > 95°C
LTR-IV
nm
T_m 50.4 ± 0.6

C

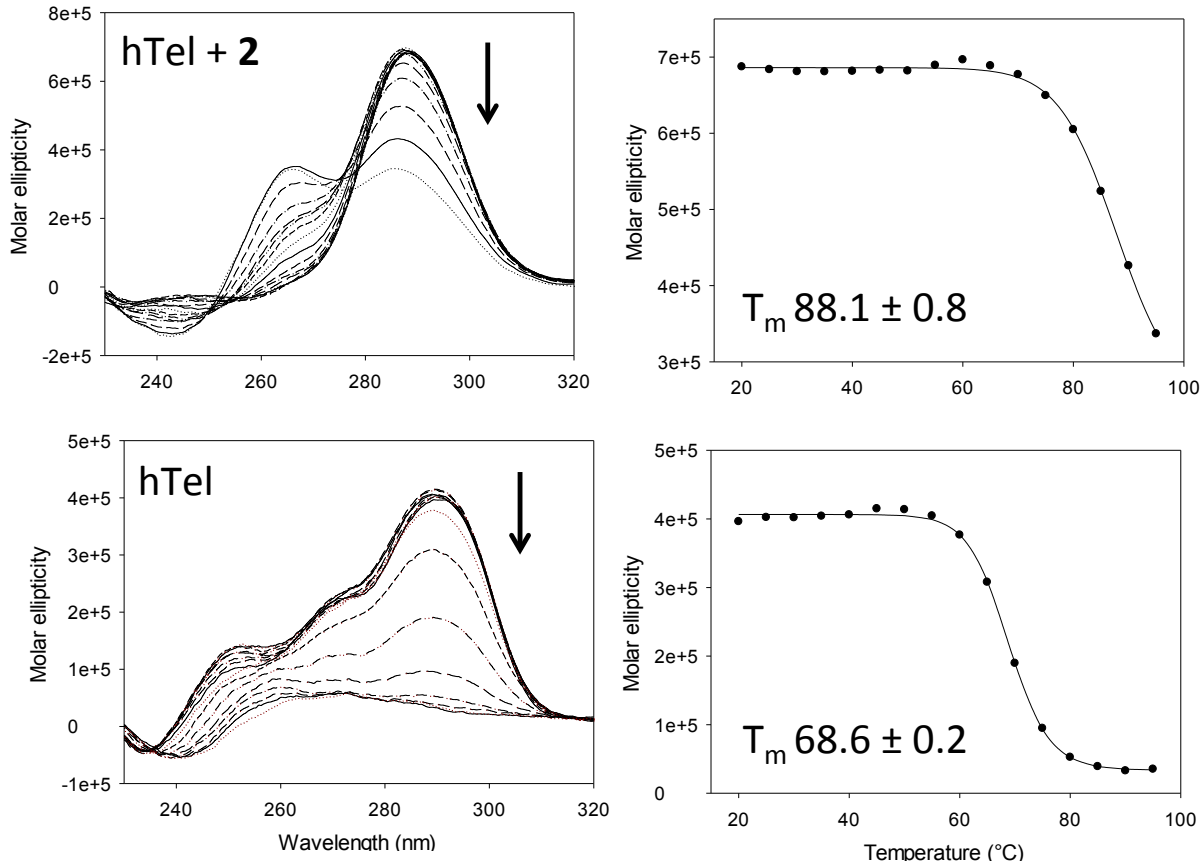

hTel + 2
Molar ellipticity
8e+5
6e+5
4e+5
2e+5
0
-2e+5
240
260
280
300
320
7e+5
6e+5
5e+5
4e+5
3e+5
T_m 88.1 ± 0.8
20
40
60
80
100
hTel
5e+5
4e+5
3e+5
2e+5
1e+5
0
-1e+5
Wavelength (nm)
T_m 68.6 ± 0.2
Temperature (°C)

Figure S3. CD analysis of G4-folded oligonucleotides in MS buffer conditions.

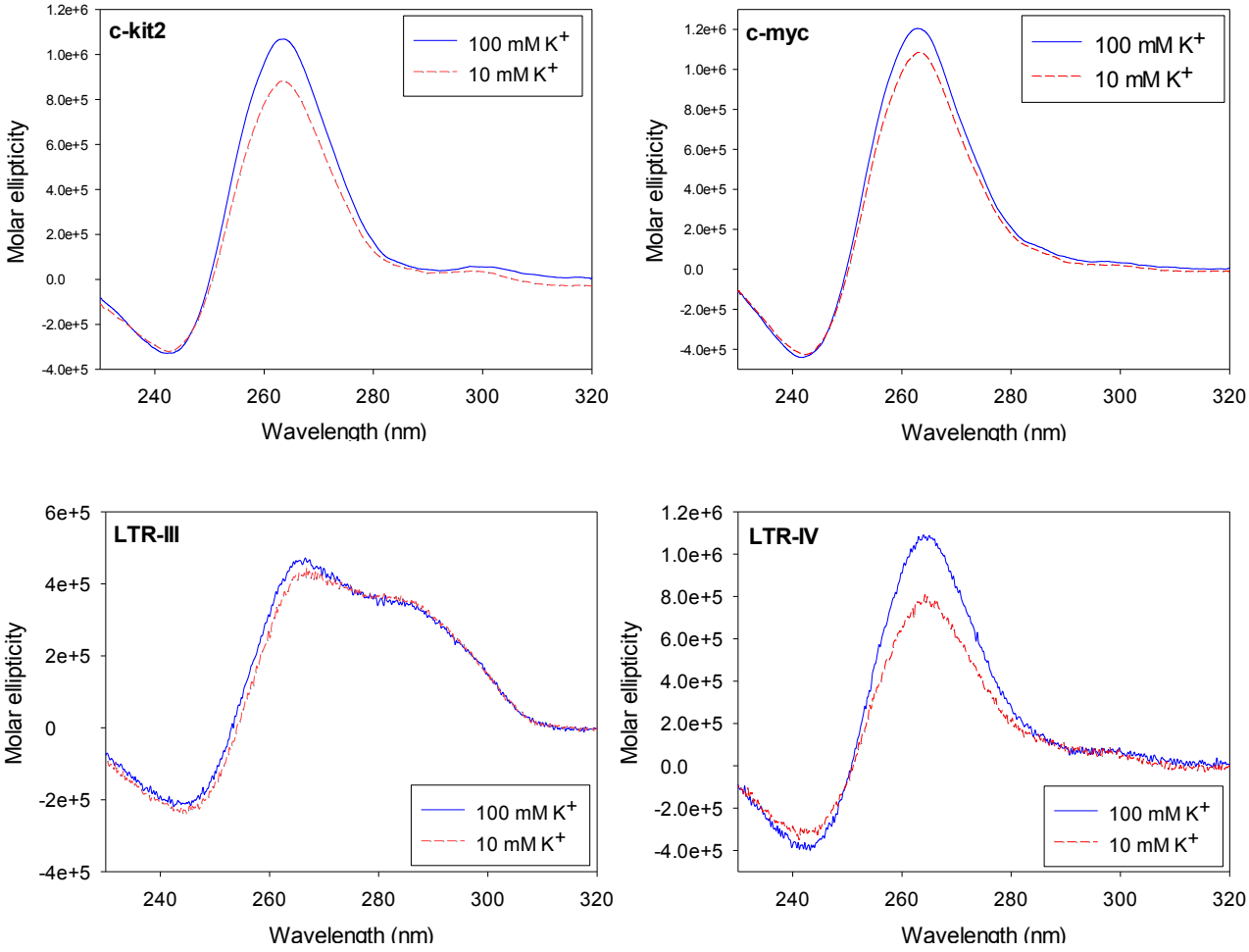


Figure S4. ESI/MS competition analysis. Equimolar amounts of LTR-III/hTel (A), LTR-IV/hTEl (B), LTR-III/LTR-IV (C) were incubated with 2-fold excess of **2** (left panels) or **4** (right panels) and analysed by ESI/MS. The identity of m/z signals is shown above each signal. RI stands for relative intensity.

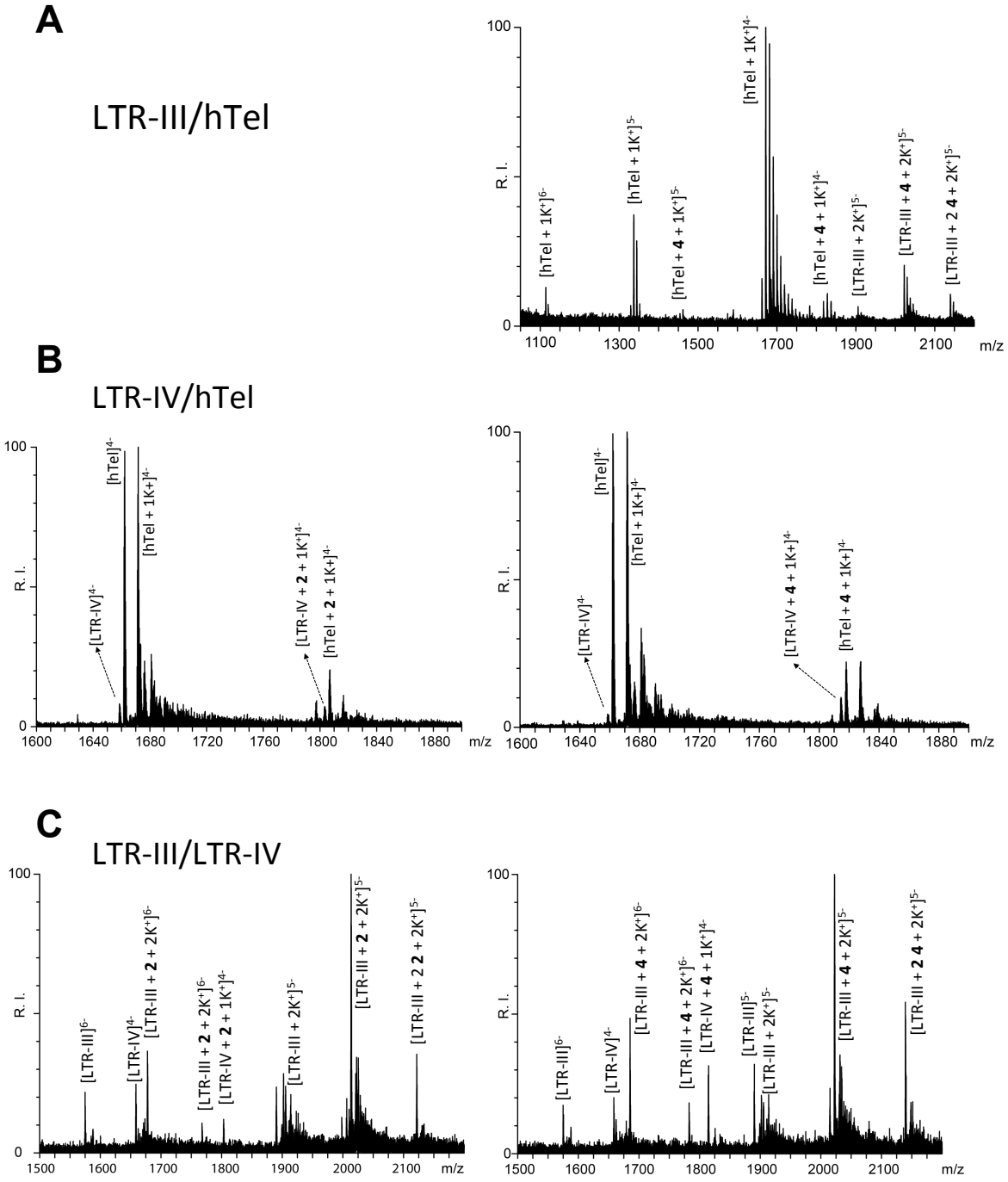


Figure S5. ESI/MS analysis of **c-exNDI 2** binding stoichiometry at saturating concentration (10-fold molar excess) towards the G4 structures. Representative spectrum of **c-exNDI 2** and LTR-III G4.

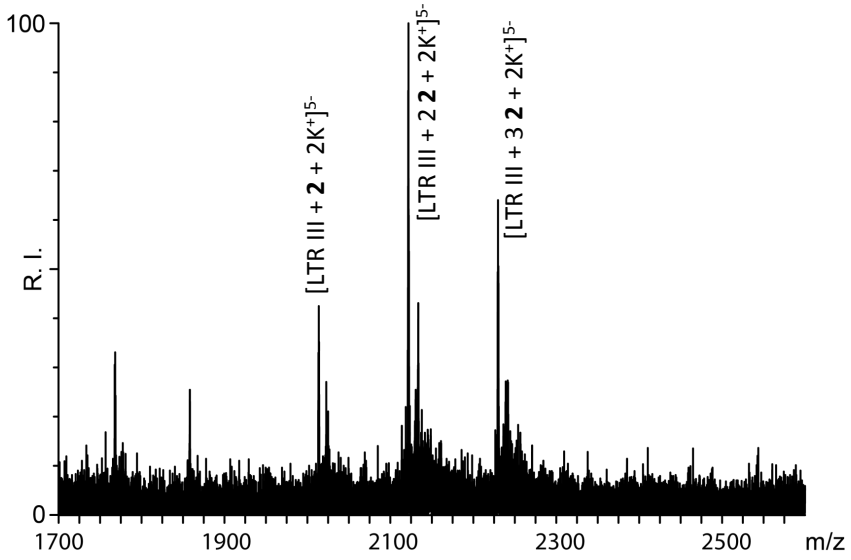
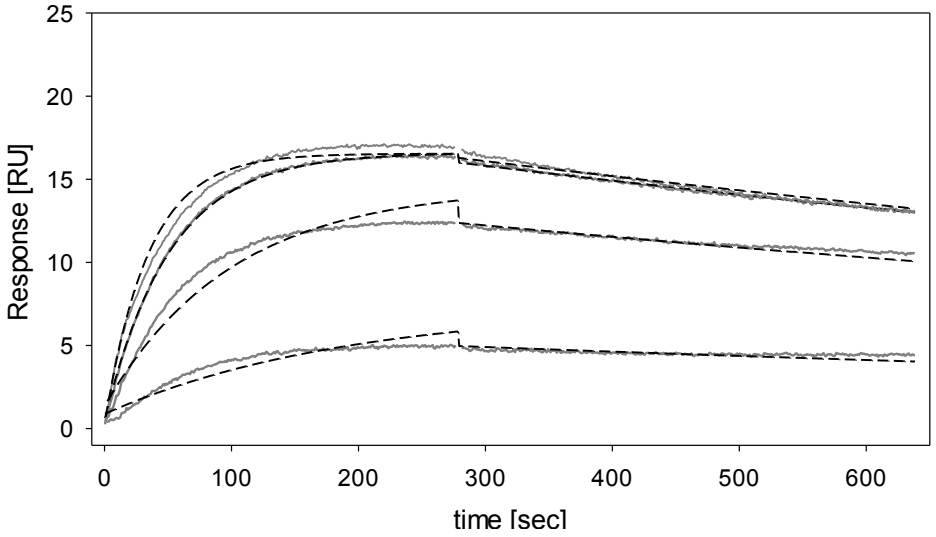


Figure S6. SPR analysis of **2** binding towards LTR-IV (upper panel) and hTel (lower panel) G4s. Biosensor data were collected at 25°C; the compound was injected at increasing concentrations (50, 100, 200, 300 nM). Sensograms are shown as grey lines. Dotted black lines represent local 1:1 binding fits obtained for a calculated R_{max} of 33.6 (LTR-IV) and 34.8 (hTel) RU (Response Units).



