

**Pharmacophore modeling, Docking and Integrated use of ligand-based and structure based  
Virtual screening approach for the novel DNA gyrase inhibitors: Synthetic and Biological  
Evaluation studies**

Deepti Mathpal<sup>aΨ</sup>, Mukesh Masand<sup>bΨ</sup>, AnishaThomas<sup>c</sup>, Irfan Ahmad<sup>d</sup>, Mohd Saeed<sup>e</sup>, Gaffar Sarwar Zaman<sup>d</sup>, Mehnaz Kamal<sup>f</sup>, Talha Jawaid<sup>g</sup>, Pramod K. Sharma<sup>a</sup>, Madan M. Gupta<sup>h</sup>, Santosh Kumar<sup>i</sup>, Swayam Prakash Srivastava<sup>\*b,j</sup>, Vishal M. Balaramnavar<sup>\*a</sup>

<sup>a</sup>Department of Pharmacy, Faculty of Medicine and Allied Sciences, Galgotias University, Gautam Buddha Nagar, Uttar Pradesh, India, 226001.

<sup>b</sup>Sanskriti University, School of Pharmacy and Research, 28 K. M. Stone, Mathura - Delhi Highway, Chhata, Mathura Uttar Pradesh (U.P.) Pin – 281401

<sup>c</sup>Department of Chemistry, School of Advanced Sciences, VIT, Vellore India.

<sup>d</sup>Department of Clinical Laboratory Science, College of Applied Medical Sciences, King Khalid University, Abha, Saudi Arabia

<sup>e</sup>Department of Biology College of Sciences, University of Hail, Saudi Arabia

<sup>f</sup>Department of Pharmaceutical Chemistry, College of Pharmacy, Prince Sattam bin Abdulaziz University, P.O. Box No. 173, Al Kharj 11942, Kingdom of Saudi Arabia

<sup>g</sup>Department of Pharmacology, College of Medicine, Al Imam Mohammad ibn Saud Islamic University (IMSIU), Othman ibn Affan Street, Riyadh 13317, Kingdom of Saudi Arabia

<sup>h</sup>School of Pharmacy, Faculty of the West Indies St Augustine, Trinidad and Tobago West Indies

<sup>i</sup>Government Degree College, Hansaur, Barabanki, Uttar Pradesh (U.P.) Pin – 225415

<sup>j</sup>Department of Pediatrics, Yale University School of Medicine, New Haven, Connecticut, 06511, USA

\*Corresponding author E-Mail Address: V.M.B.-[\\*v.balaramnavar@gmail.com](mailto:v.balaramnavar@gmail.com)

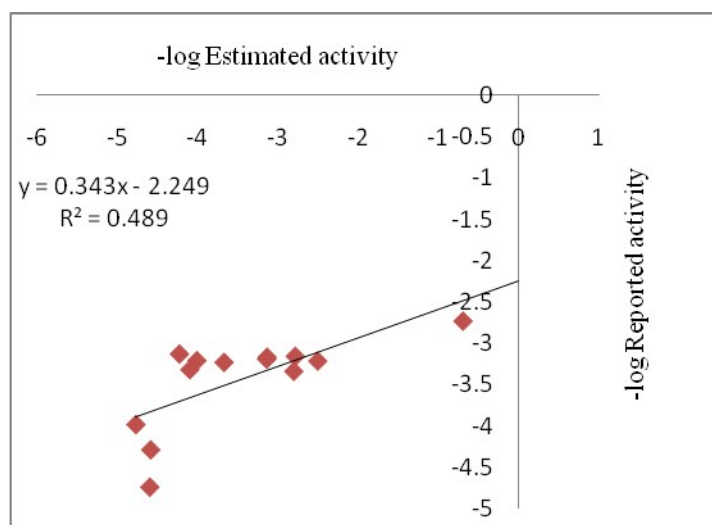
ΨThese two authors have equal contributions for the manuscript.

## Supporting Information

<b>Sr. No.</b>	<b>Title</b>	<b>Page No.</b>
<b>S1</b>	Validation of the Pharmacophore model using structurally diverse and Potent DNA gyrase inhibitors	<b>3</b>
	Figure S1: Graph of prediction in terms of actual and estimated activity with ( $r^2 = 0.49$ ).	<b>3</b>
<b>S2</b>	Design of combinatorial library of Ciprofloxacin derivatives	<b>4</b>
	Table S1: The in-silico developed library of the ciprofloxacin derivatives along with their docking scores with PDB IDs 3ILW and 3FOF.	<b>4</b>
<b>S3</b>	Correlation of binding affinity and activity	<b>5</b>
	Table S2: The comparative docking scores of all the compounds from the series with Pose Energy, Rerank score, Binding Affinity, Gold score, MLR and ANN and $-\log$ MIC values.	<b>6</b>
<b>S4</b>	Spectroscopic data VM-2	<b>8-9</b>
<b>S5</b>	Spectroscopic data VM-2	<b>10-11</b>
<b>S6</b>	HPLC Data	<b>12-15</b>

### S1: Validation of the Pharmacophore model using structurally diverse and Potent DNA gyrase inhibitors

The validated predictive pharmacophore model can be used as a query tool to search the databases of diverse drug-like compounds to identify new molecules with potent DNA gyrase inhibitory activity. To identify new leads is of prime importance in pharmaceutical companies and patent them if already known, which can add value to the company. Therefore before going onto the virtual screening of the diverse databases we have initiated a validation study to check the expediency of the model. We had prepared dataset of 13 known potent DNA gyrase inhibitors which have already been clinically introduced. The rationale behind this study was that if the pharmacophore was really predictive, then it should predict the established DNA gyrase inhibitors as active in terms of both fit value and activity in the range of one log activity. Their mapping was carried out using the Best Fit option in Catalyst Discovery Studio. Results of this study are indicated in the Figure S1. The selected hypothesis mapped onto three features of the model (two hydrogen bond acceptors and one hydrophobe) for all the molecules except prulifloxacin that maps well on this model. This can be explained by the fact that certain known inhibitors have molecular weights  $\geq 500$ . Even though smaller molecules do not map onto the fourth acceptor, the molecules are active provided they map onto the rest of the three features. The model showed good prediction in terms of actual and estimated activity with ( $r^2 = 0.49$ ).

