Discovery of new nicotinamides as apoptotic VEGFR-2 inhibitors: virtual screening, synthesis, anti-proliferative, immunomodulatory, ADMET, toxicity, and molecular dynamic simulation studies

Reda G.Yousef[a], Albaraa Ibrahim[a], Mohamed M. Khalifa[a], Wagdy M. Eldehna[b,c], Ibraheem M. M. Gobaara[d], Ahmed B. M. Mehany[d], Eslam B. Elkaeed[e],Aisha A. Alsfouk[f], Ahmed M. Metwaly[g,h]* Ibrahim. H. Eissa[a]*,

[a] Pharmaceutical Medicinal Chemistry & Drug Design Department, Faculty of Pharmacy (Boys), Al-Azhar University, Cairo11884, Egypt.

[b] School of Biotechnology, Badr University in Cairo, Badr City, Cairo 11829, Egypt

[c]Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Kafrelsheikh University, P.O. Box 33516, Kafrelsheikh, Egypt.

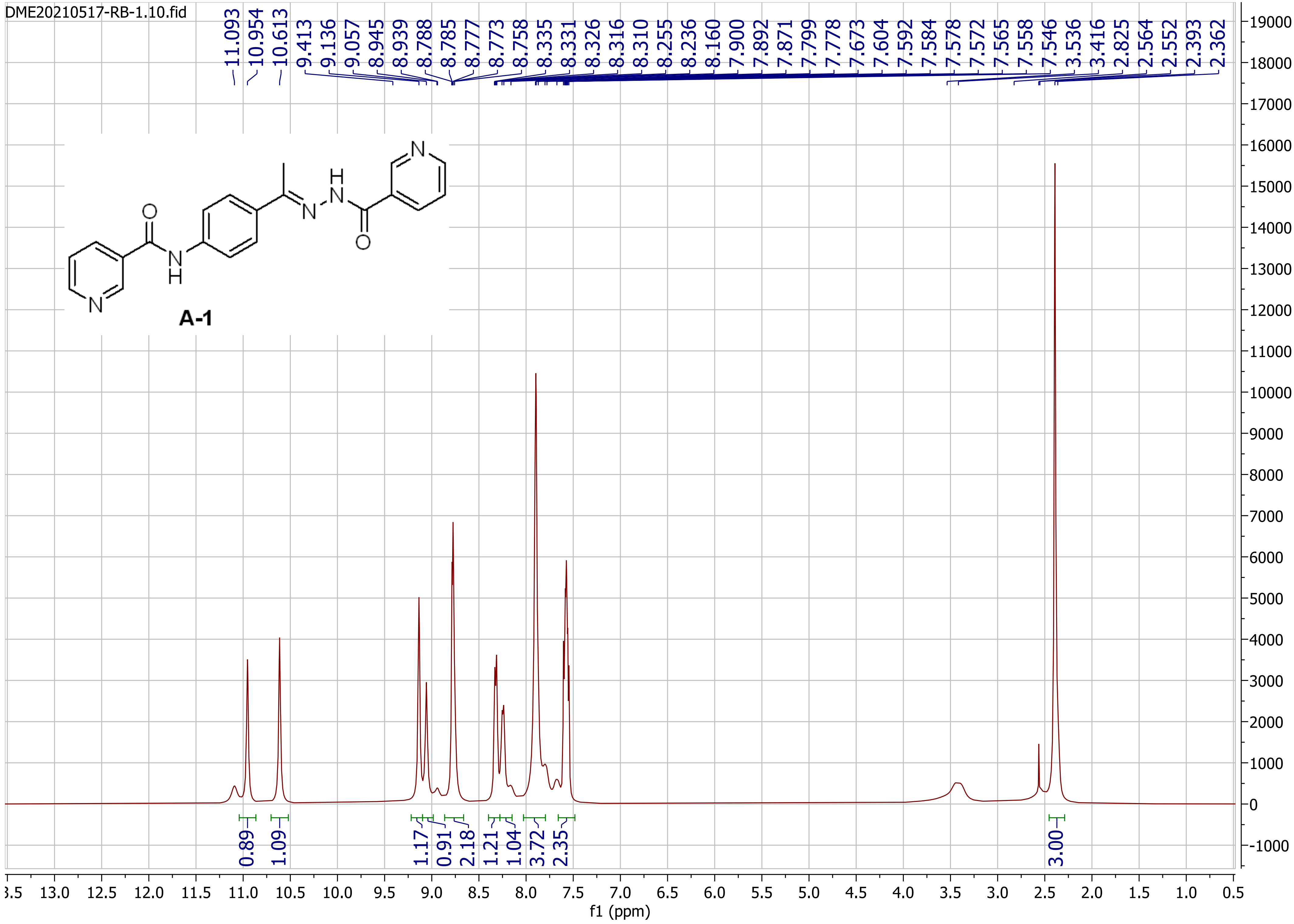
[d] Zoology Department, Faculty of Science (Boys), Al-Azhar University, Cairo11884, Egypt.