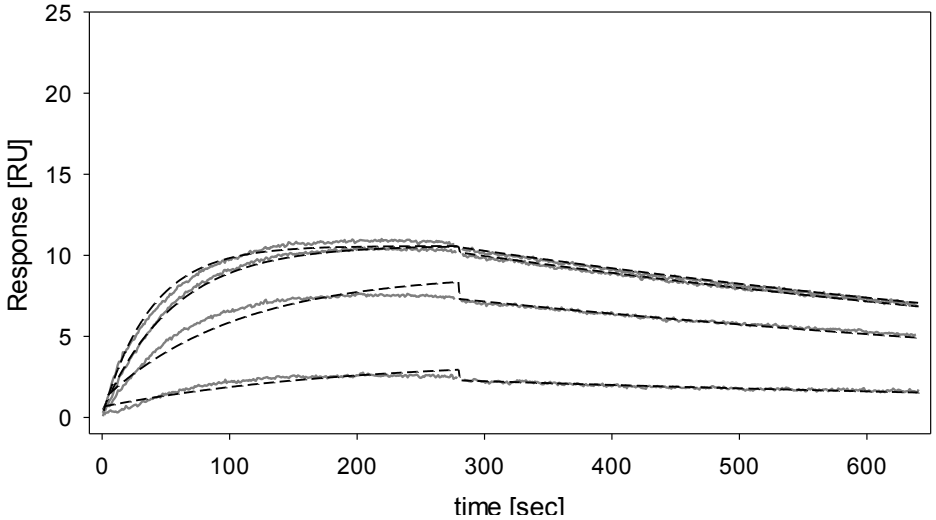
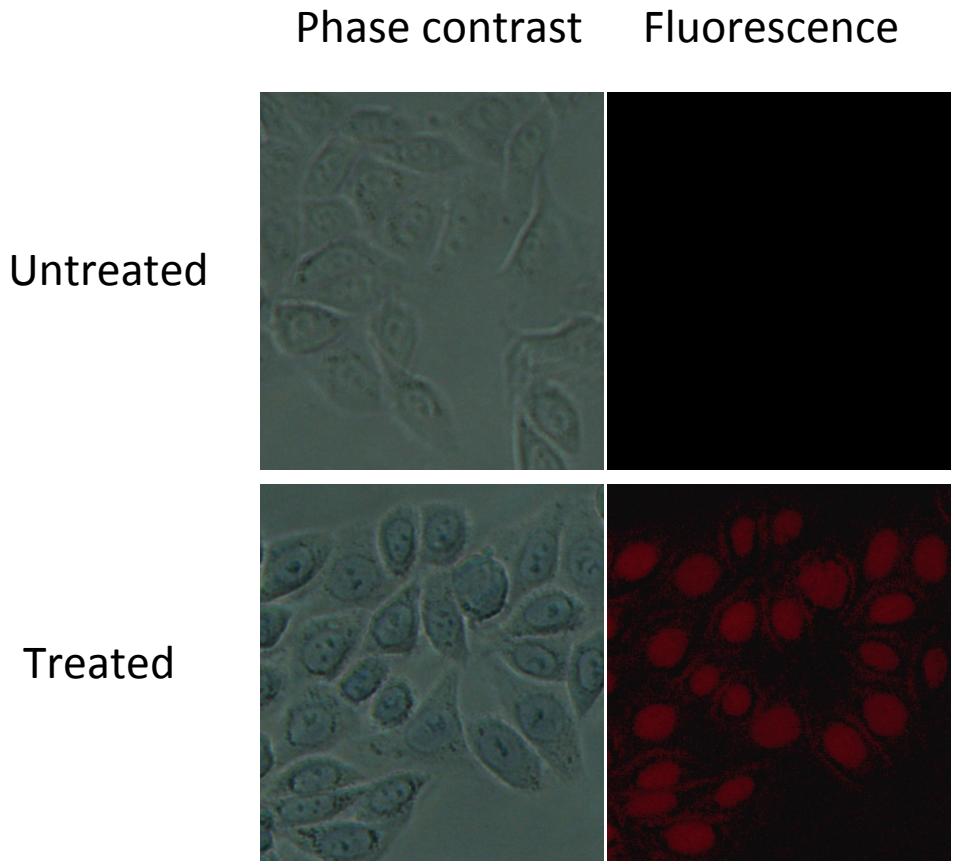
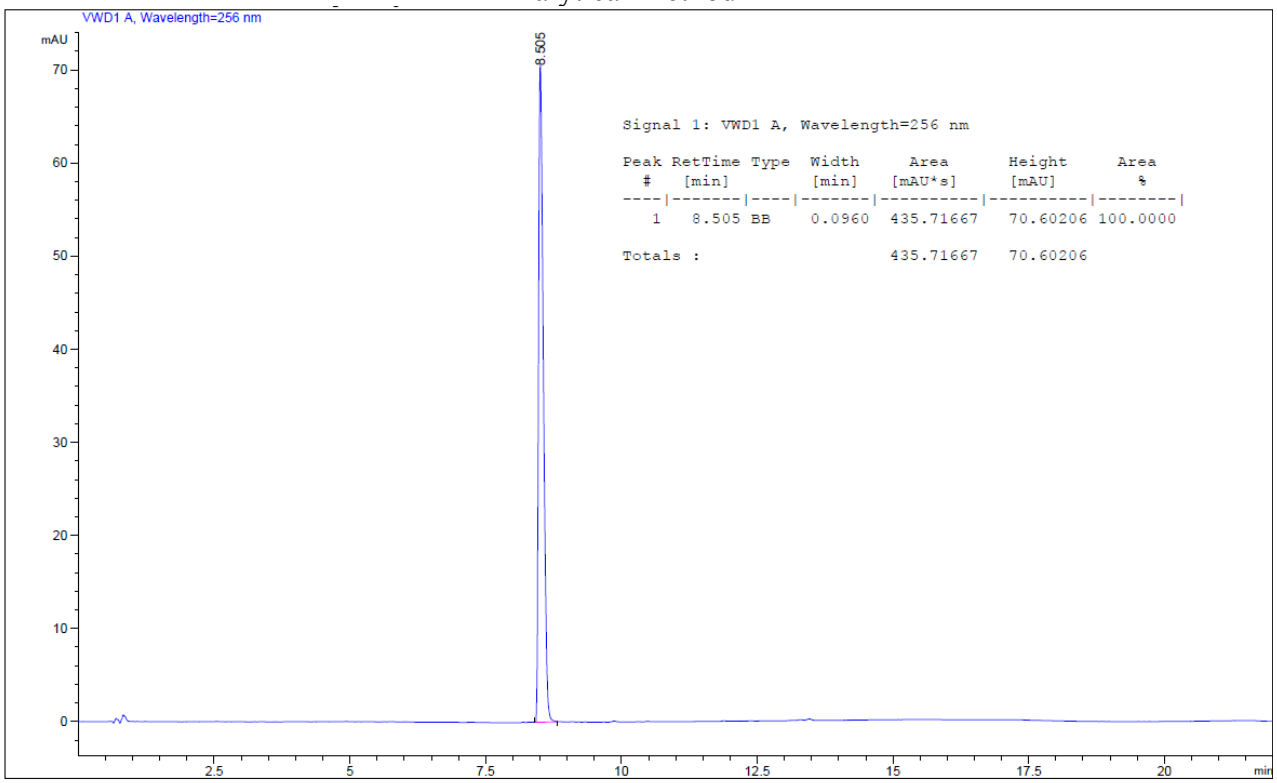


Figure S7. Assessment of **2** cell entry by fluorescence microscopy. TZM-bl cells were seeded (15000 cells/well) in μClear black 96 well plate (CELLSTAR®, Greiner Bio-One GmbH,Frickenhausen, Germany) and treated with **2** (12.5 μM). After 30 min, the medium was replaced with PBS 1X and cells were observed with the fluorescence microscope Leica DFC 420C (Leica Microsystems Srl, Milano, Italy) equipped with a mercury AC lamp (Leistungselektronik JENA GmbH, Jena, Germany) and an appropriate filter to visualize the red fluorescence of the compound with a 20X ocular objective. Images were acquired with Leica Application Suite V3.8 program (Leica Microsystems Srl, Milano, Italy).

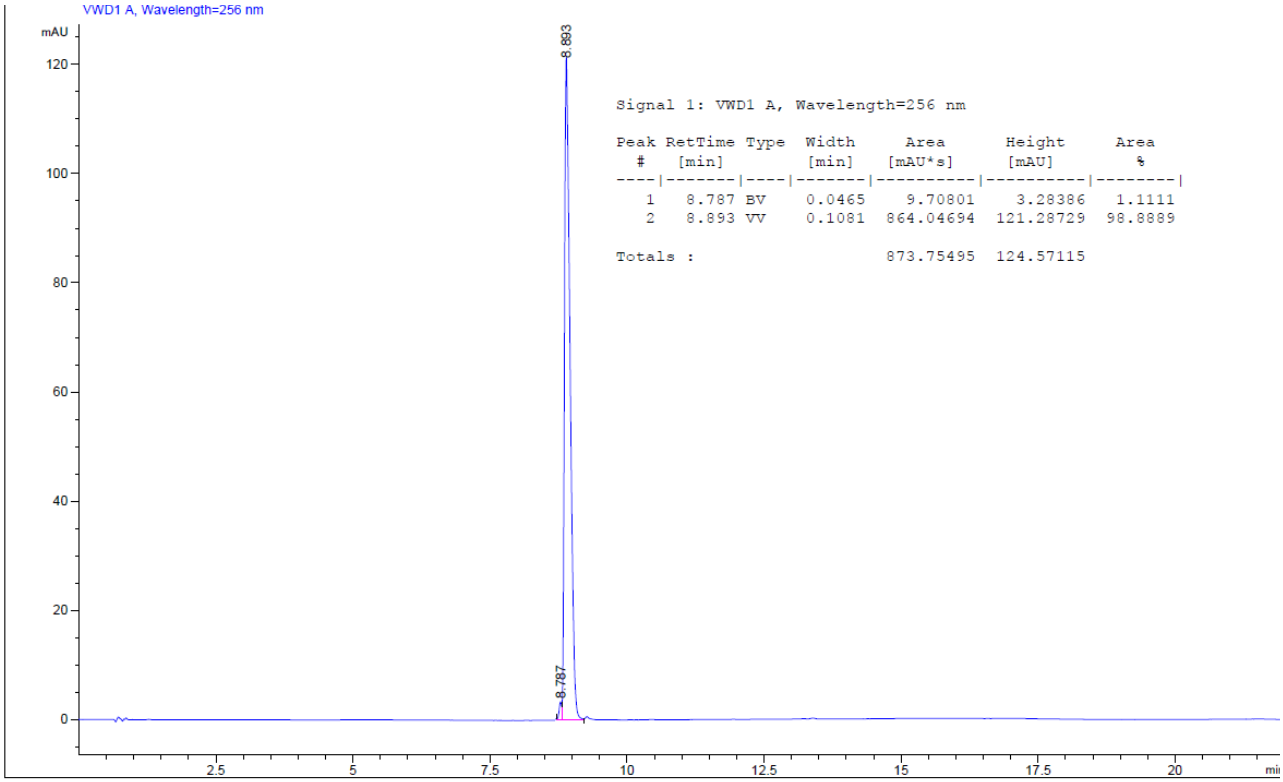


HPLC PURITY DATA:

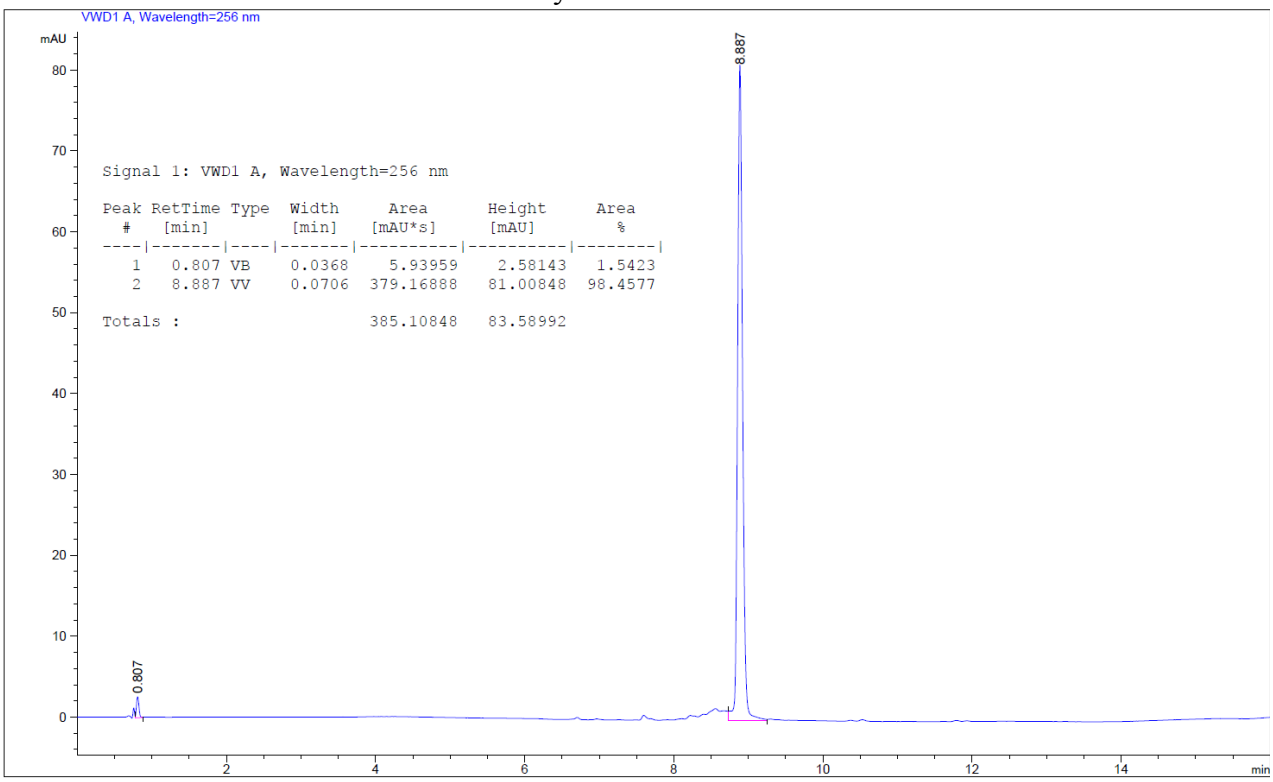
2·2HCl
Analytical Method-A



3·2HCl
Analytical Method-A



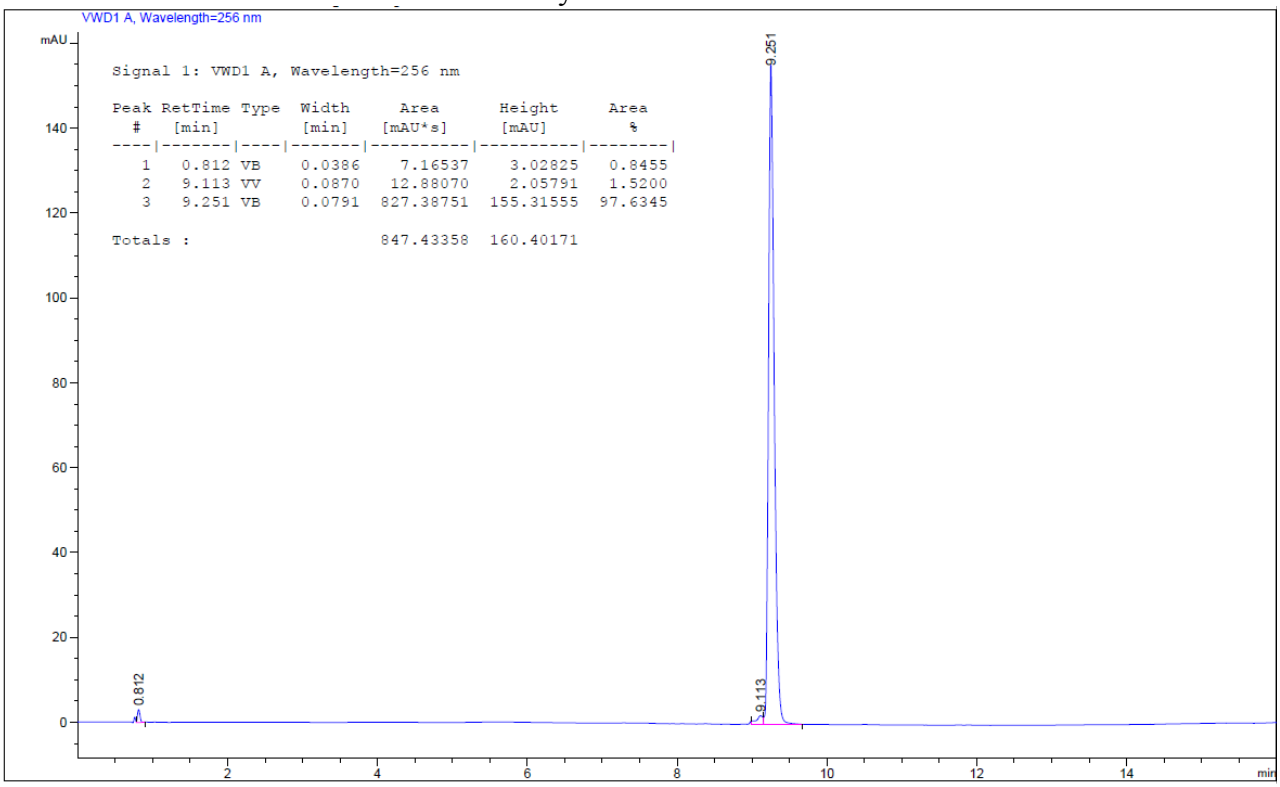
4·2HCl
Analytical Method-A



Signal 1: VWD1 A, Wavelength=256 nm

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	0.807	VB	0.0368	5.93959	2.58143	1.5423
2	8.887	VV	0.0706	379.16888	81.00848	98.4577
Totals :				385.10848	83.58992	

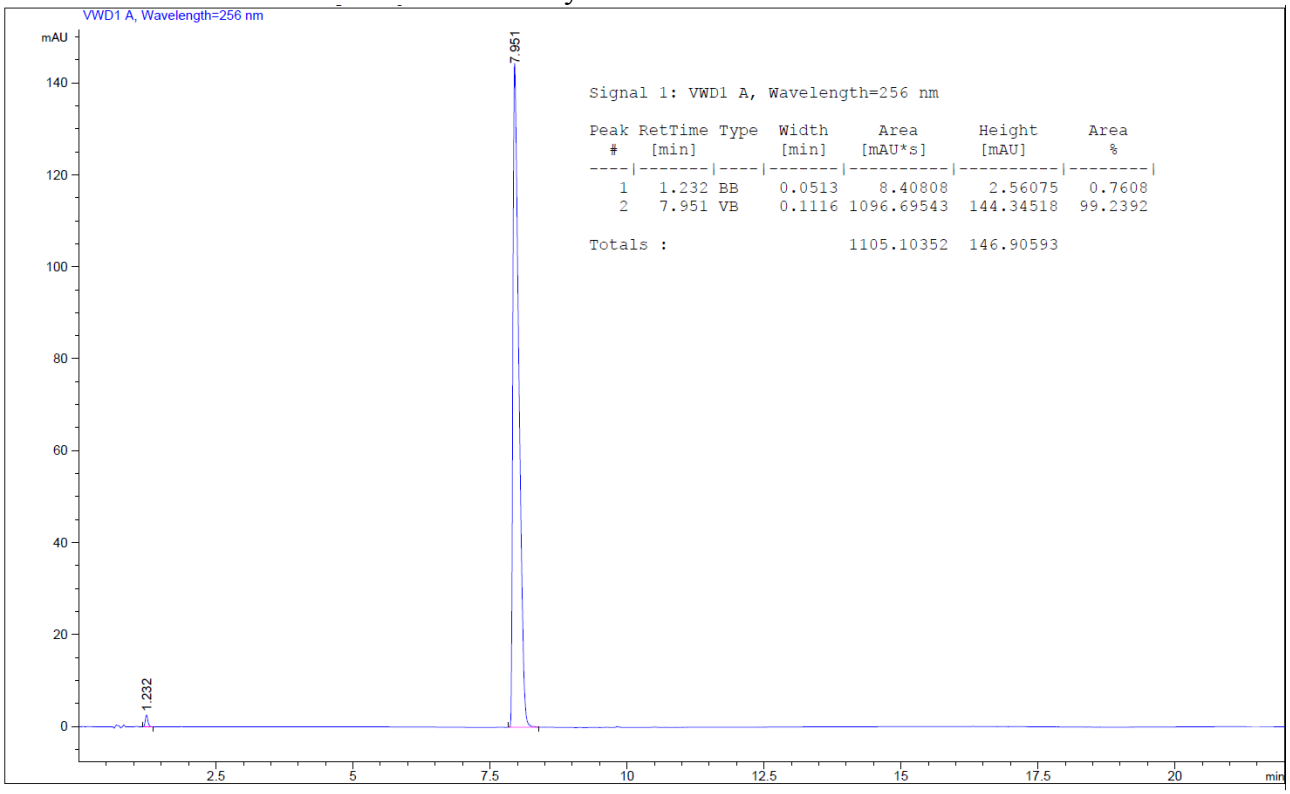
5·2HCl
Analytical Method-A



Signal 1: VWD1 A, Wavelength=256 nm

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	0.812	VB	0.0386	7.16537	3.02825	0.8455
2	9.113	VV	0.0870	12.88070	2.05791	1.5200
3	9.251	VB	0.0791	827.38751	155.31555	97.6345
Totals :				847.43358	160.40171	

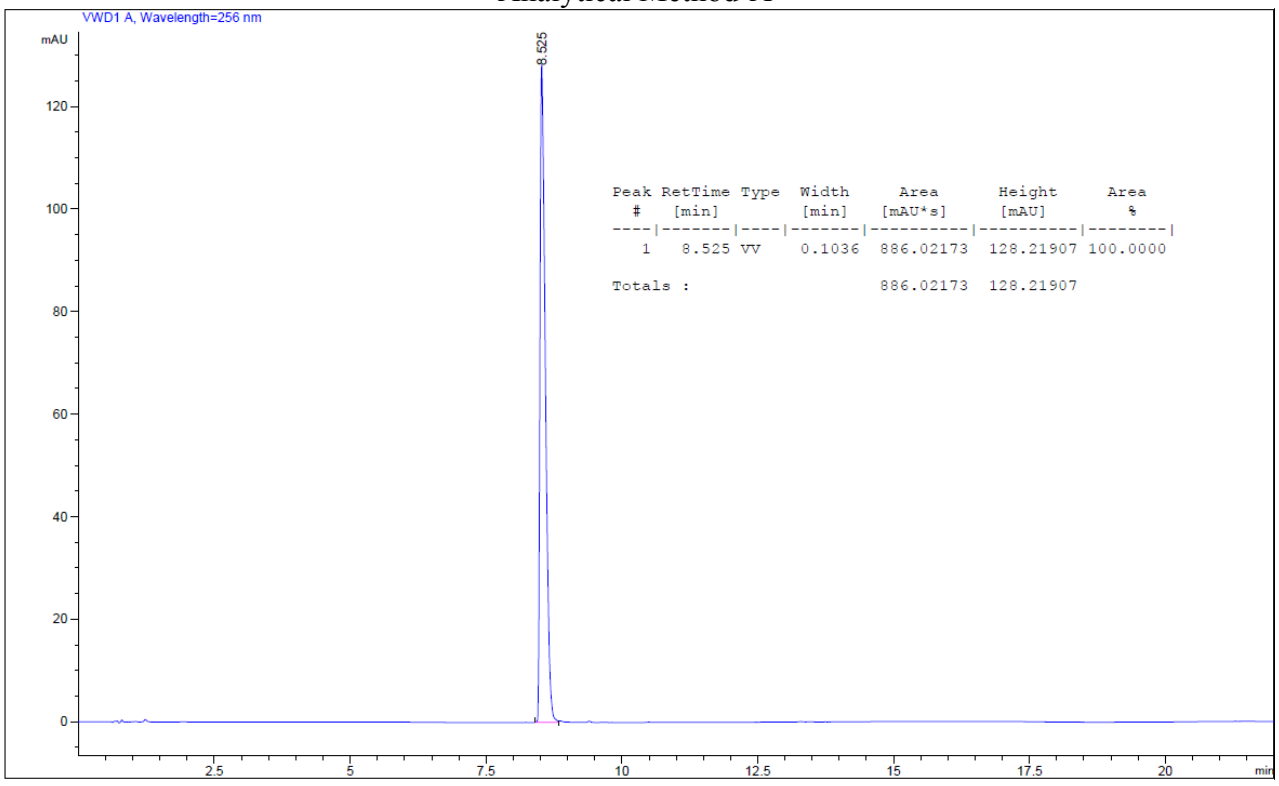
6·2HCl
Analytical Method-A



Signal 1: VWD1 A, Wavelength=256 nm

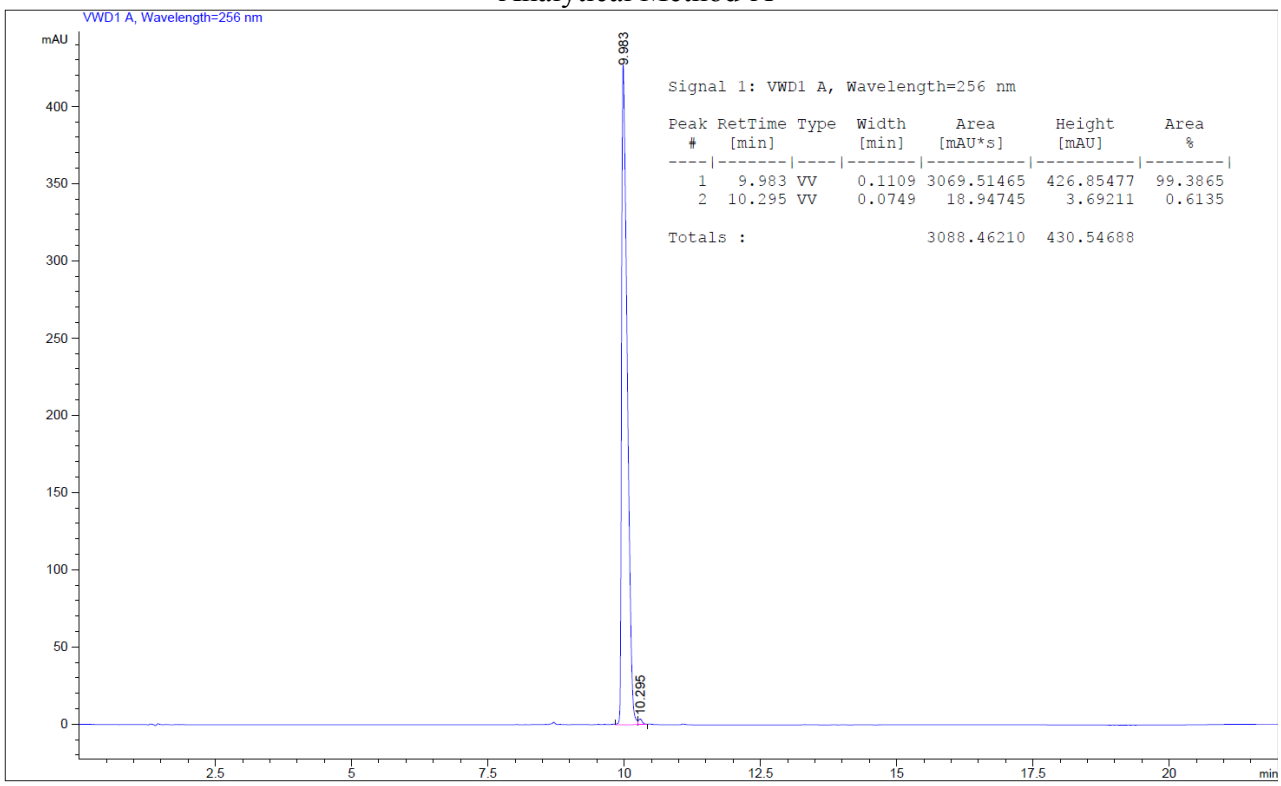
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	1.232	BB	0.0513	8.40808	2.56075	0.7608
2	7.951	VB	0.1116	1096.69543	144.34518	99.2392
Totals :				1105.10352	146.90593	

7·2HCl
Analytical Method-A



Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	8.525	VV	0.1036	886.02173	128.21907	100.0000
Totals :				886.02173	128.21907	

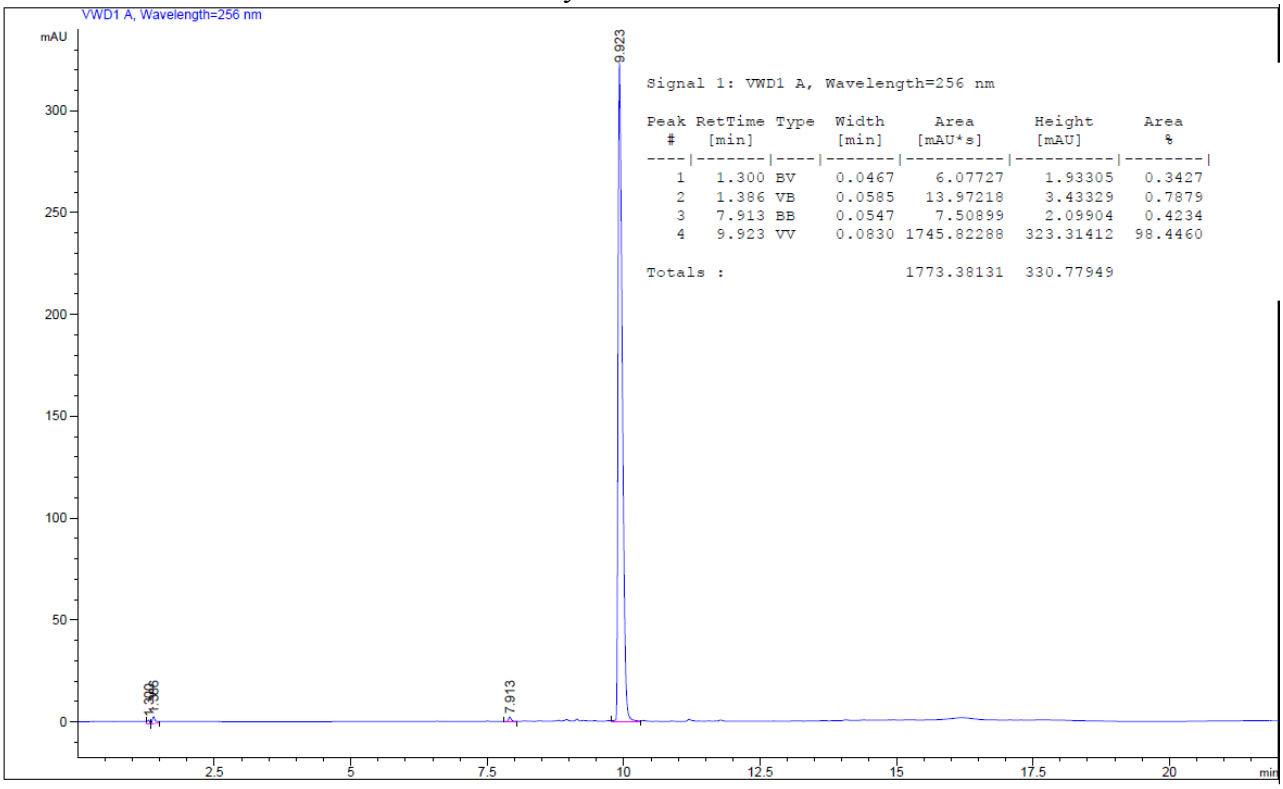
8·2HCl
Analytical Method-A



Signal 1: VWD1 A, Wavelength=256 nm

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	9.983	VV	0.1109	3069.51465	426.85477	99.3865
2	10.295	VV	0.0749	18.94745	3.69211	0.6135
Totals :				3088.46210	430.54688	