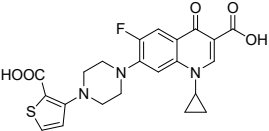
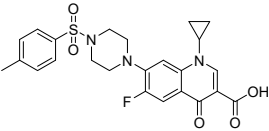
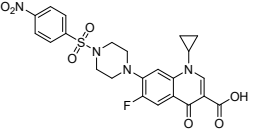


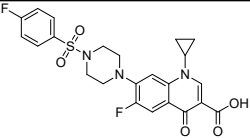
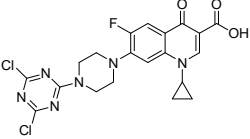
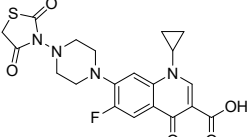
**Figure S1:** Graph of prediction in terms of actual and estimated activity with ( $r^2 = 0.49$ ).

## S2 Design of combinatorial library of Ciprofloxacin derivatives

In this case we had generated the library of ciprofloxacin derivatives by using prior knowledge of the binding site, 3D pharmacophore features and using the ciprofloxacin as fluoroquinolones for substitution at C7 piperazine nitrogen. All the R1, R3, R5, R7 were kept as such for analysis of these molecules and respective docking scores. The pharmacophore mapping and docking analysis of the binding site for quinolones reveal the importance of acceptor functions to impart the activity of the existing ligands. In this attempt we tried some substitutions to fulfil the requirements of pharmacophore as well as binding site in terms of fit values and docking scores. In the present study, we had applied state of art techniques to design new motifs in terms of simple substitutions on the piperazine nucleus to discover some important modifications on the existing drugs. We had designed six ligands that can be synthesized using simple substitution reactions and has meaningful results in terms of activity. We had considered the fit values, estimated activities; docking score in terms of both Gold score and Rerank score to verify our approach. According to the requirements of the binding site the results of the designed ligands are condensed in the Table S1 below. The structures of the ligands were shown in the table below.

**Table S1:** The in-silico developed library of the ciprofloxacin derivatives along with their docking scores with PDB IDs 3ILW and 3FOF.

Code	Structure	Fit value	Estimated activity (nM)	Moldock Score		Rerank Score		Gold score	
				3ILW	3FOF	3ILW	3FOF	3ILW	3FOF
13.1		10.818	9.583	-133.69	-103.10	-76.55	-66.21	34.48	63.73
13.3		8.645	1430.36	-107.23	-94.72	-86.21	-75.10	53.08	73.15
13.5		11.407	2.47	-122.4	-100.22	-102.16	-16.76	53.57	75.46

13.6		8.61	1550.34	-107.147	-97.43	-85.56	-67.24	52.12	69.14
13.7		11.39	2.571	-106.335	-108.85	-76.58	-72.49	54.11	67.24
13		9.20	391.428	-112.197	-87.77	-77.28	-66.21	40.36	69.38

### S3 Correlation of binding affinity and activity

Moldock predicts binding affinities and interaction energies. Binding affinity is supposed to be providing deeper insights into structural insights in understanding the drug receptor interactions. The binding affinity prediction contributes some of the solvent interactions and entropy which are difficult to handle in simple models in docking protocols. The rerank score in the MVD provides estimates in terms of strength of interactions but lacks the consideration of entropy in estimation. Another drawback of using the rerank score is that the rerank score may be applied for ranking of the different poses of the same ligand but can't be applied to the ranking of poses of different ligands. The binding affinity can be measured from a data analyzer in MVD. These protocols use the terms extracted from Moldock score and also the static descriptors not using the 3D conformations of the pose (like MW and or no. of nitrogen atoms). The Molegro virtual docker inbuilt binding affinity prediction protocol was based on the coefficients for binding affinity terms derived from multiple linear regression. The model was based on more than 200 structurally diverse complexes (From PDB data bank) with known binding affinities in kJ/mol. The Pearson correlation coefficient was 0.60 with 10 fold cross validation. To avoid the false positive results in terms of binding affinity we had used the combined approach to analyze the virtual library as the model for binding affinities was based on known complexes and it may sometimes give strong affinities for weakly active compounds also. In this attempt, we got the good correlation of docking results in terms of rerank score and  $PIC_{50}$  values ( $r^2$  0.6142). We also got the good correlation of binding affinity and MLR and binding affinity and ANN that were predicted by the model developed in the Molegro software Table S2.

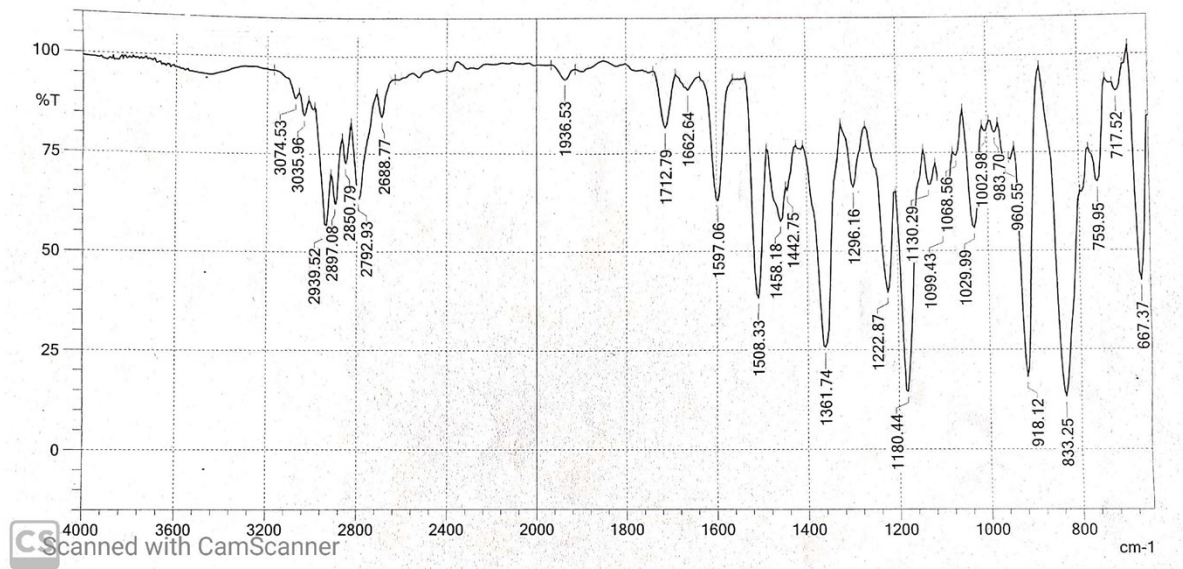
**Table S2:** The comparative docking scores of all the compounds from the series with Pose Energy, Rerank score, Binding Affinity, Gold score, MLR and ANN and  $-\log$  MIC values.