[e] Department of Pharmaceutical Sciences, College of Pharmacy, AlMaarefa University, Riyadh 13713, Saudi Arabia.
[f] Department of Pharmaceutical Sciences, College of Pharmacy, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia.

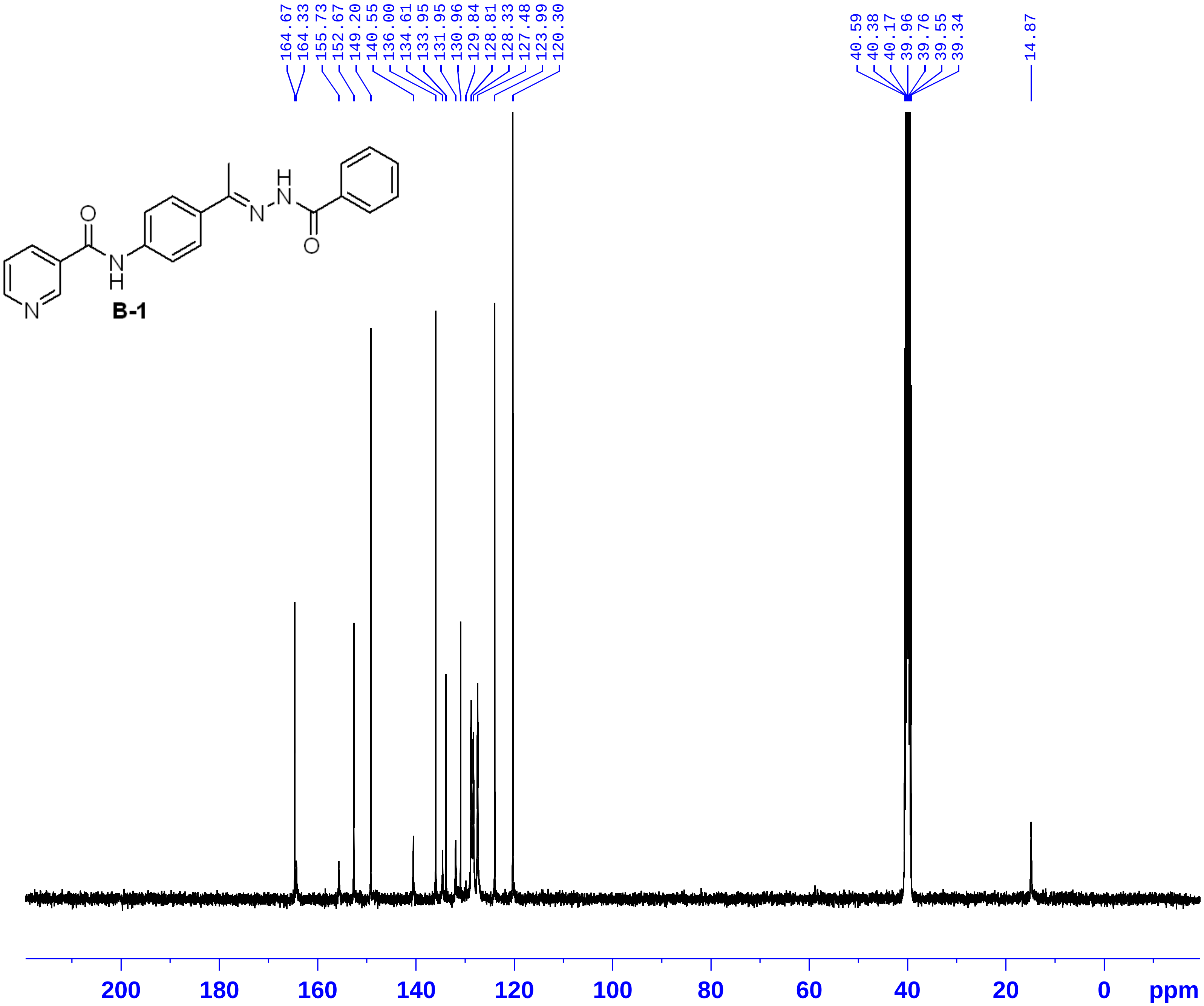
[g] Pharmacognosy and Medicinal Plants Department, Faculty of Pharmacy (Boys), Al-Azhar University, Cairo 11884, Egypt.
[h]Biopharmaceutical Products Research Department, Genetic Engineering and Biotechnology Research Institute, City of Scientific Research and Technological Applications (SRTA-City), Alexandria, Egypt