9·3HCl
Analytical Method-A

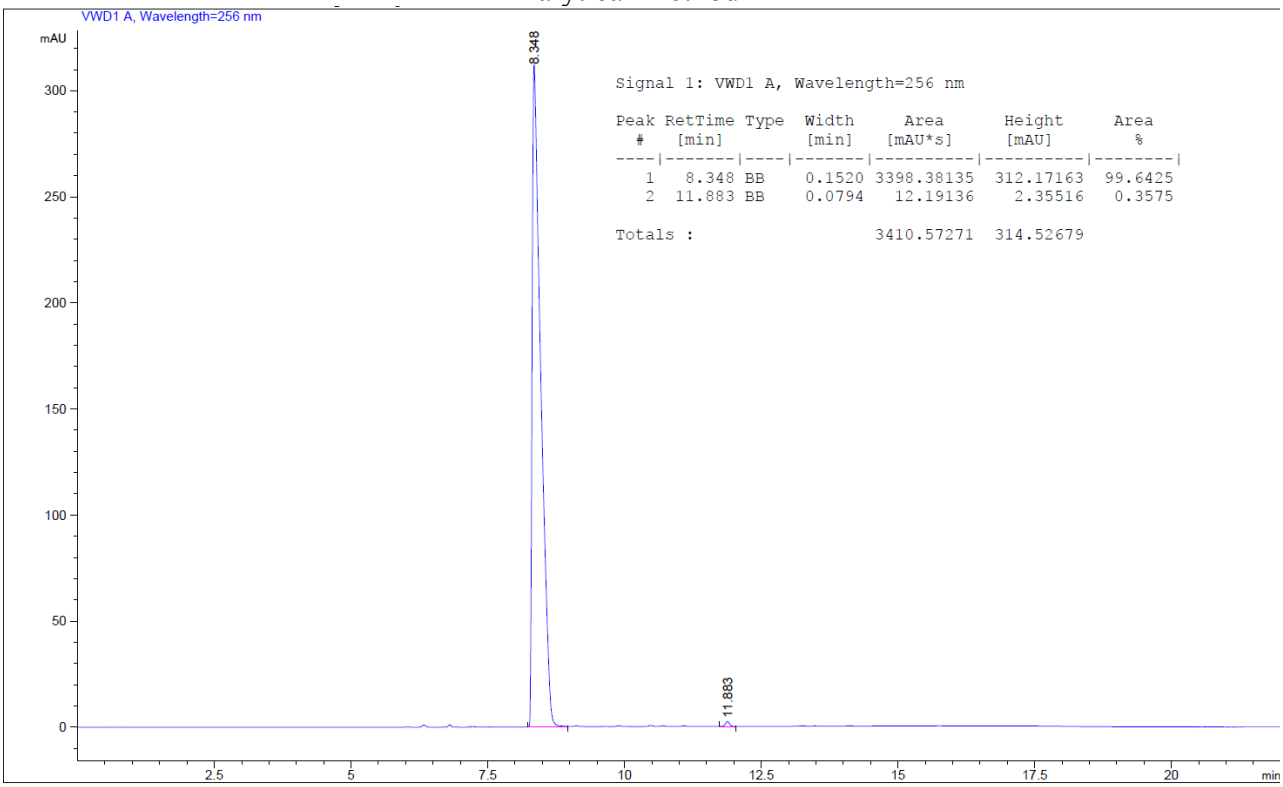


Signal 1: VWD1 A, Wavelength=256 nm

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	1.300	BV	0.0467	6.07727	1.93305	0.3427
2	1.386	VB	0.0585	13.97218	3.43329	0.7879
3	7.913	BB	0.0547	7.50899	2.09904	0.4234
4	9.923	VV	0.0830	1745.82288	323.31412	98.4460
Totals :				1773.38131	330.77949	

10

Analytical Method-A

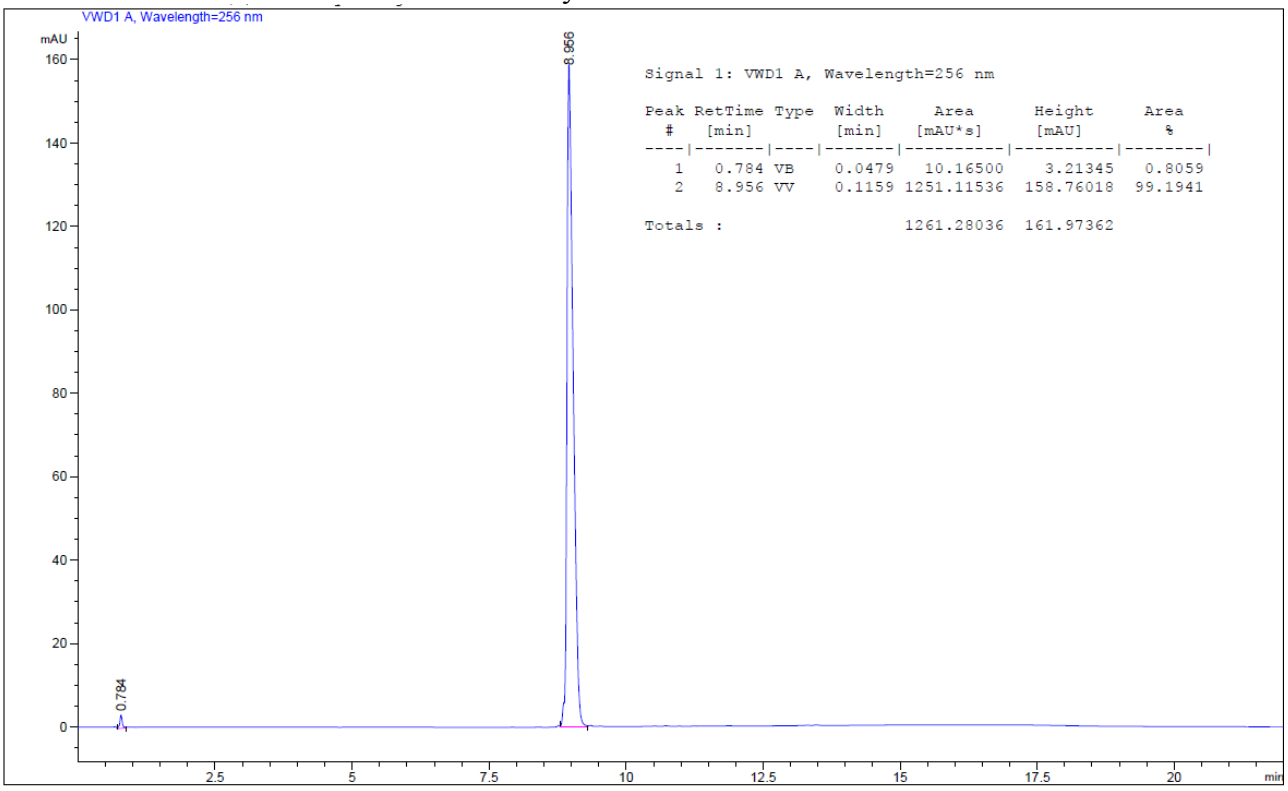


Signal 1: VWD1 A, Wavelength=256 nm

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	8.348	BB	0.1520	3398.38135	312.17163	99.6425
2	11.883	BB	0.0794	12.19136	2.35516	0.3575
Totals :				3410.57271	314.52679	

11

Analytical Method-A



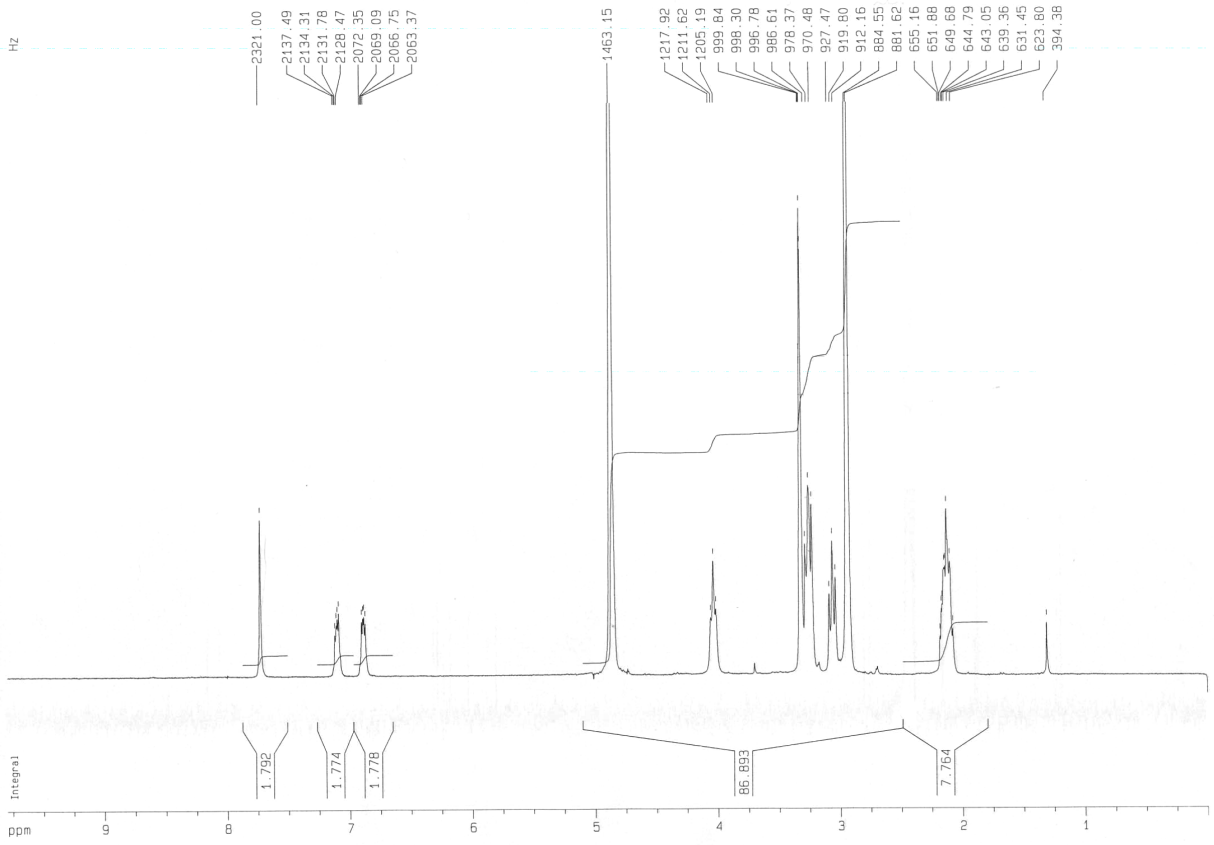
Signal 1: VWD1 A, Wavelength=256 nm

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	0.784	VB	0.0479	10.16500	3.21345	0.8059
2	8.956	VV	0.1159	1251.11536	158.76018	99.1941
Totals :				1261.28036	161.97362	

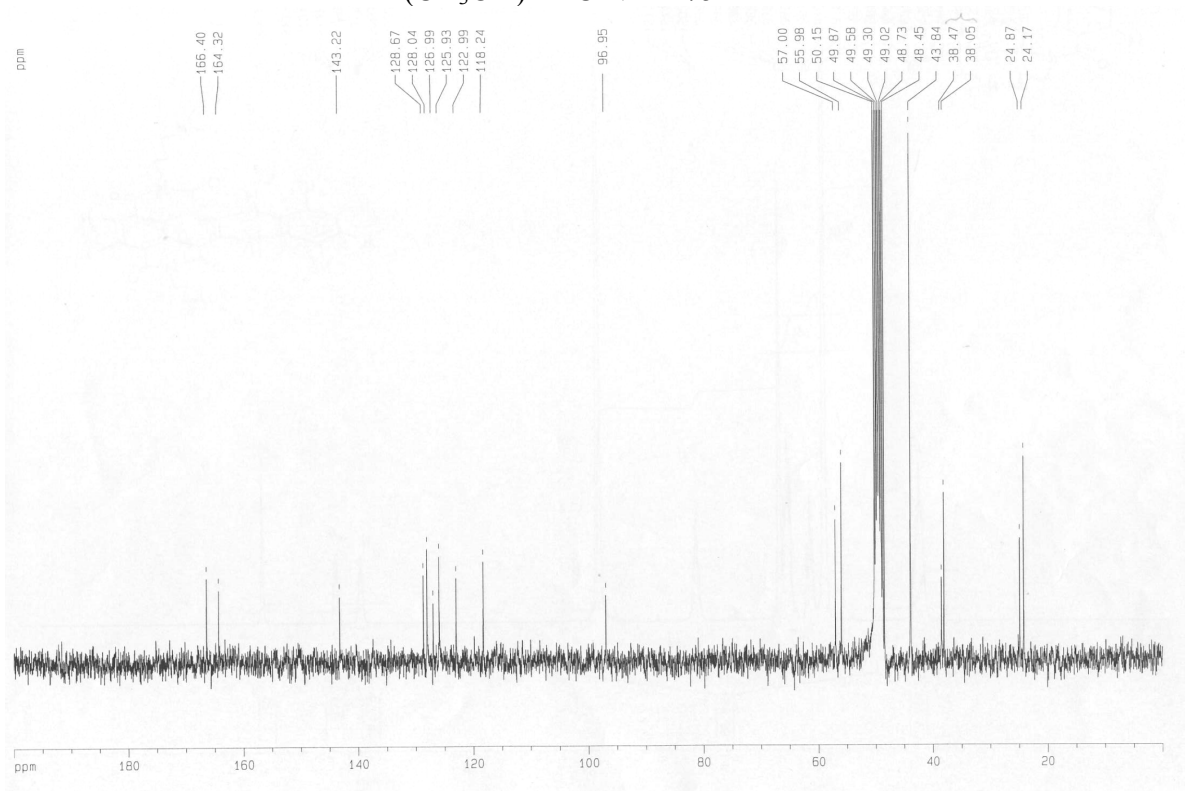
NMR Spectra:

2·2HCl

(CD_3OD) ^1H NMR 300MHz

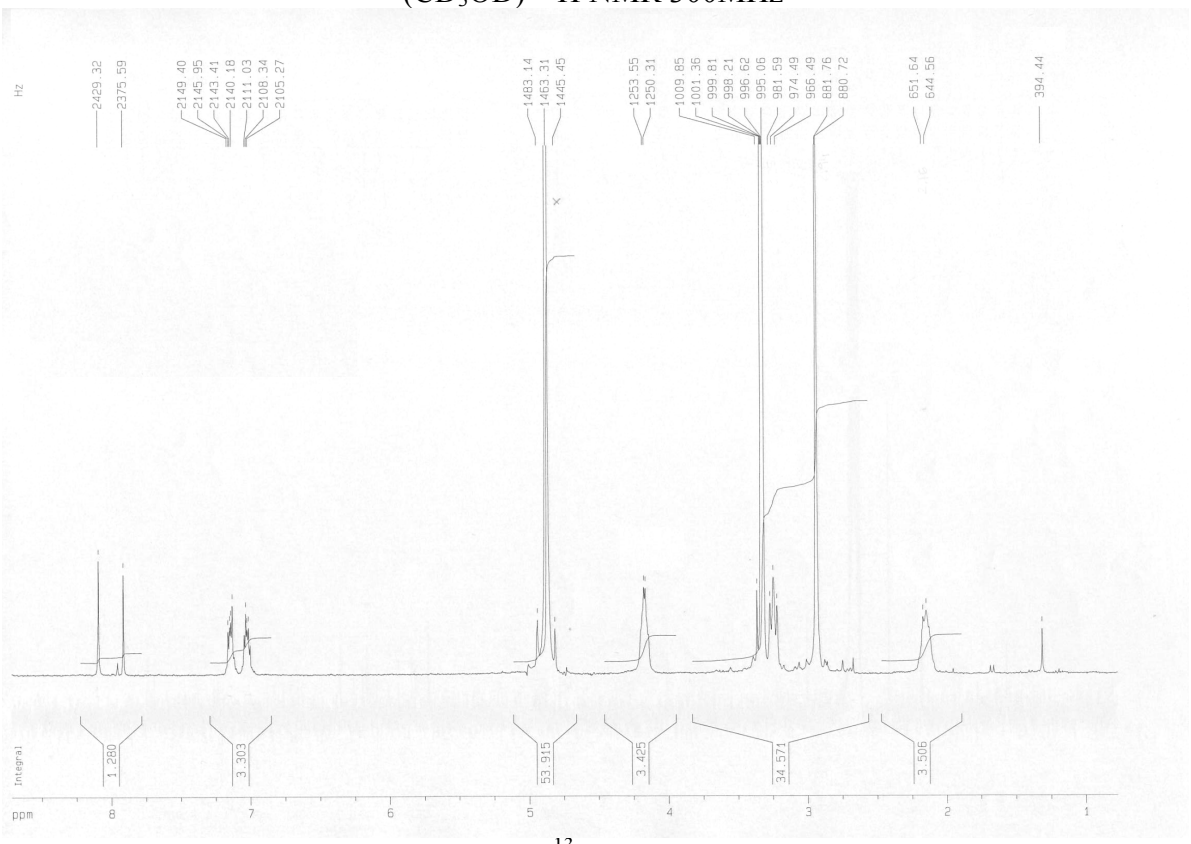


(CD_3OD) ^{13}C NMR 75MHz

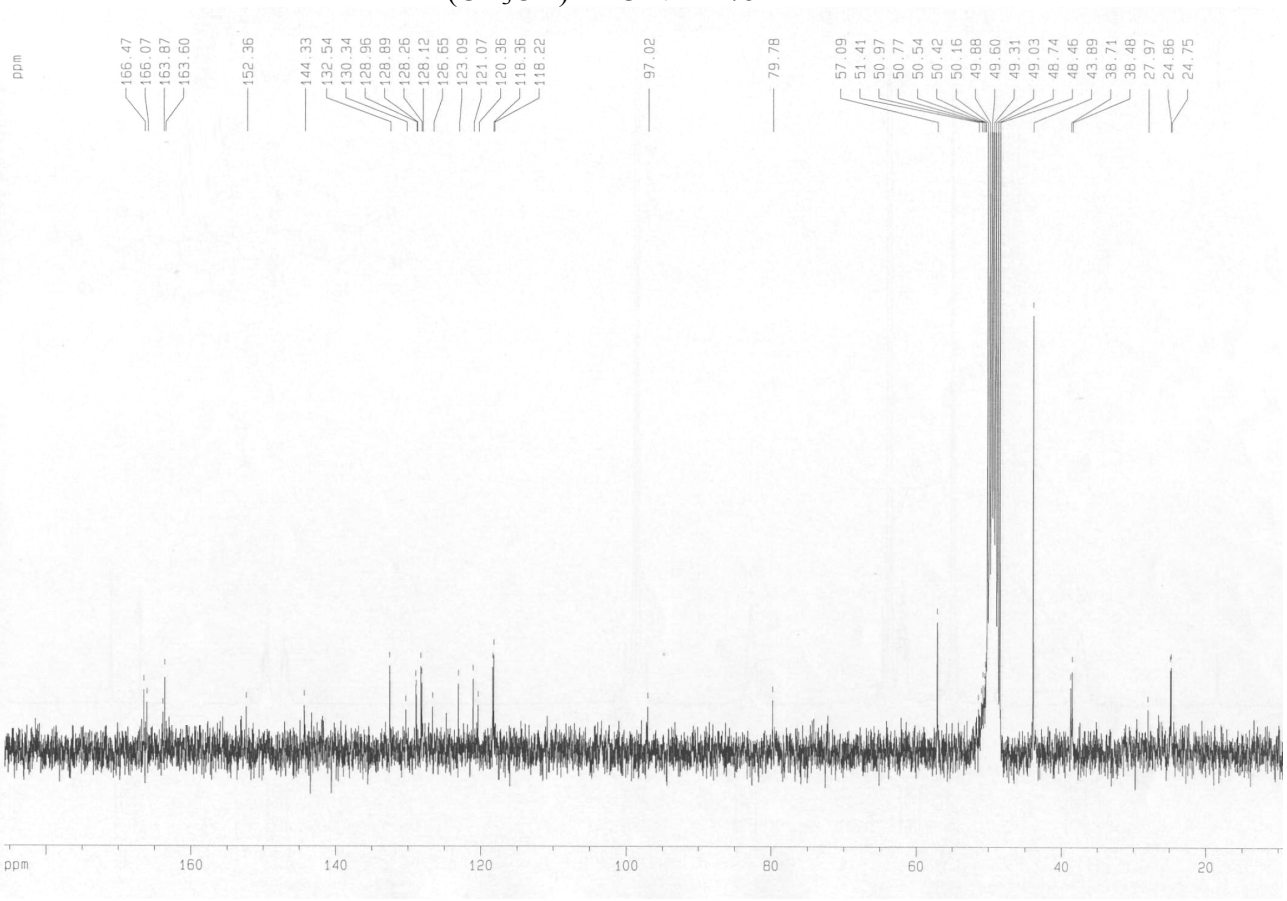


3·2HCl

(CD_3OD) ^1H NMR 300MHz

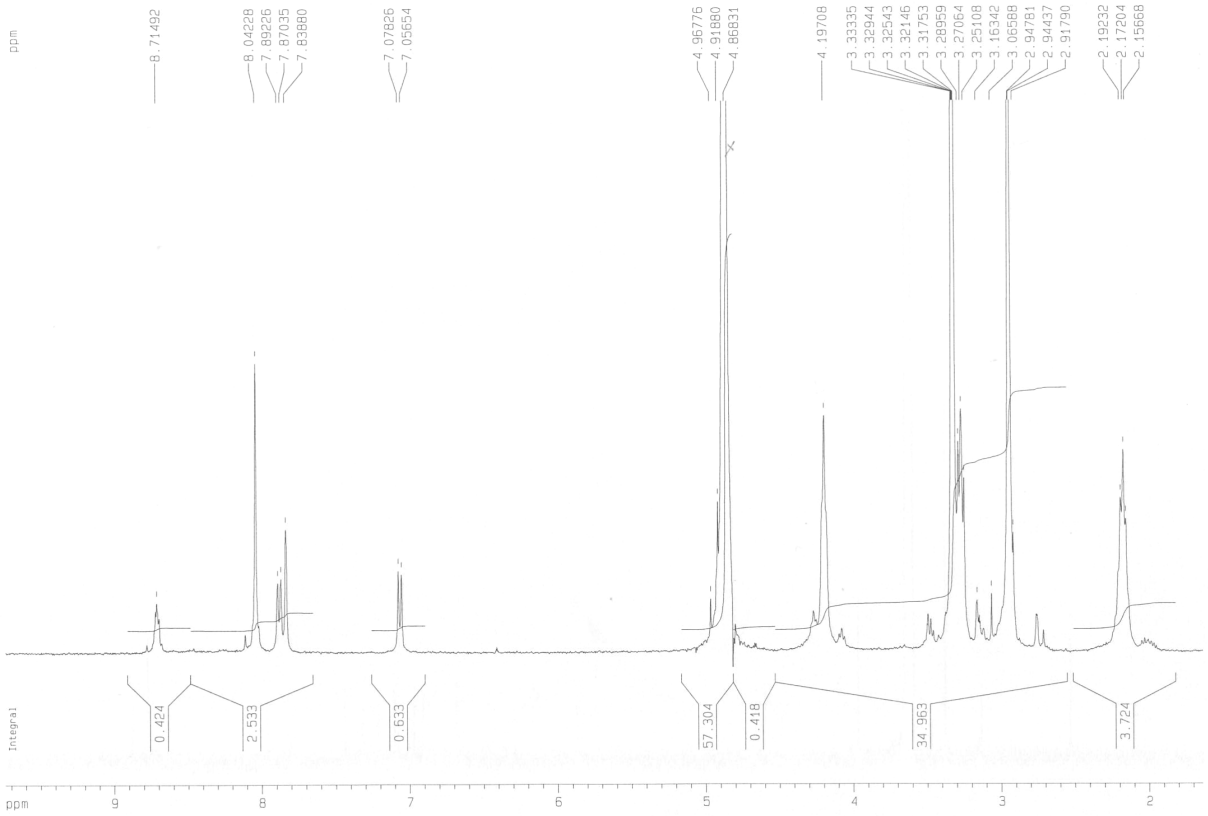


(CD_3OD) ^{13}C NMR 75MHz

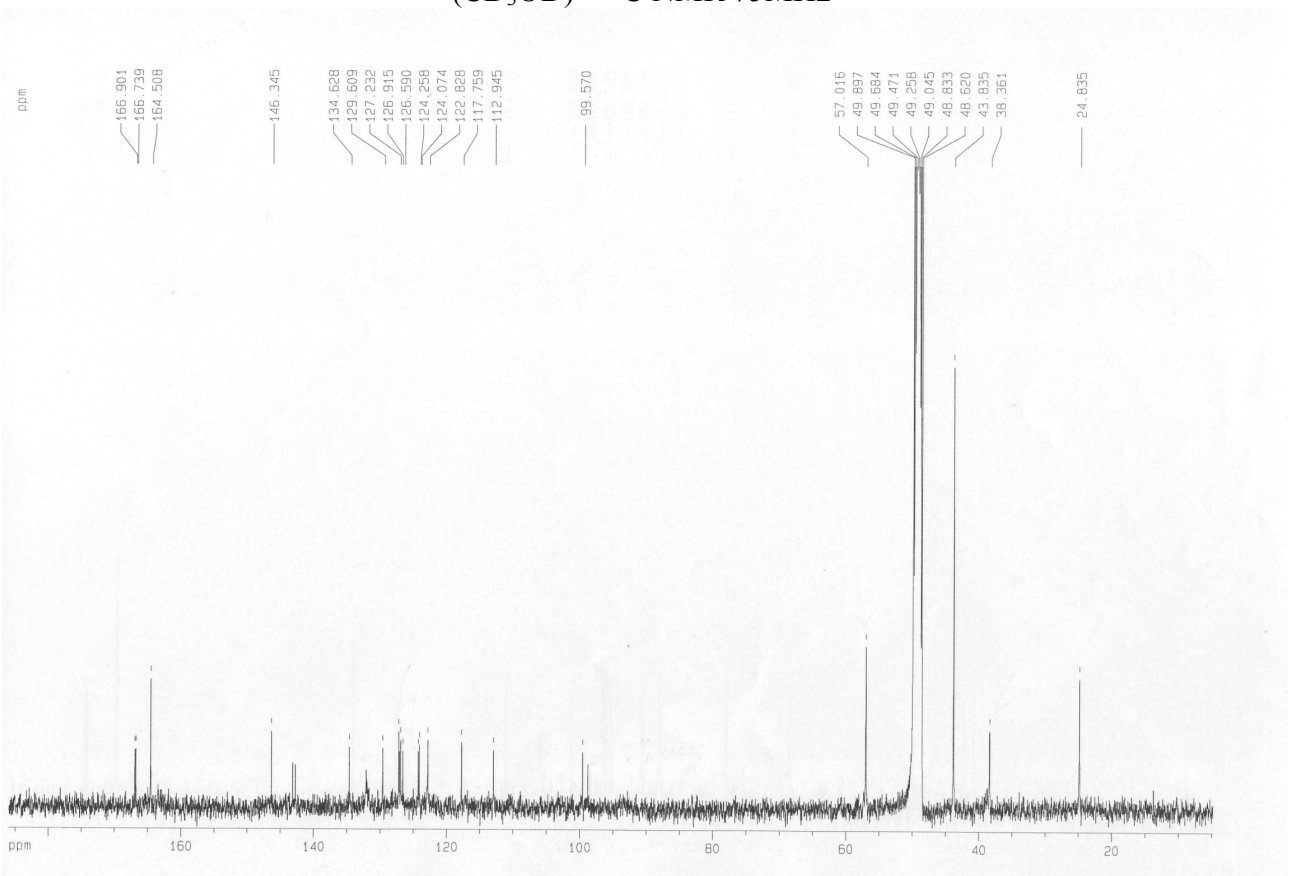


4·2HCl

(CD_3OD) ^1H NMR 300MHz