Name	Pose Energy	Rerank Score	Binding Affinity	MLR-Train	ANN-Train	Gold score	$-\log$ MIC
122_1A	-85.41	-64.96	-20.94	12.61	13.37	44.52	-4.02
122_1B	-85.50	-65.66	-20.87	12.61	13.36	43.65	-4.32
122_1J	-86.82	-69.66	-21.33	11.80	12.26	40.36	-3.73
122_1L	-97.14	-64.11	-21.38	11.30	10.91	49.82	-4.00
122_1M	-97.58	-75.38	-21.39	11.23	10.89	44.98	-3.70
122_1N	-130.97	-72.15	-23.85	11.48	11.04	55.93	-4.21
122_1O	-98.21	-71.66	-20.47	11.14	10.78	38.72	-4.16
122_1P	-89.36	-69.47	-23.86	10.30	10.25	44.54	-3.71
122_1Q	-101.94	-76.05	-21.87	9.14	9.44	41.83	-3.69
122_1R	-99.88	-73.55	-20.68	9.06	9.28	59.67	-3.37
122_1S	-109.66	-81.85	-20.24	8.65	9.16	58.44	-3.35
122_1T	-102.35	-78.78	-20.69	9.07	9.26	61.46	-3.35
122_1U	-108.82	-71.21	-21.80	14.46	15.40	56.41	-3.07
122_1V	-92.37	-70.92	-23.06	15.20	15.02	54.30	-3.35
122_1W	-107.99	-76.24	-24.99	14.43	15.17	55.85	-3.86
122_1X	-111.97	-66.87	-21.19	14.46	15.07	57.36	-4.24
126_6B	-93.81	-69.13	-21.38	12.72	13.22	49.02	-4.29
126_6C	-95.57	-61.47	-21.40	12.74	13.20	46.94	-4.29
126_7F	-132.50	-76.16	-19.15	13.88	14.71	56.39	-3.96
126_7G	-106.07	-83.02	-24.91	20.81	21.15	62.40	-3.37
148_3C	-82.29	-63.77	-21.31	12.86	13.19	52.27	-4.07
148_3D	-93.00	-65.92	-19.89	13.96	13.54	49.91	-3.47
148_3E	-90.46	-63.15	-20.92	12.70	12.88	55.35	-3.74
148_3F	-89.22	-70.33	-20.77	12.60	13.06	53.79	-4.05
148_3J	-80.60	-63.03	-23.06	11.72	11.10	51.65	-4.36
90_1A	-101.26	-70.18	-25.83	10.97	11.07	42.89	-3.09
90_1B	-101.17	-70.12	-29.44	11.33	10.92	45.50	-3.06
90_1C	-100.54	-67.34	-26.50	10.97	11.06	47.96	-3.04
90_1E	-100.15	-77.14	-26.06	10.73	11.16	47.28	-3.01
90_1F	-85.91	-65.92	-25.90	11.09	11.34	48.13	-3.26
3A_C	-121.14	-93.08	-23.32	15.14	15.38	62.97	-0.20
3B_C	-130.33	-85.53	-27.14	15.11	15.23	64.77	-0.47
3C_C	-126.37	-82.73	-27.98	14.90	15.33	65.98	-0.44
3D_C	-127.63	-85.50	-24.93	14.99	15.16	55.59	-0.49
3E_C	-139.07	-87.80	-29.24	15.88	16.11	42.76	-0.45
3F_C	-158.98	-86.54	-32.14	15.60	15.99	51.75	-0.43
3G_C	-134.96	-96.46	-31.79	15.59	15.98	61.29	-0.09
3H_C	-148.02	-92.66	-26.01	15.90	15.95	53.76	-0.14