***Corresponding authors:**
Ibrahim H. Eissa
Medicinal Chemistry Department, Faculty of Pharmacy (Boys), Al-Azhar University, Cairo 11884, Egypt. **Email:** Ibrahimeissa@azhar.edu.eg

Ahmed M. Metwaly
Pharmacognosy and Medicinal Plants Department, Faculty of Pharmacy (Boys), Al-Azhar University, Cairo 11884, Egypt. **Email:** ametwaly@azhar.edu.eg


A-1
DME20210517-RB-1.10.fid
11.093
10.954
10.613
9.413
9.136
9.057
8.945
8.939
8.788
8.785
8.777
8.773
8.758
8.335
8.331
8.326
8.316
8.310
8.255
8.236
8.160
7.900
7.892
7.871
7.799
7.778
7.673
7.604
7.592
7.584
7.578
7.572
7.565
7.558
7.546
3.536
3.416
2.825
2.564
2.552
2.393
2.362
0.89
1.09
1.17
0.91
2.18
1.21
1.04
3.72
2.35
3.00
f1 (ppm)
13.5
13.0
12.5
12.0
11.5
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
19000
18000
17000
16000
15000
14000
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000
O
N
H
N
H
N
N
O
N

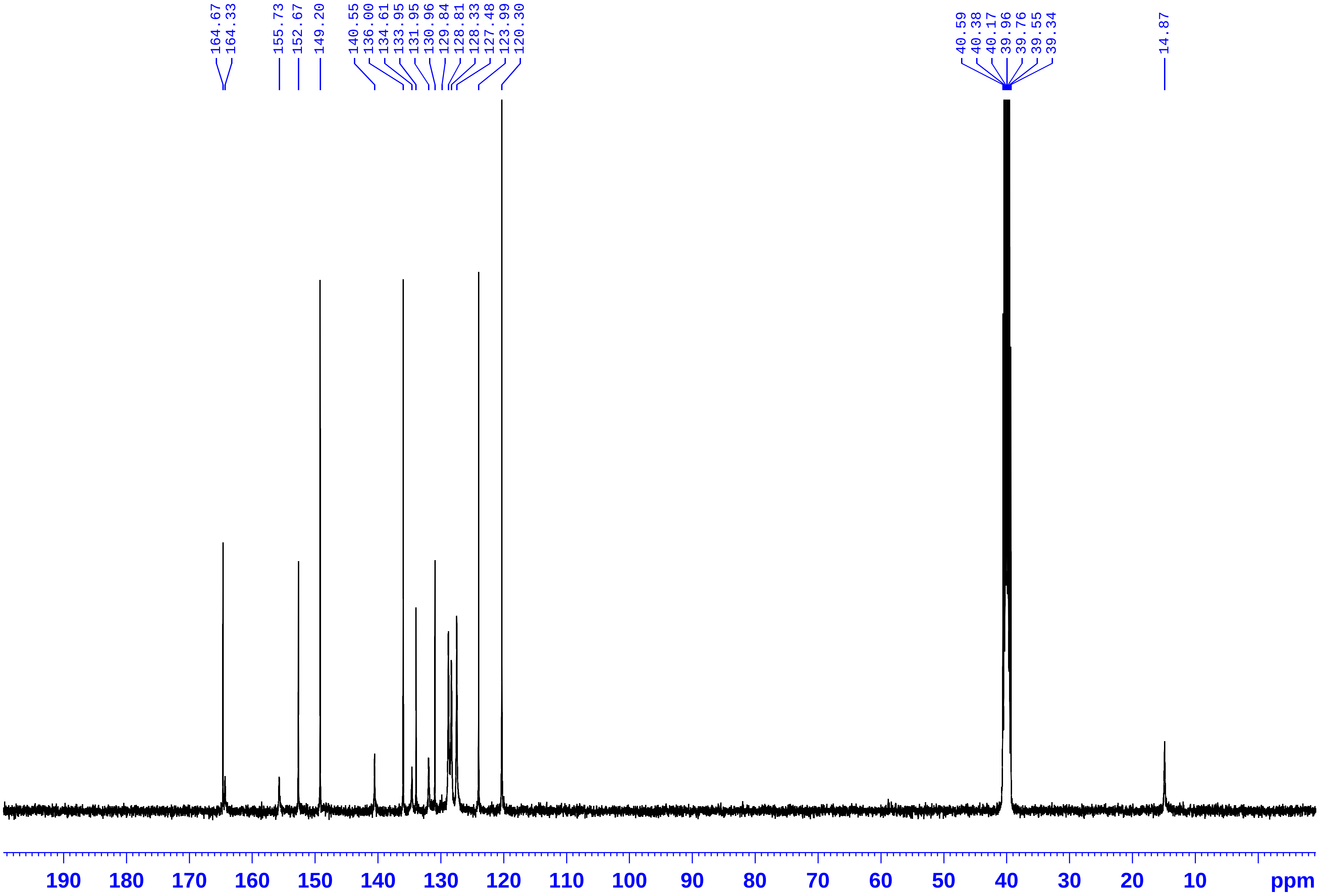
Wagdy eldehna -RB4-carbon-Es



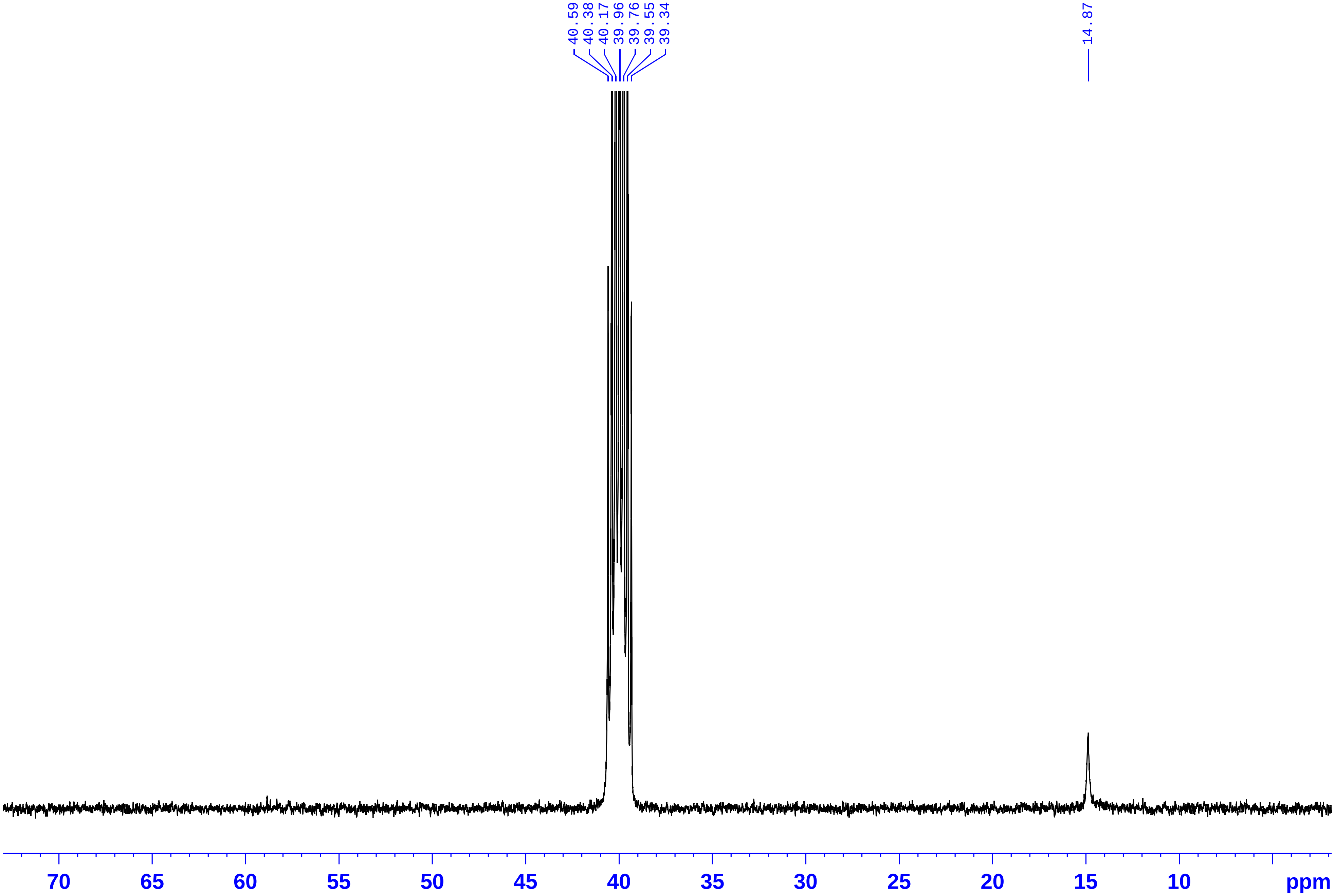
B-1

```
Current Data Parameters
NAME     Wagdy eldehna -RB4-carbon-Es
EXPNO                     10
PROCNO                     1

F2 - Acquisition Parameters
Date_               20210627
Time                   19.58 h
INSTRUM                spect
PROBHD    Z108618_0945 (
PULPROG              zgpg30
TD                     65536
SOLVENT                 DMSO
NS                      2200
DS                         4
SWH                24038.461 Hz
FIDRES              0.733596 Hz
AQ                 1.3631488 sec
RG                    197.77
DW                    20.800 usec
DE                      6.50 usec
TE                     297.5 K
D1                2.00000000 sec
D11               0.03000000 sec
TD0                        1
SFO1             100.6404331 MHz
NUC1                     13C
P1                     10.00 usec
PLW1             47.00000000 W
SFO2             400.2016008 MHz
NUC2                      1H
CPDPRG[2             waltz16
PCPD2                  90.00 usec
PLW2             13.00000000 W
PLW12             0.29249999 W
PLW13             0.14713000 W

F2 - Processing parameters
SI                     32768
SF               100.6303700 MHz
WDW                       EM
SSB                        0
LB                      1.00 Hz
GB                         0
PC                      1.40
```