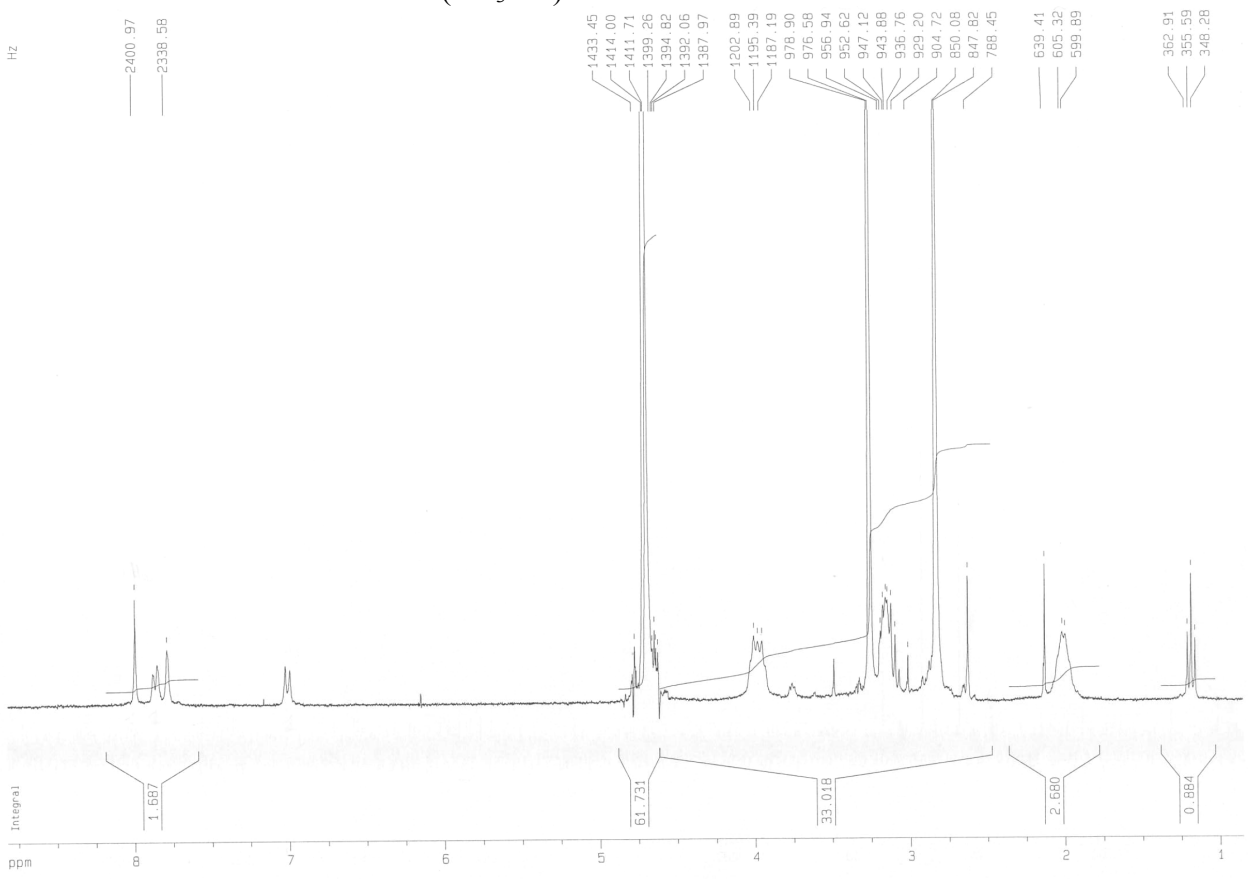


(CD_3OD) ^{13}C NMR 75MHz

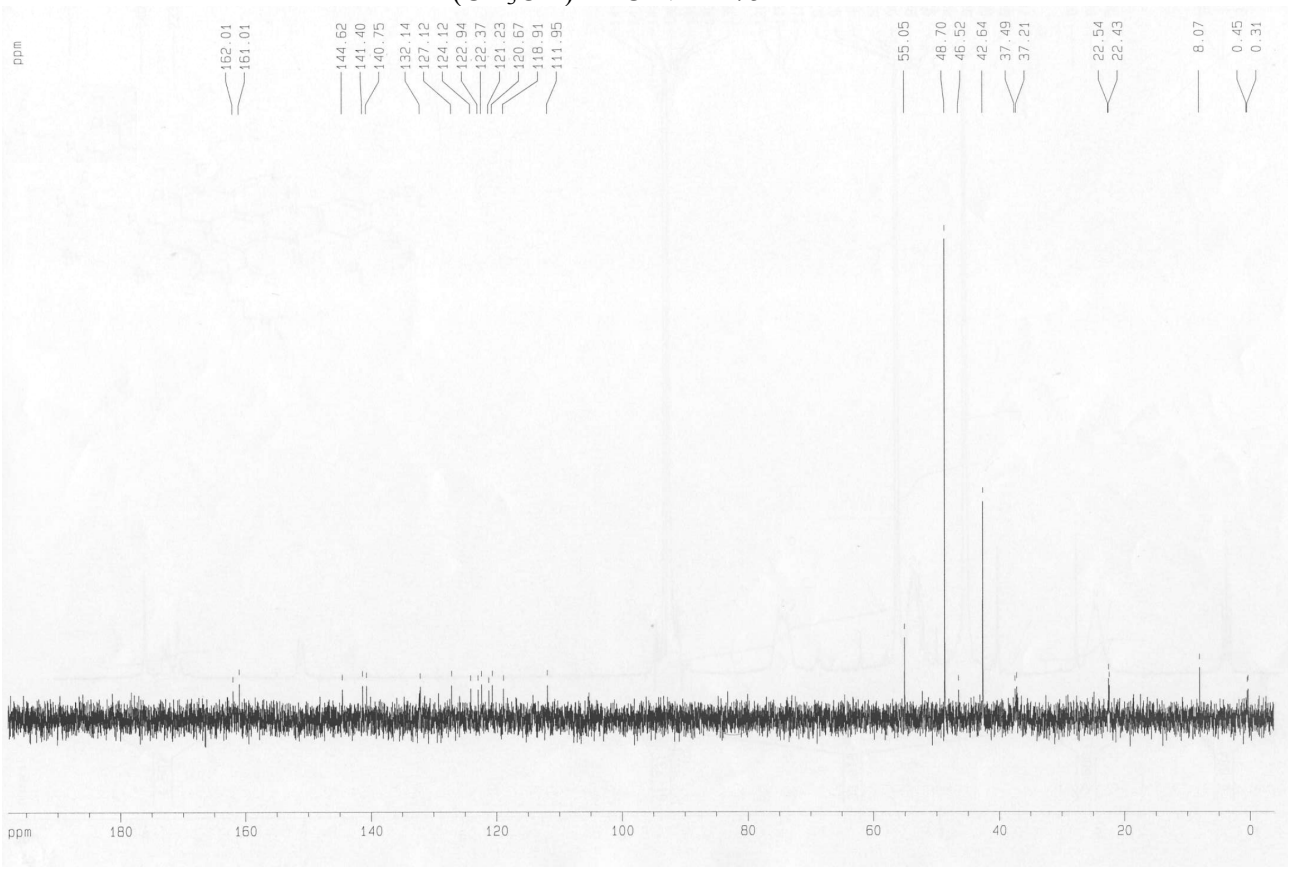


5·2HCl

(CD_3OD) ^1H NMR 300MHz

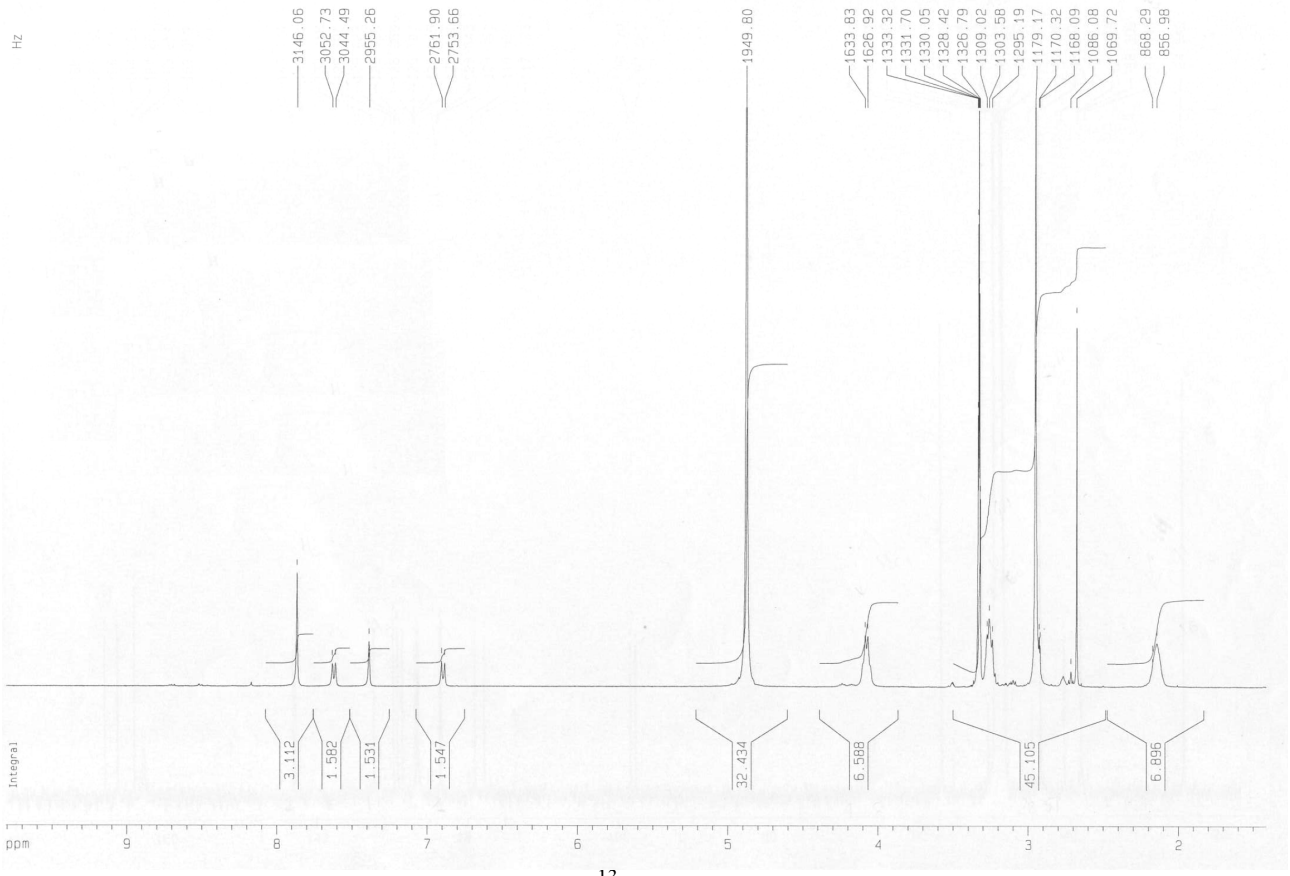


(CD_3OD) ^{13}C NMR 75MHz

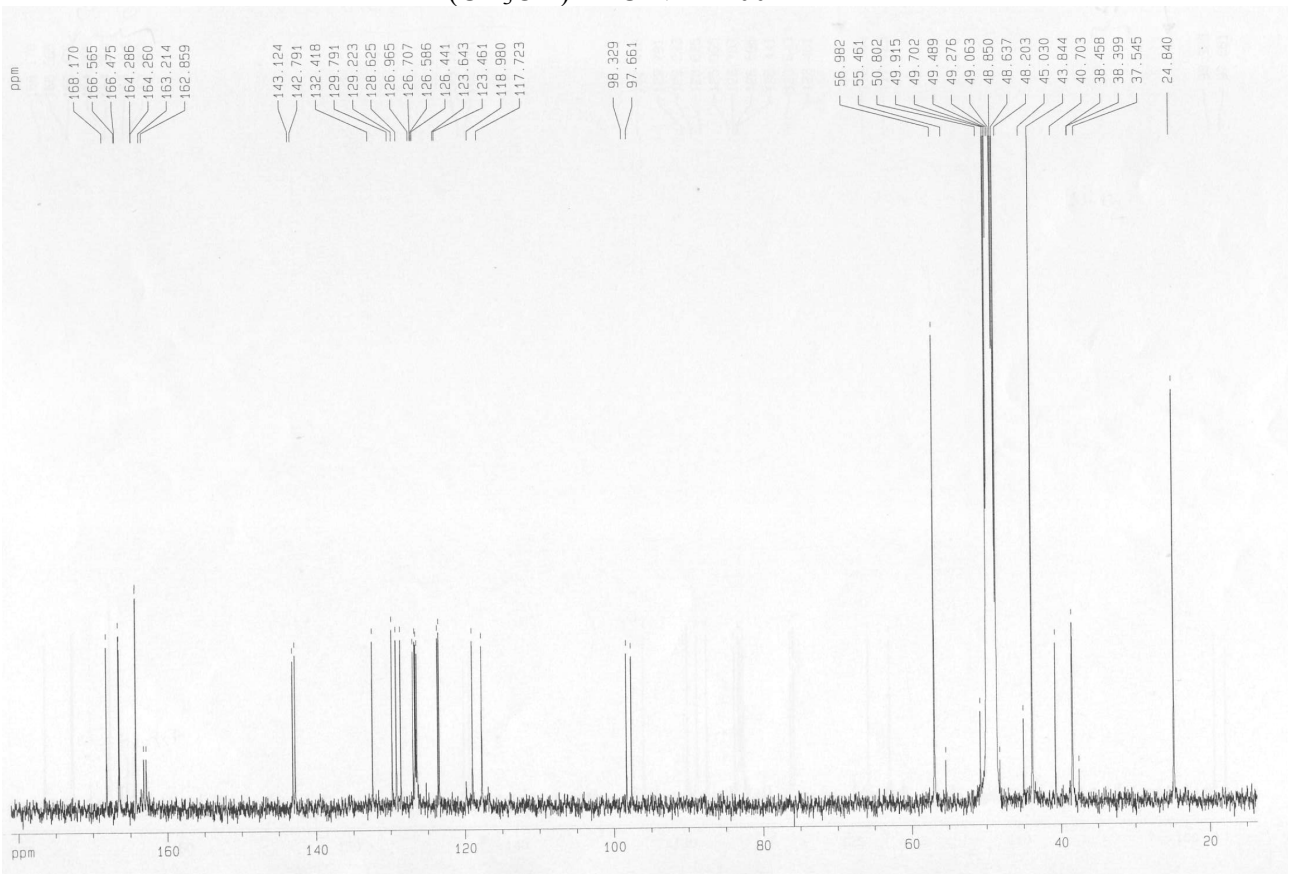


6·2HCl

(CD_3OD) ^1H NMR 400MHz

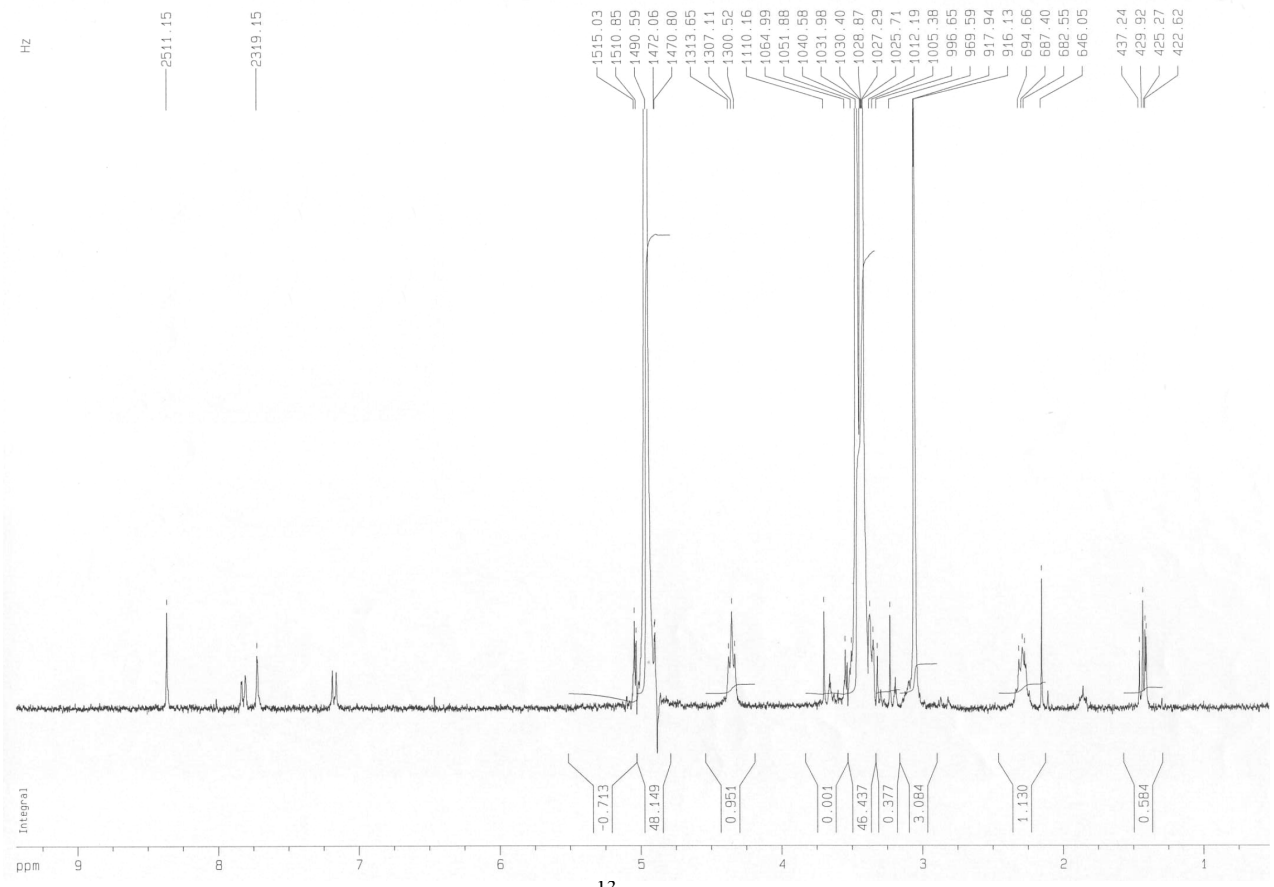