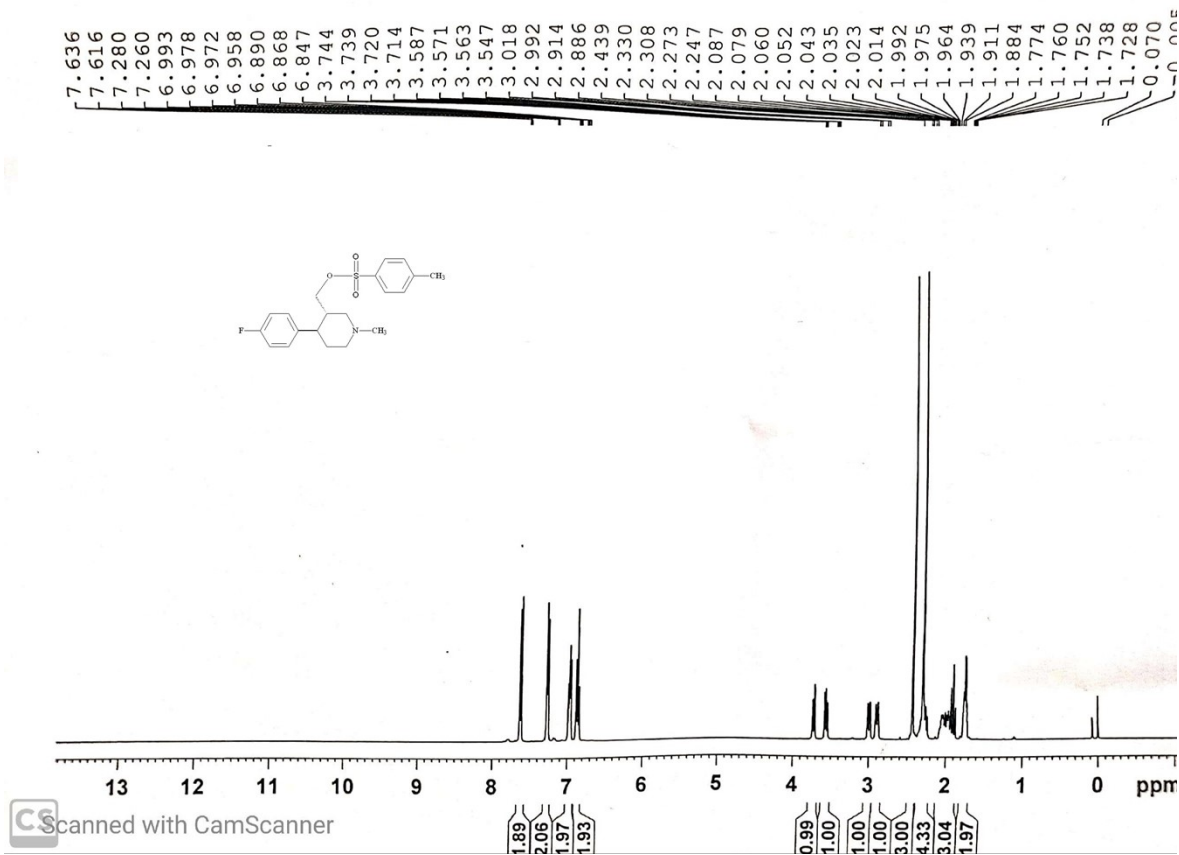
3I_C	-125.43	-92.01	-26.94	16.91	17.08	51.69	-0.14
3J_C	-163.41	-91.65	-30.79	16.85	17.41	42.09	-0.42
3K_C	-136.42	-90.90	-30.34	16.74	17.17	72.28	-0.08
3L_C	-141.46	-81.29	-26.15	17.04	17.02	42.95	-1.03
1	-142.53	-76.14	-41.48	25.43	25.87	22.50	-3.59
2	-146.13	-80.90	-41.91	25.63	25.94	48.05	-3.28
3	-169.07	-64.24	-42.04	25.98	26.00	28.22	-3.57
4	-154.85	-81.44	-44.91	26.09	25.88	33.44	-3.57
5	-134.73	-75.24	-47.38	25.75	25.88	19.57	-3.56
6	-147.68	-80.10	-44.15	26.01	26.20	20.55	-3.55
7	-184.89	-70.96	-41.49	25.96	26.19	24.78	-3.61
8	-161.00	-81.71	-43.96	26.18	26.01	33.68	-3.59
9	-152.46	-64.43	-43.31	26.14	26.33	22.12	-3.58
10	-150.26	-77.37	-44.05	25.48	25.89	46.22	-3.61
11	-204.99	-77.00	-39.79	26.60	26.15	46.22	-3.61
12	-132.18	-72.26	-45.67	25.76	26.12	30.86	-3.60
3A_G	-136.55	-81.32	-32.92	16.56	16.08	25.91	-2.52
3B_G	-133.80	-73.57	-29.51	16.14	16.02	33.98	-2.81
3C_G	-136.76	-78.34	-32.22	16.46	16.09	26.57	-2.79
3D_G	-133.24	-79.74	-28.57	16.39	16.07	33.14	-1.32
3E_G	-129.27	-76.63	-38.86	17.28	17.04	34.61	-2.20
3F_G	-133.31	-86.38	-28.70	17.18	17.07	36.74	-3.09
3G_G	-129.50	-77.72	-35.69	17.24	17.13	34.67	-2.78
3H_G	-145.17	-84.10	-28.42	17.21	17.07	38.22	-2.52
3J_G	-141.54	-84.11	-33.68	21.17	21.03	35.73	-3.06
3K_G	-127.42	-69.62	-31.09	20.88	20.98	32.07	-2.77
3L_G	-130.75	-76.94	-34.00	21.01	21.01	34.58	-2.46
3M_G	-140.89	-71.30	-36.15	20.63	20.97	45.63	-2.18
3N_G	-131.84	-81.25	-35.56	26.73	26.14	32.46	-2.40
3O_G	-134.49	-73.57	-39.60	26.76	26.03	35.59	-2.41
3P_G	-139.67	-72.73	-32.70	26.96	26.22	22.37	-2.12
3Q_G	-132.39	-82.10	-32.76	26.45	26.15	35.53	-2.41
90_1D	-100.79	-63.47	-26.34	10.93	11.06	46.55	-3.37