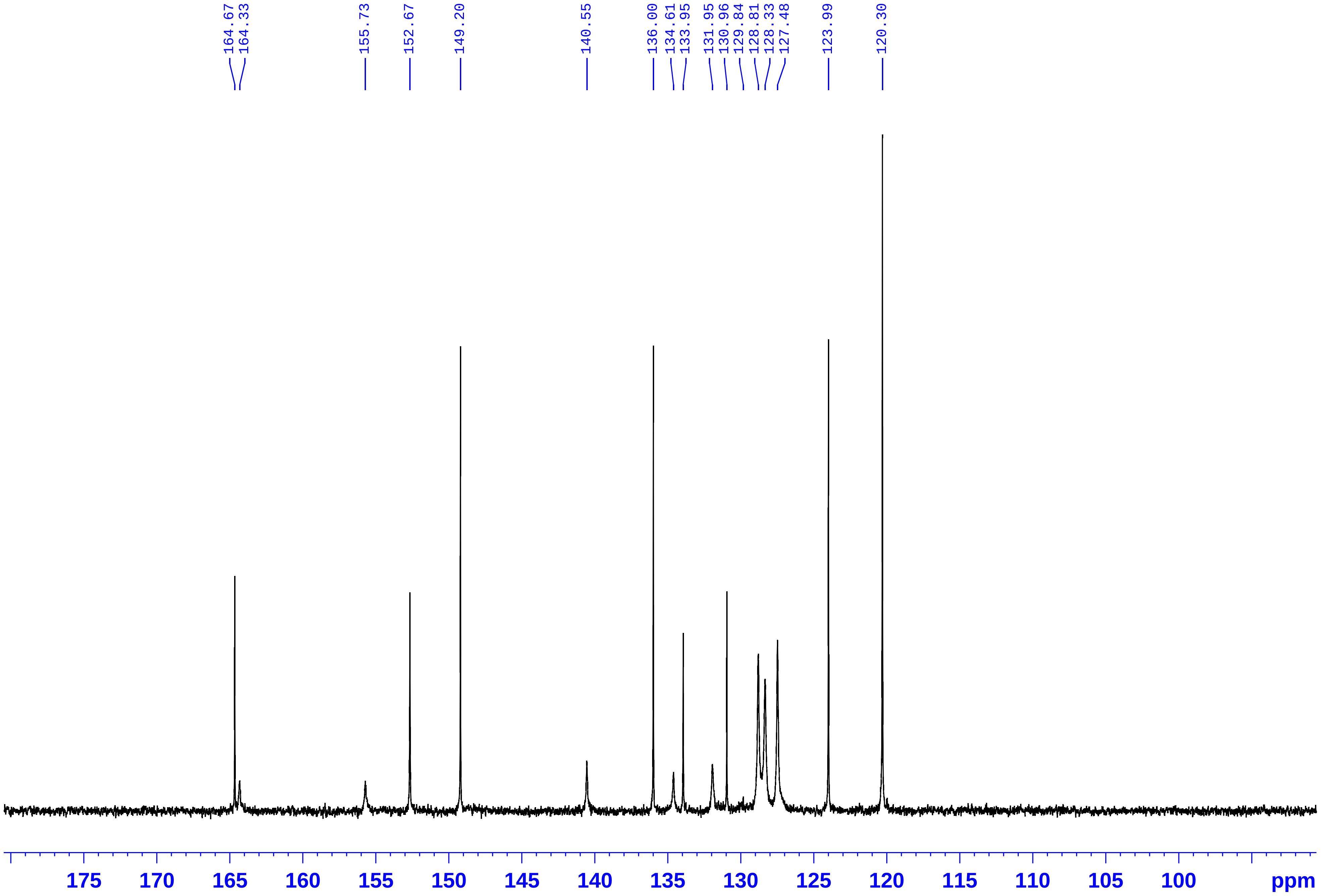
Wagdy eldehna -RB4-carbon-Es