(CD_3OD) ^{13}C NMR100MHz

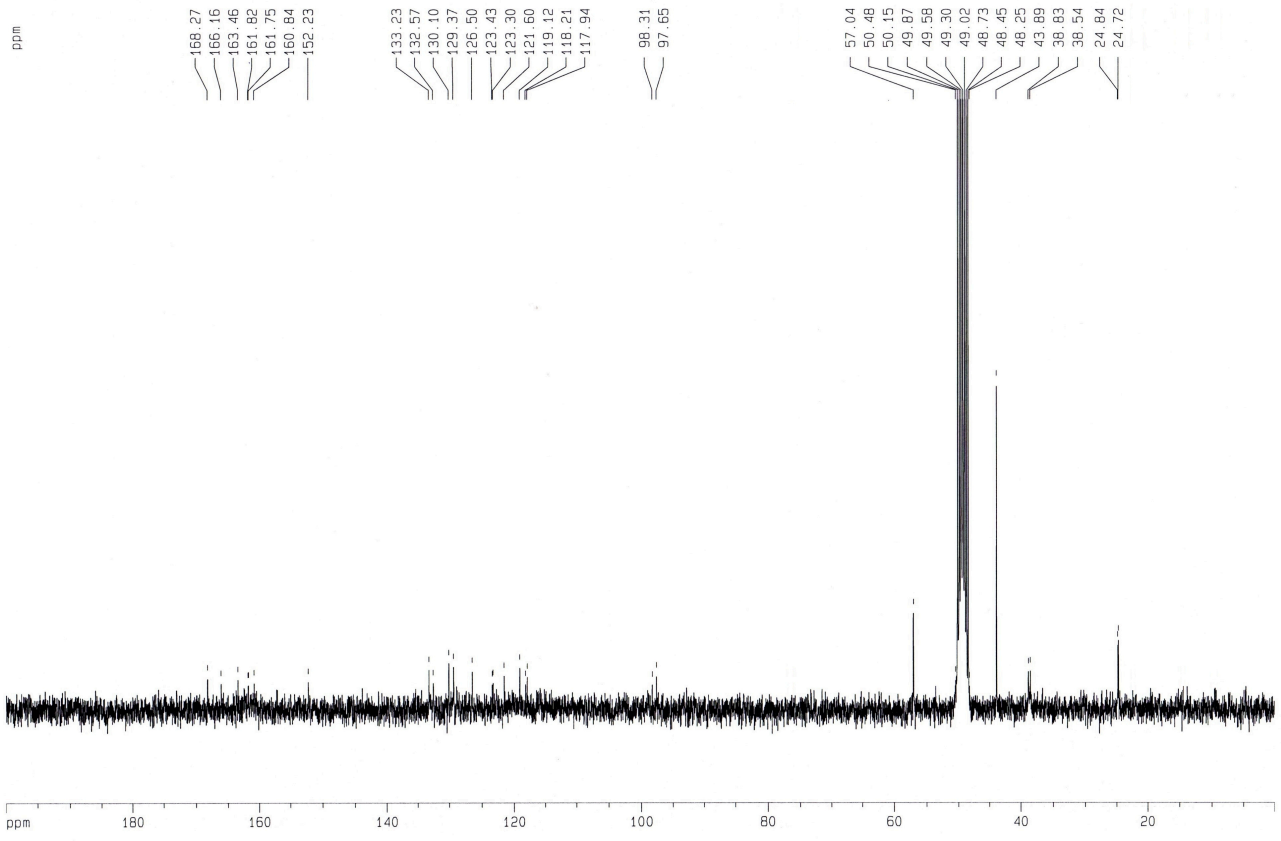


7·2HCl

(CD_3OD) ^1H NMR 300MHz

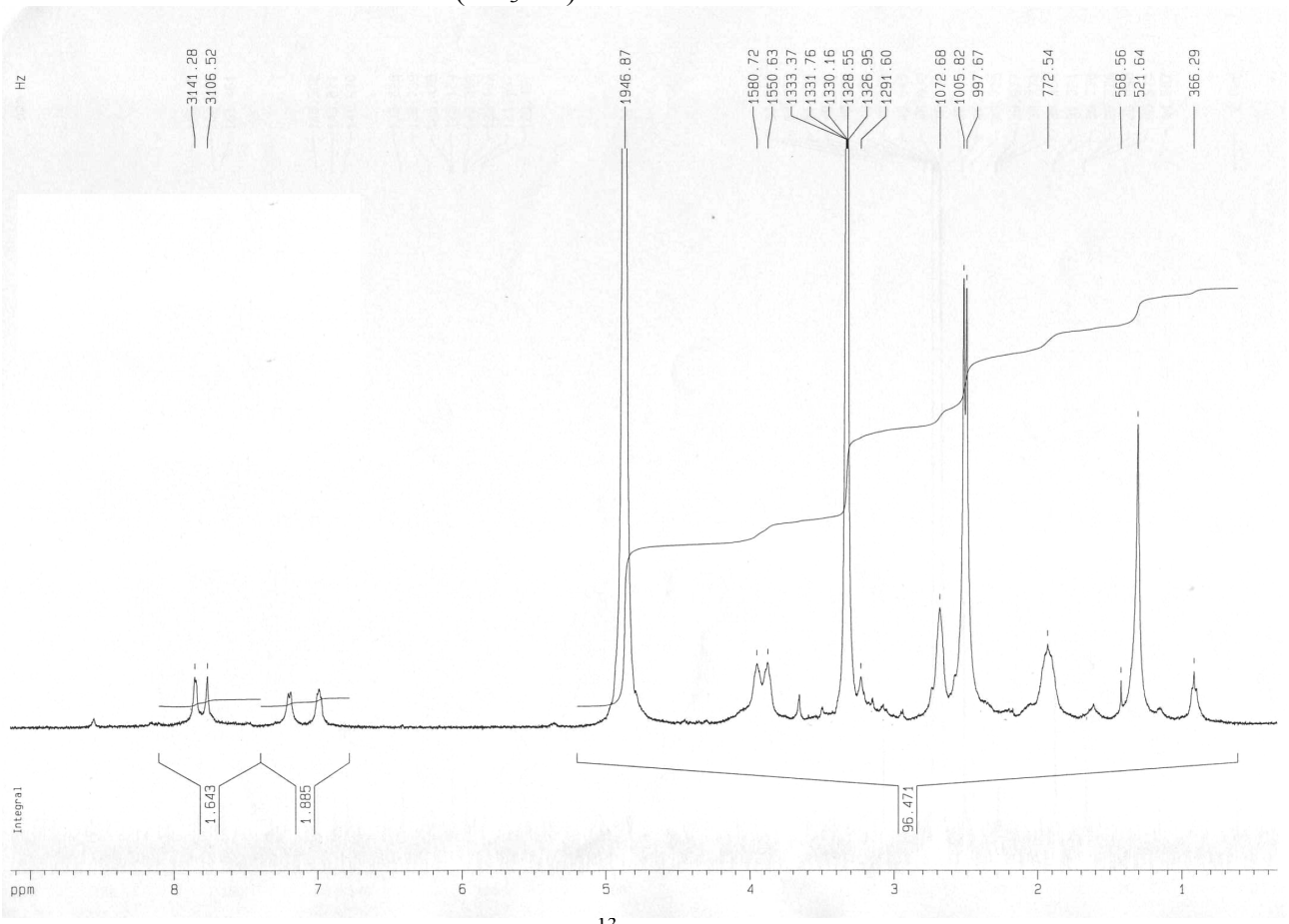


(CD_3OD) ^{13}C NMR 75MHz

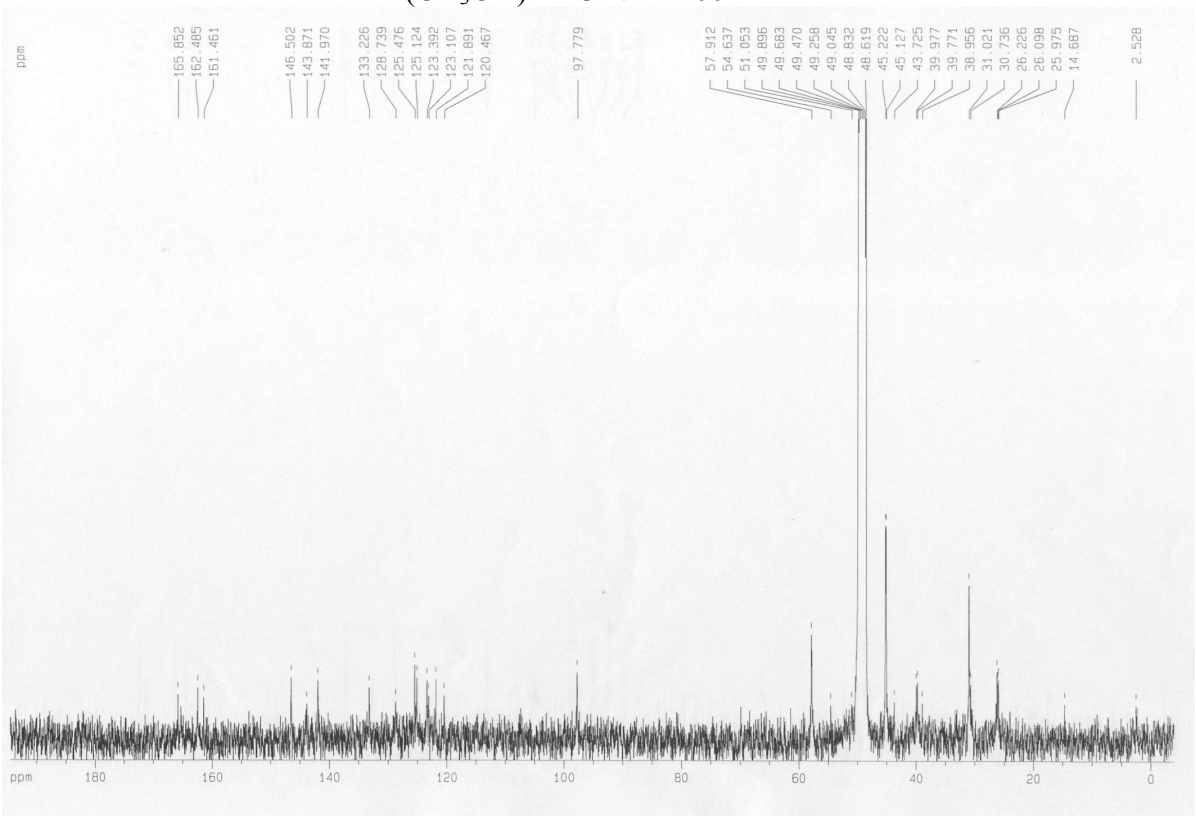


8·3HCl

(CD_3OD) ^1H NMR 400MHz

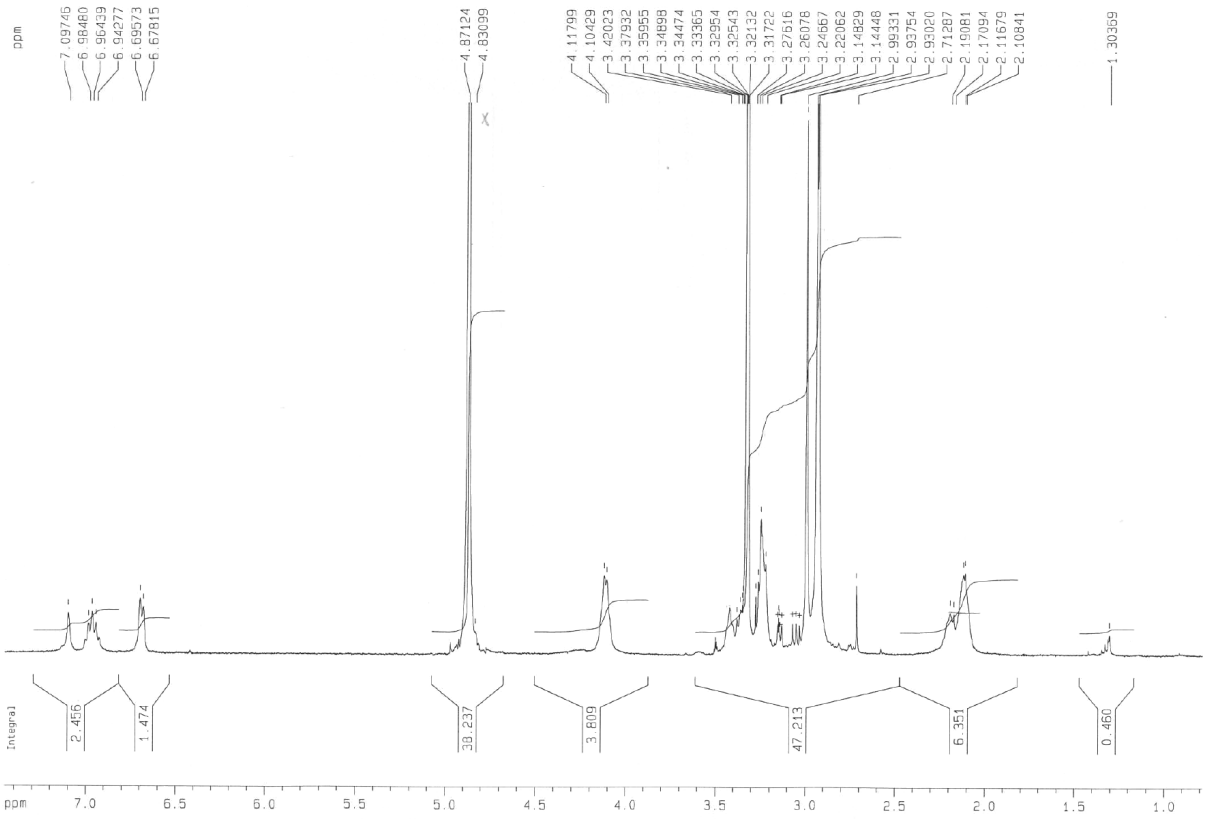


(CD_3OD) ^{13}C NMR100MHz