## S4 Spectroscopic data VM-2

### IR Spectroscopy

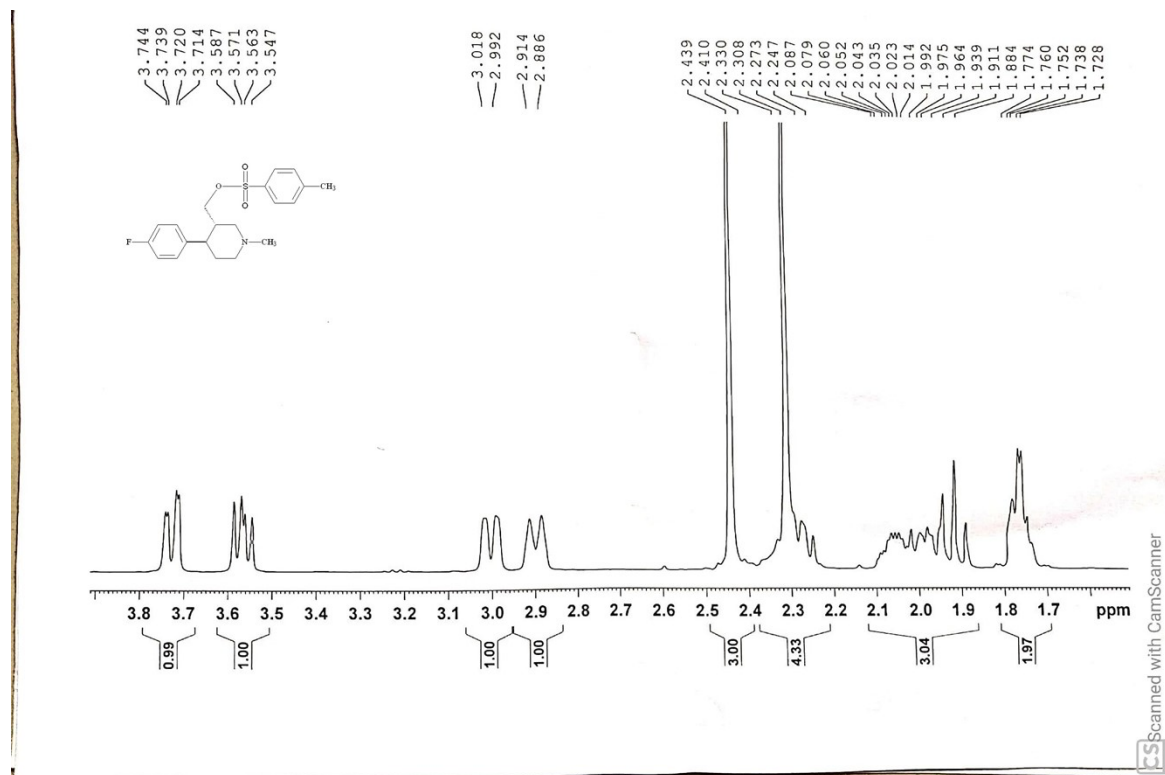


### VM-2 NMR H1 Spectra

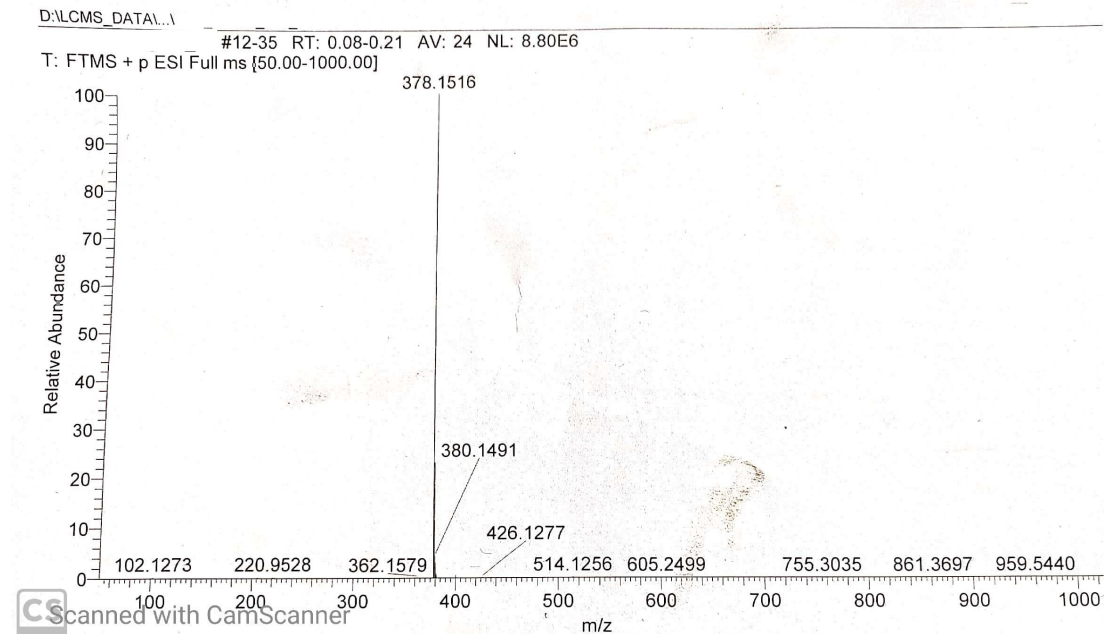




# VM-2 NMR H1 Spectra Expanded

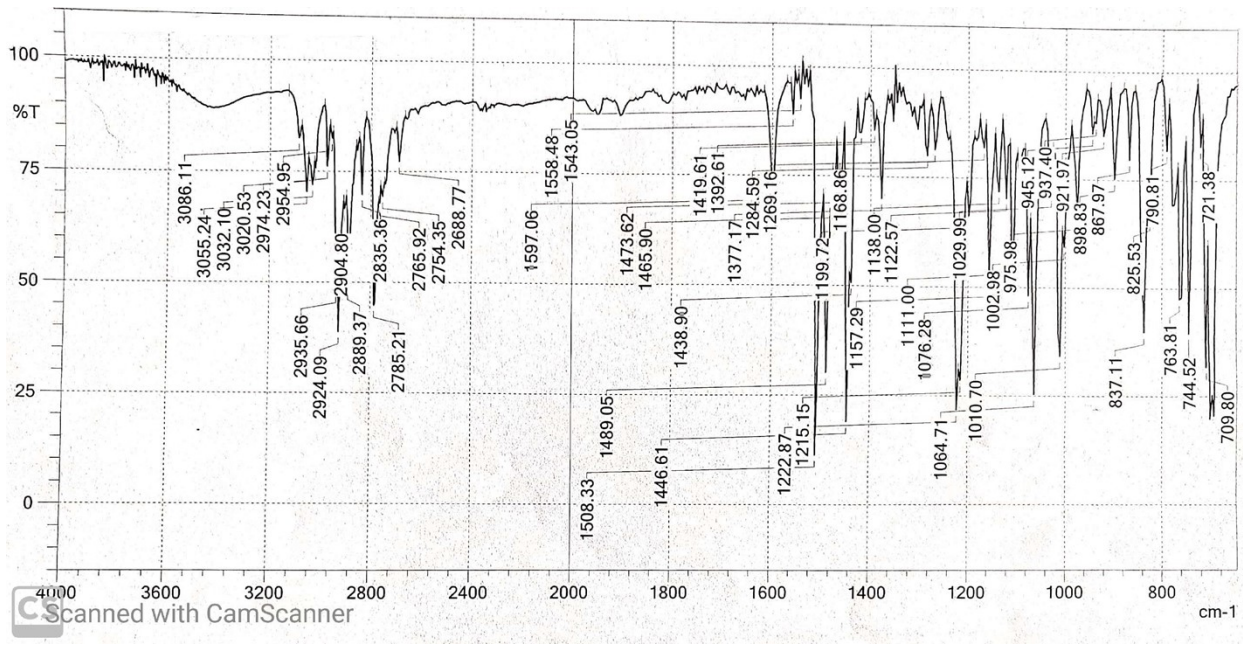


# VM-2 Mass Spectra

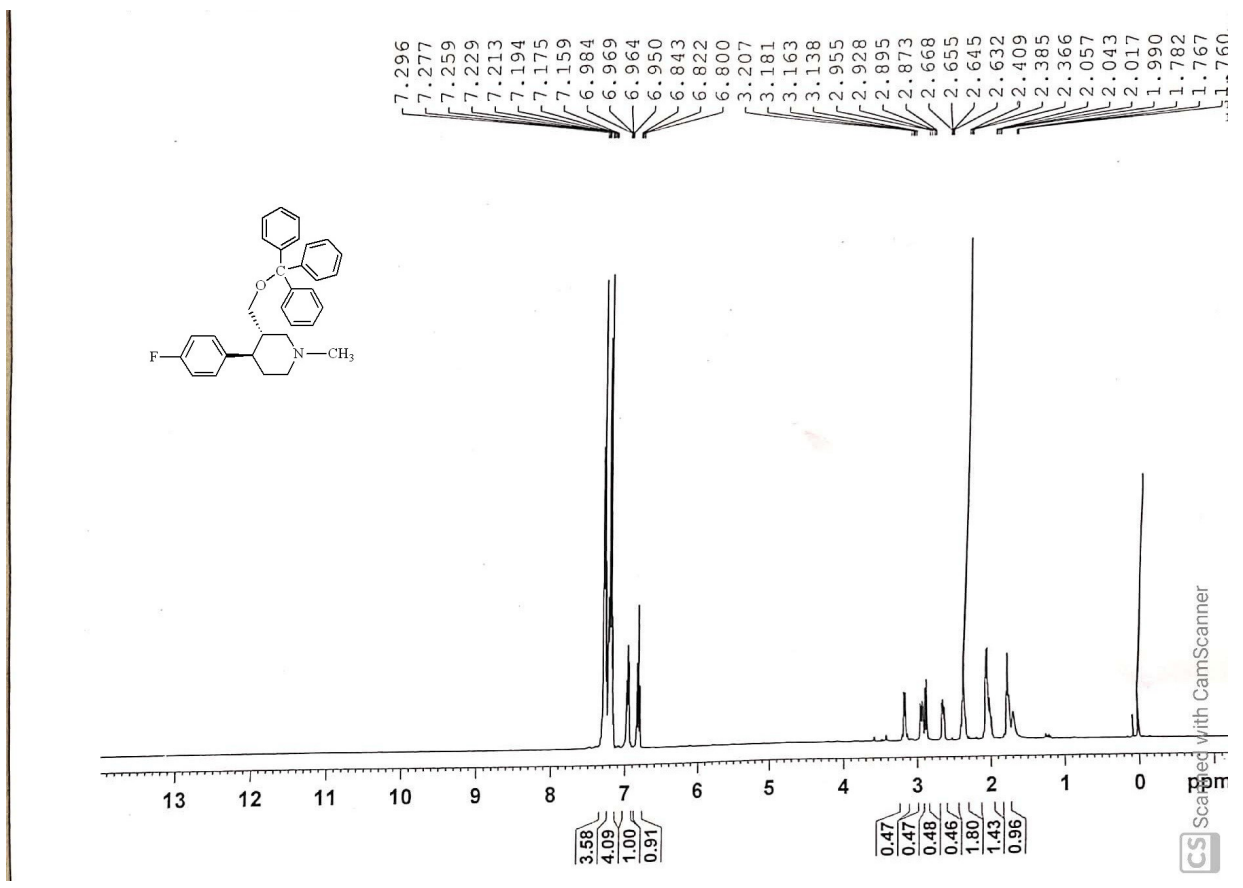


# Spectroscopic data VM-7

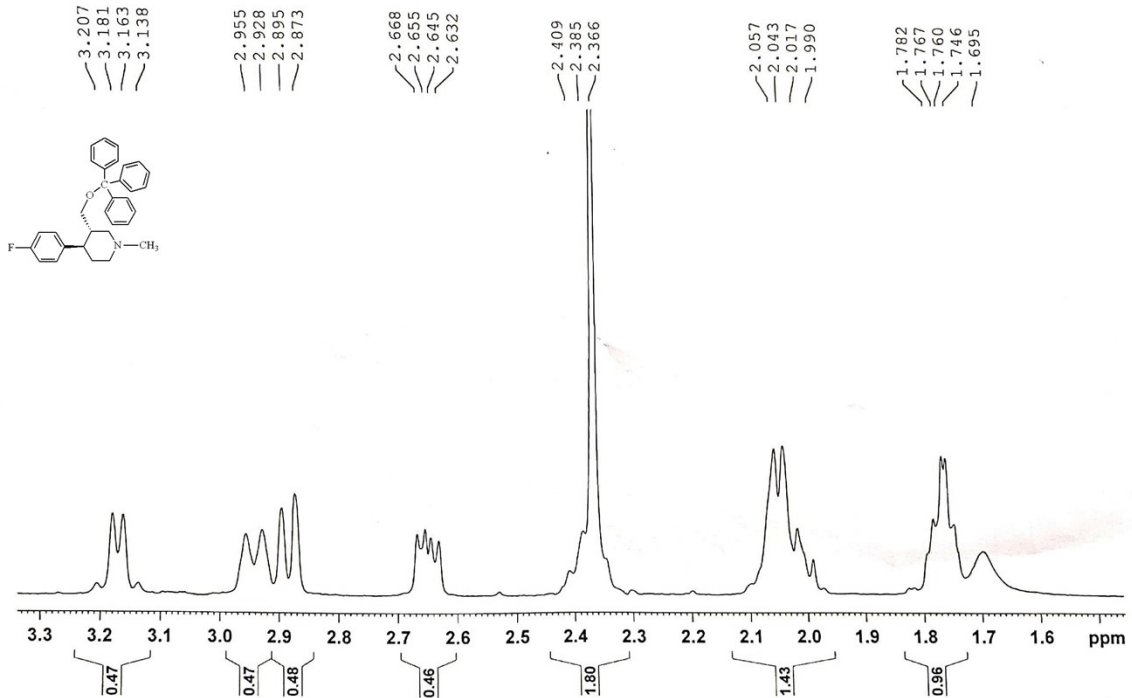