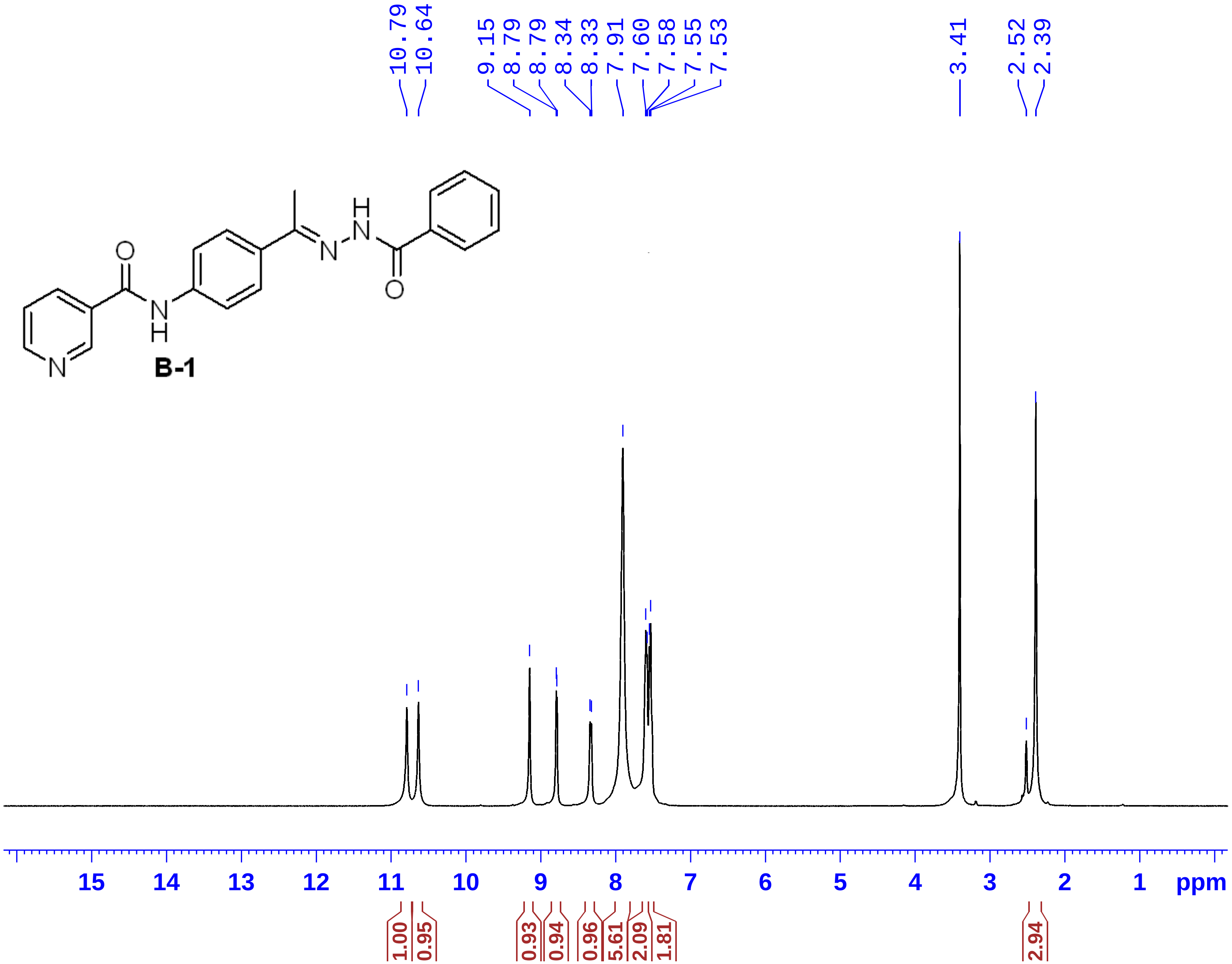
Wagdy eldehna -RB4-carbon-Es



Wagdy eldehna -RB4-carbon-Es