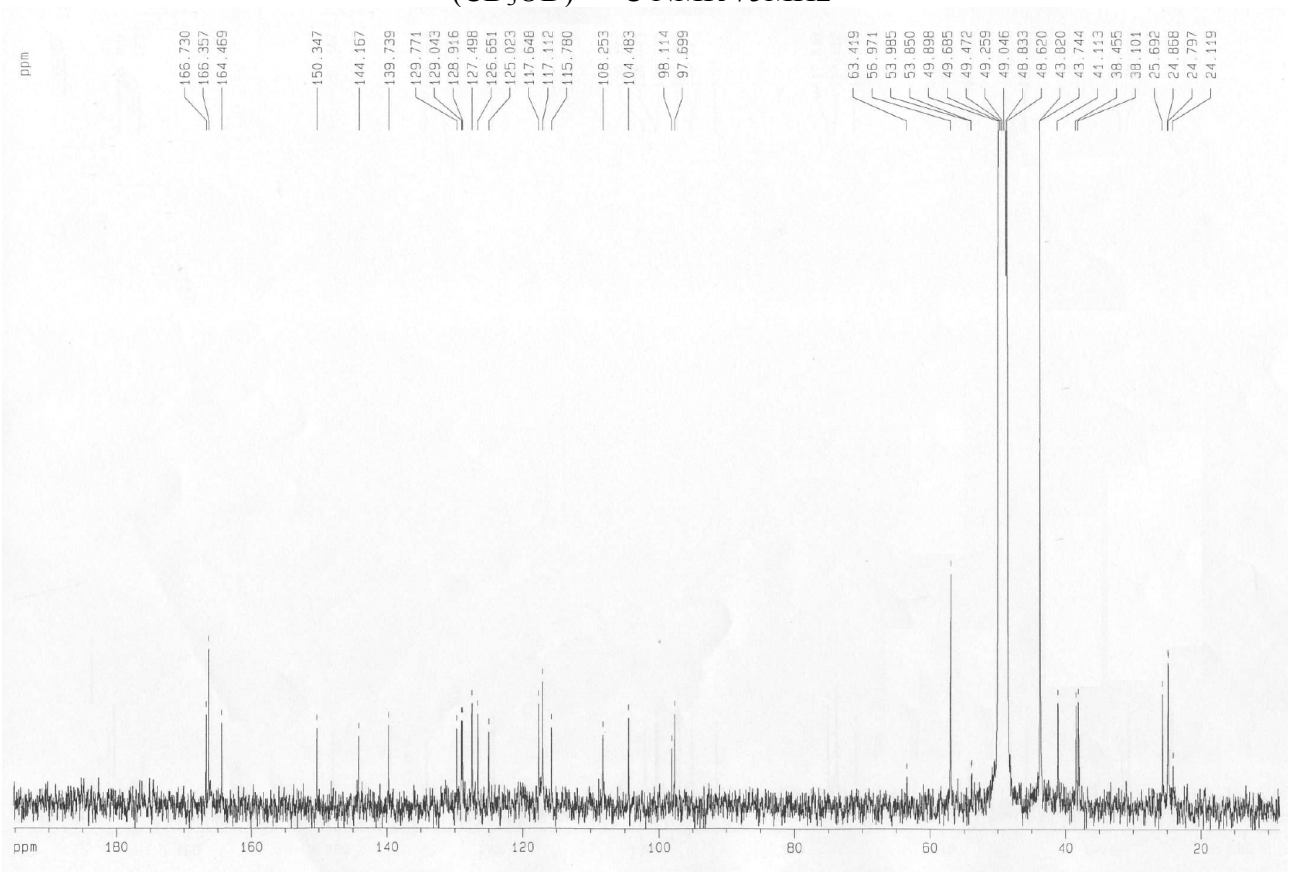


9·3HCl

(CD_3OD) ^1H NMR 300MHz

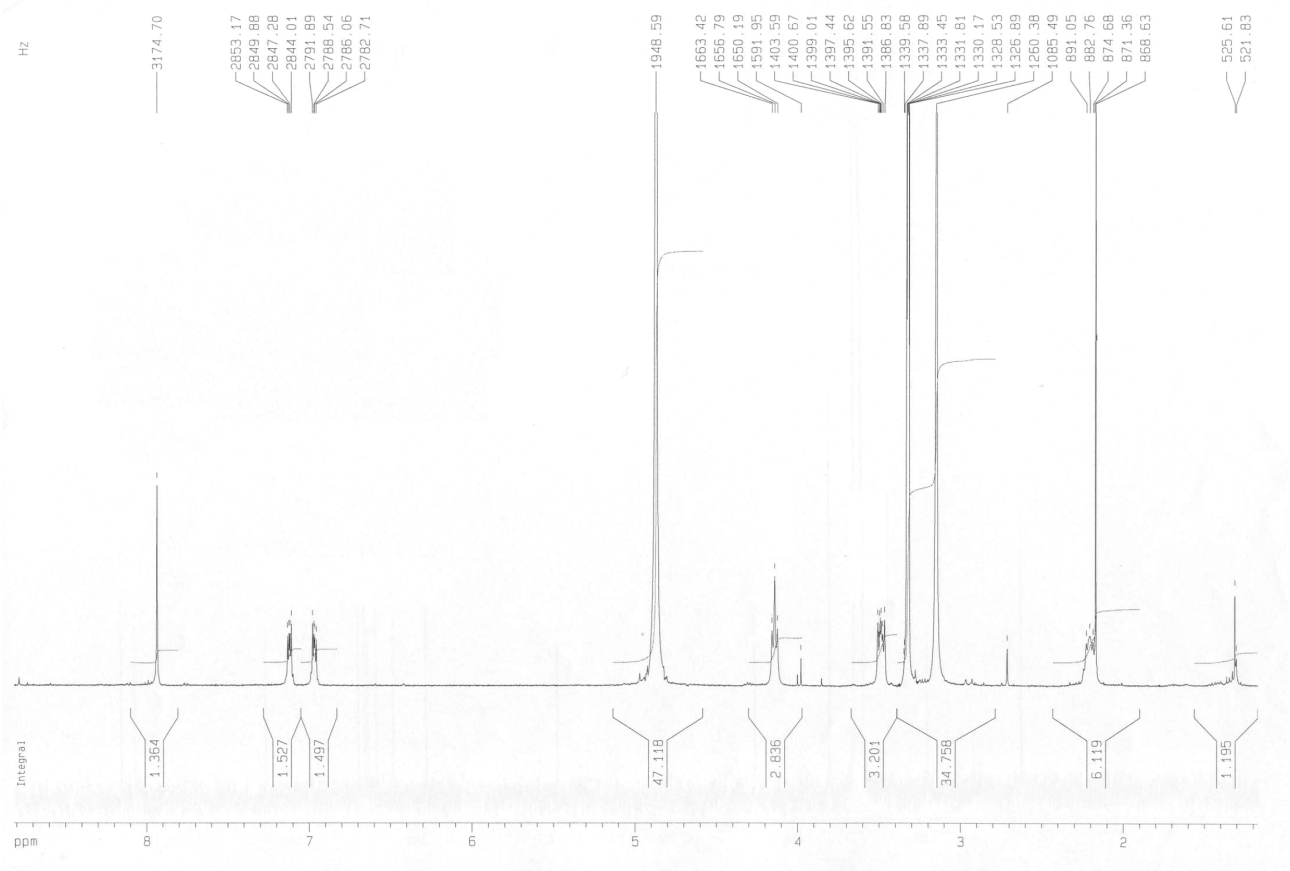


(CD_3OD) ^{13}C NMR 75MHz

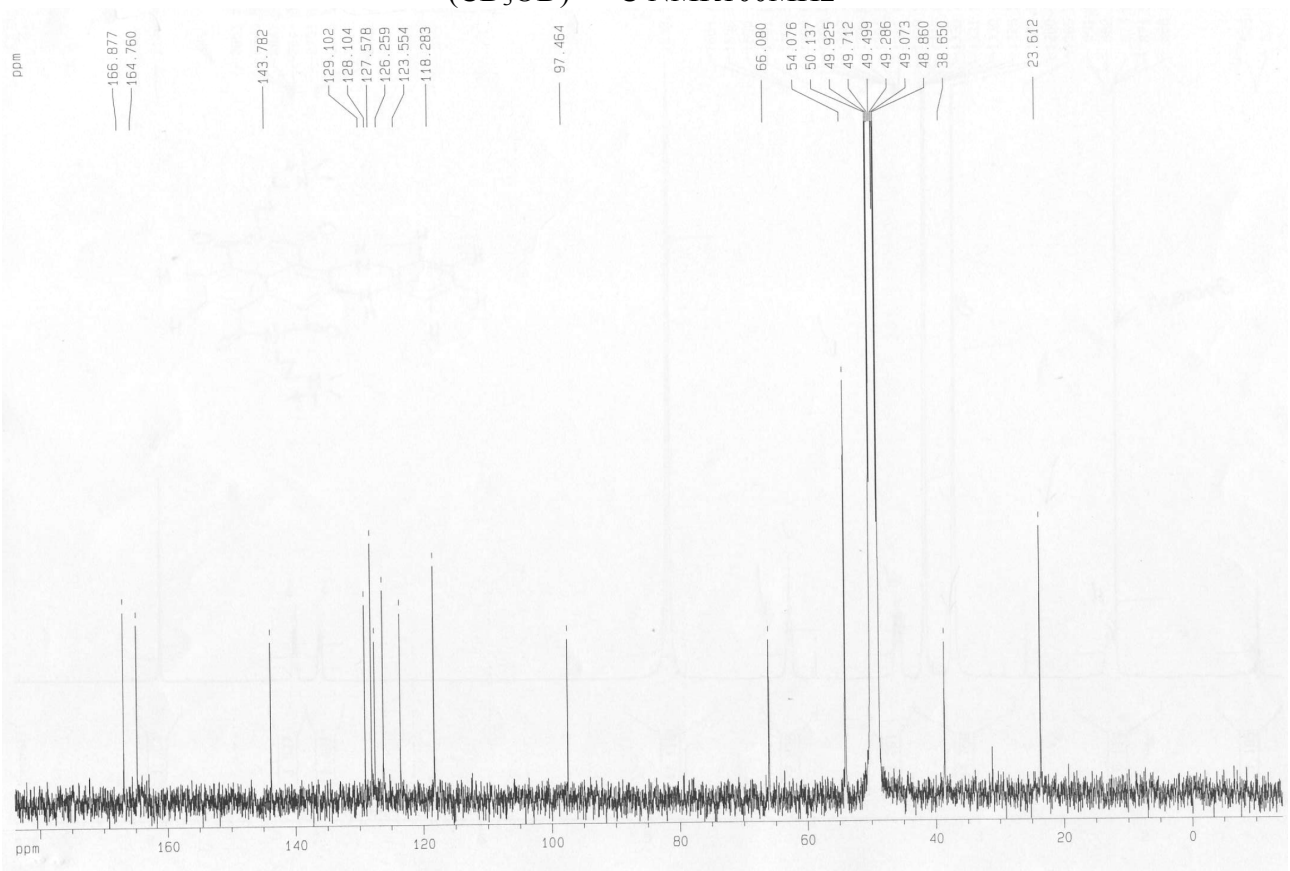


10

(CD_3OD) ^1H NMR 400MHz

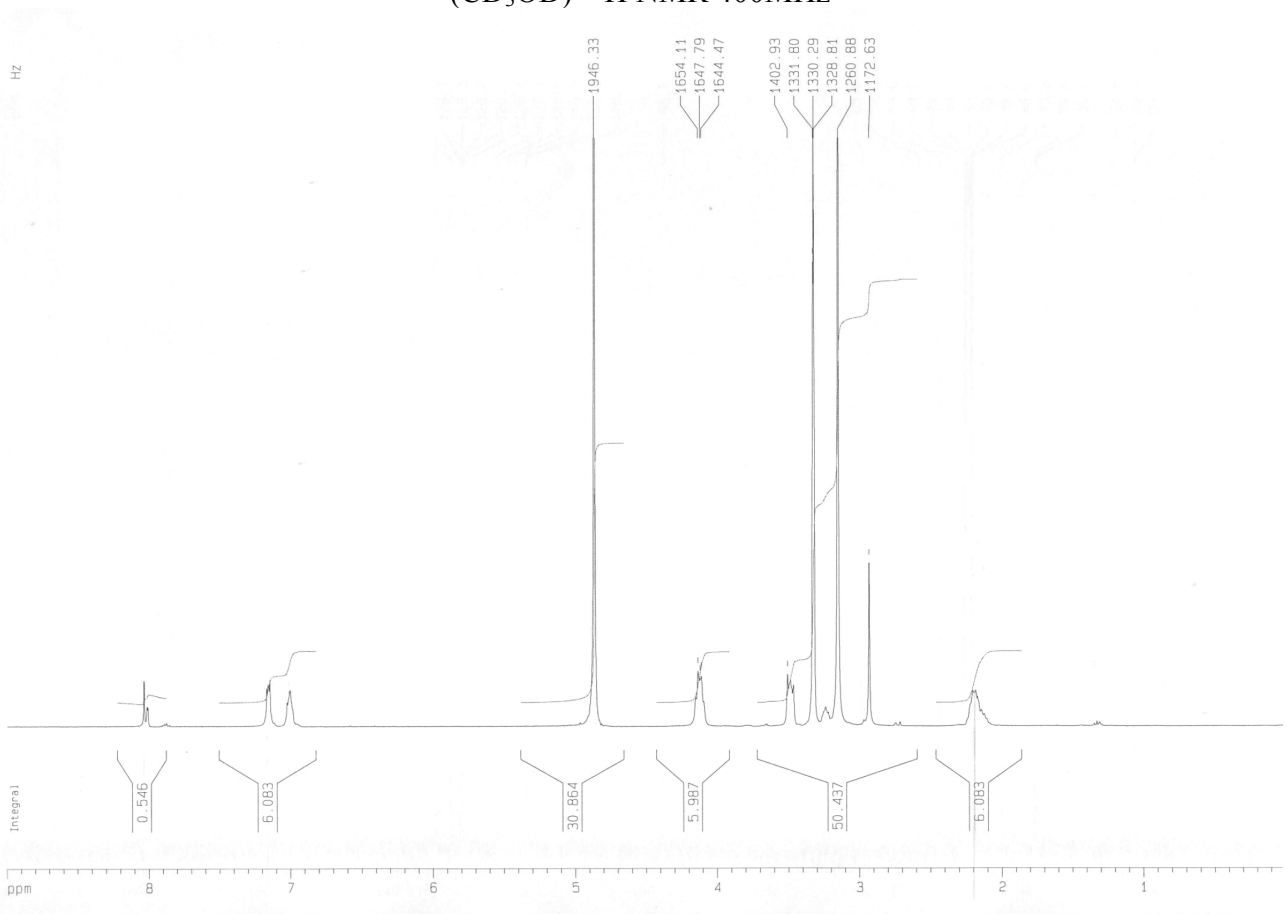


(CD_3OD) ^{13}C NMR100MHz



11

(CD_3OD) ^1H NMR 400MHz



(CD_3OD) ^{13}C NMR100MHz