## IR Spectroscopy



## VM-7 NMR H1 Spectra

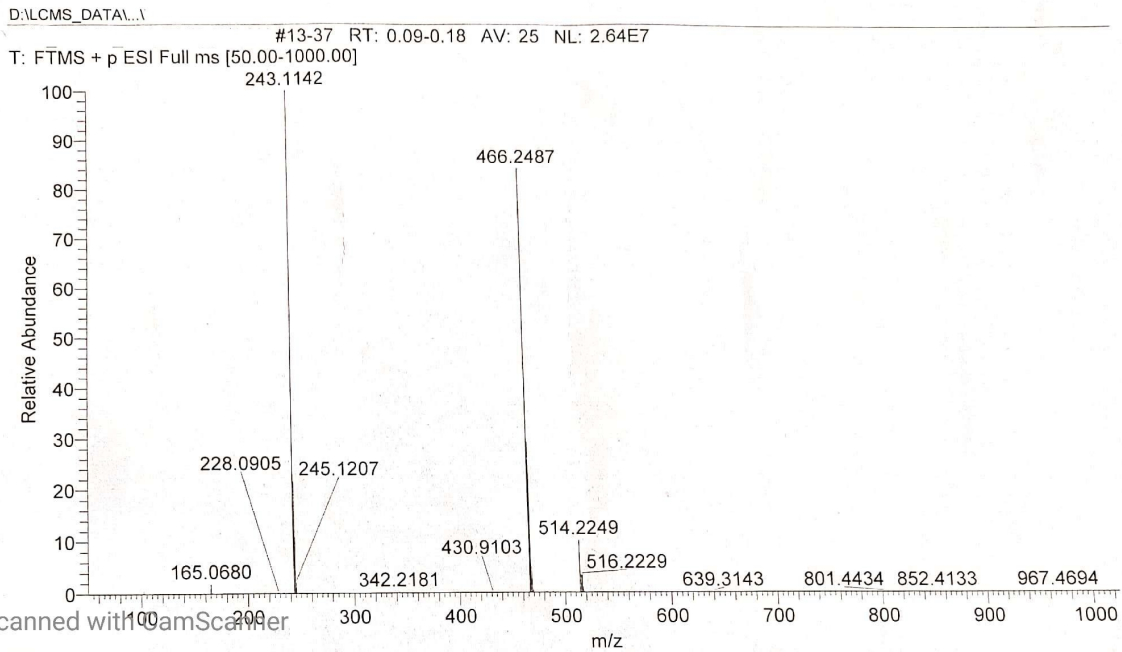


# VM-7 NMR H1 Spectra Expanded



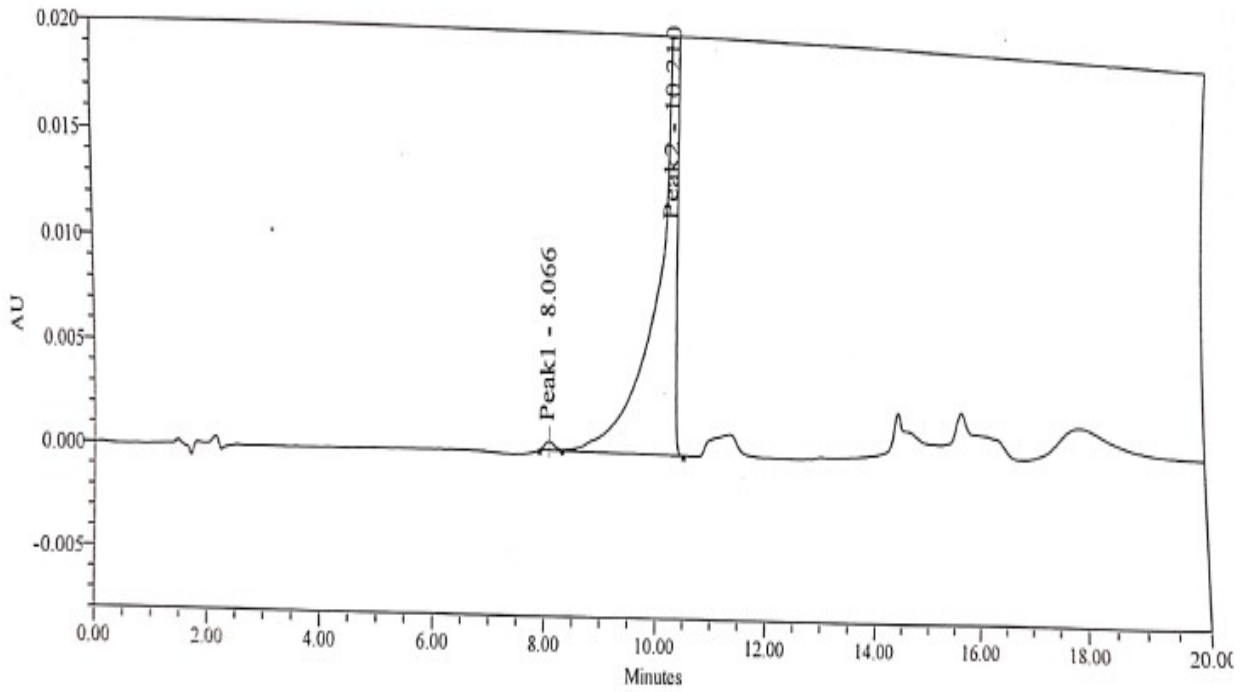
CS Scanned with CamScanner

# VM-7 Mass Spectra



CS Scanned with CamScanner

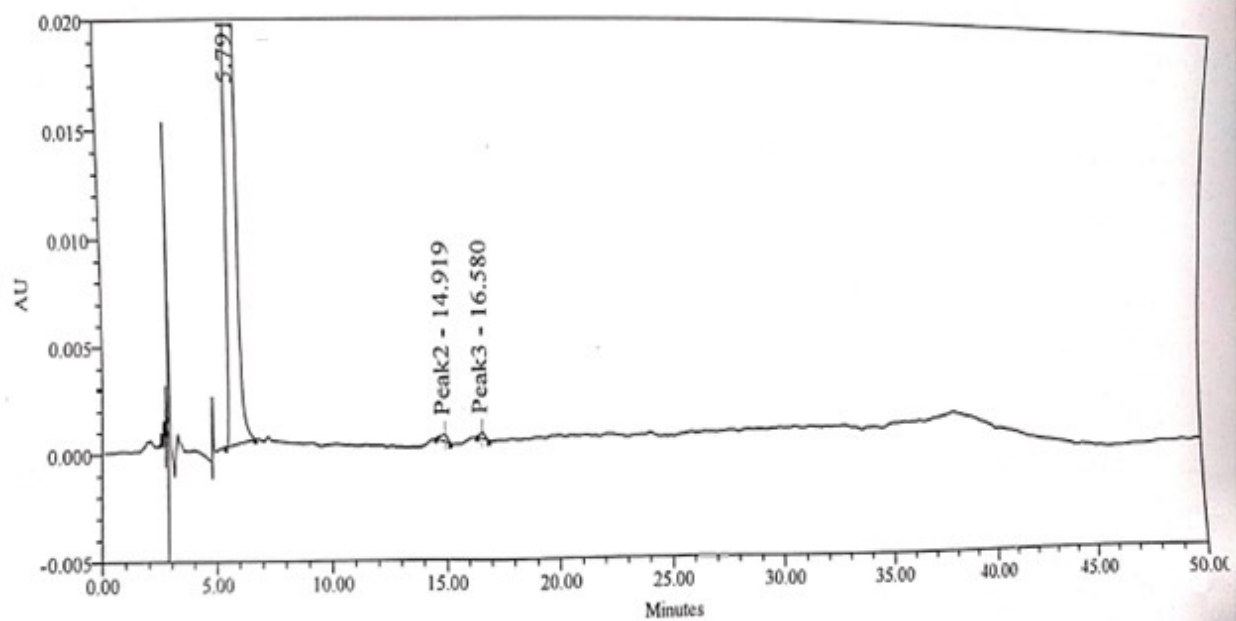
# HPLC Data of Compounds



Result # 1

Peak Results

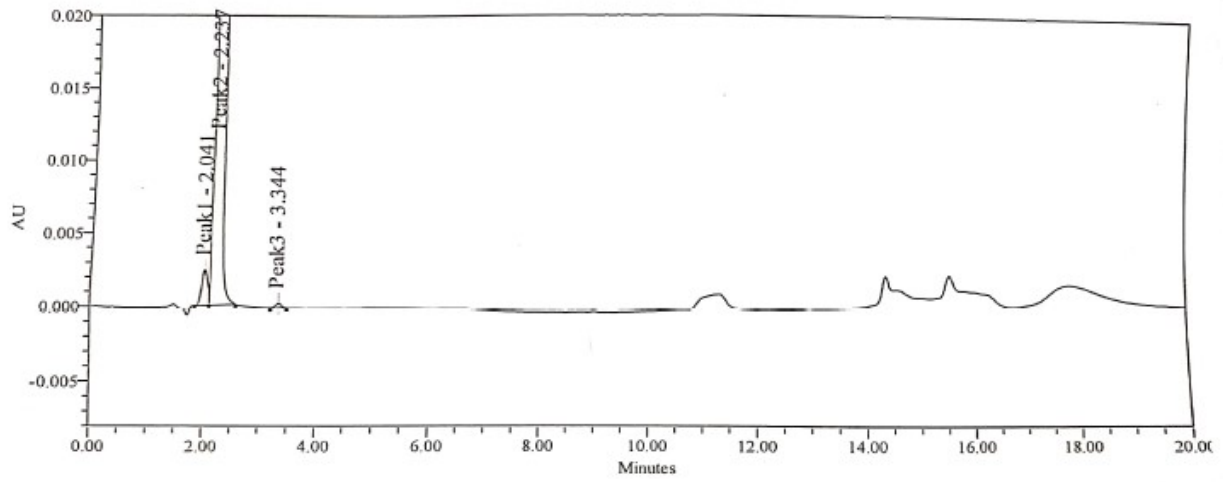
Name	Int Type	RT	RT Ratio	Area ( $\mu\text{V}^2\text{sec}$ )	% Area
1	Peak1	BV	8.07	4887	0.770
2	Peak2	VB	10.21	629835	99.230



Result # 2

Peak Results

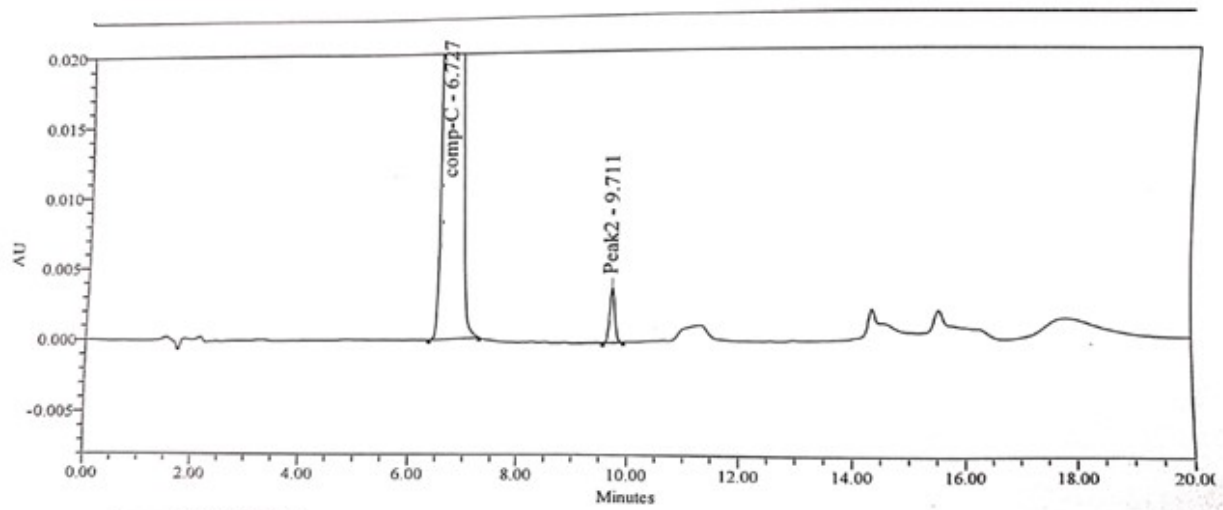
	Name	Int Type	RT	RT Ratio	Area (μV <sup>2</sup> sec)	% Area
1	Peak1	BB	5.79		11050398	99.904
2	Peak2	BB	14.92		6385	0.058
3	Peak3	BB	16.58		4255	0.038



Result # 1

Peak Results

	Name	Int Type	RT	RT Ratio	Area ( $\mu\text{V}\cdot\text{sec}$ )	% Area
1	Peak1	BV	2.04		17212	0.419
2	Peak2	VB	2.24		4088043	99.529
3	Peak3	BB	3.34		2131	0.052



— SampleName

Result # 1

**Peak Results**

	Name	Int Type	RT	RT Ratio	Area (μV*sec)	% Area
1	comp-C	BB	6.73		5076870	99.502
2	Peak2	BB	9.71		25414	0.498