Wagdy El Dahna-RB 4-DMSO-Hnmr-A



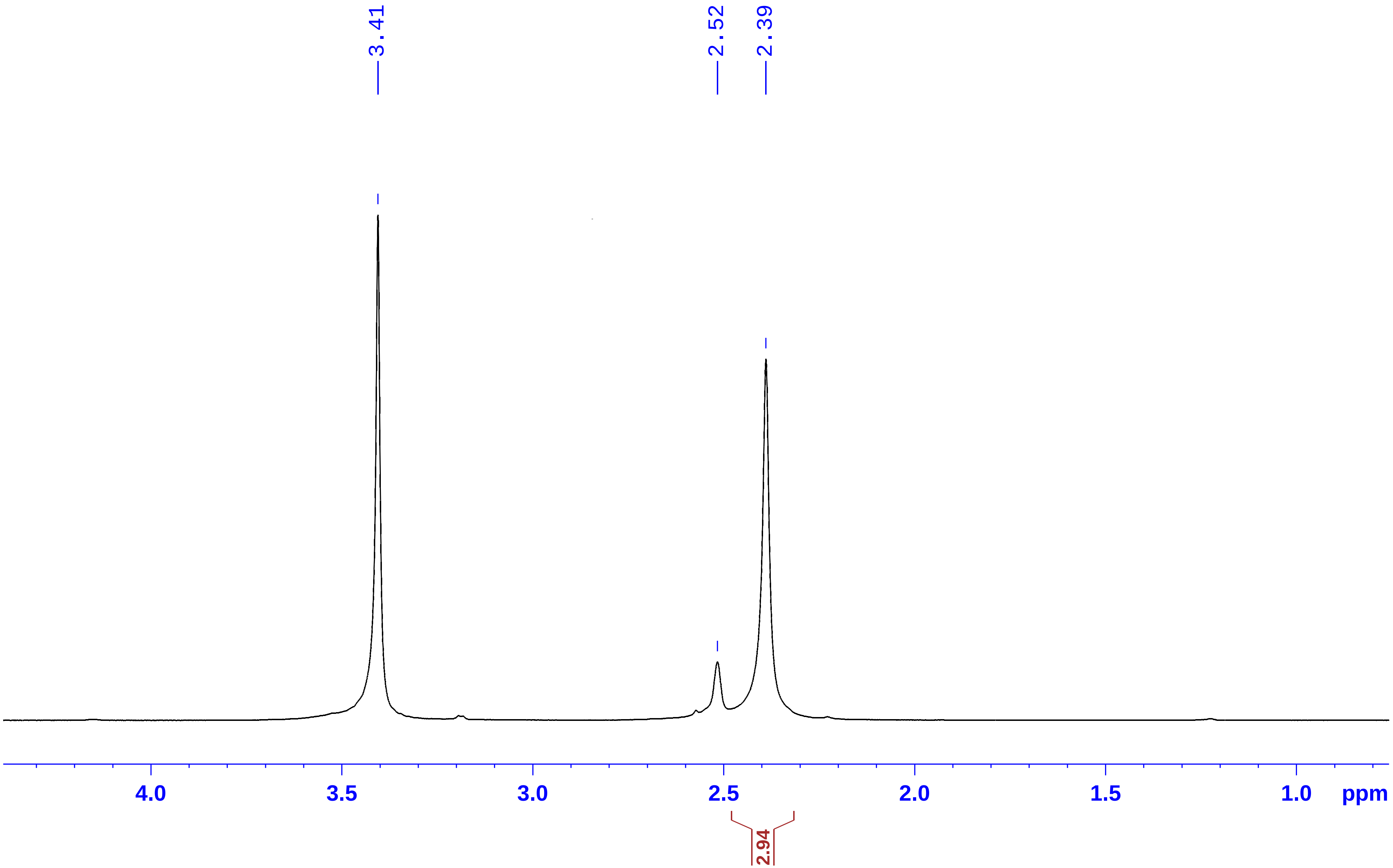
B-1

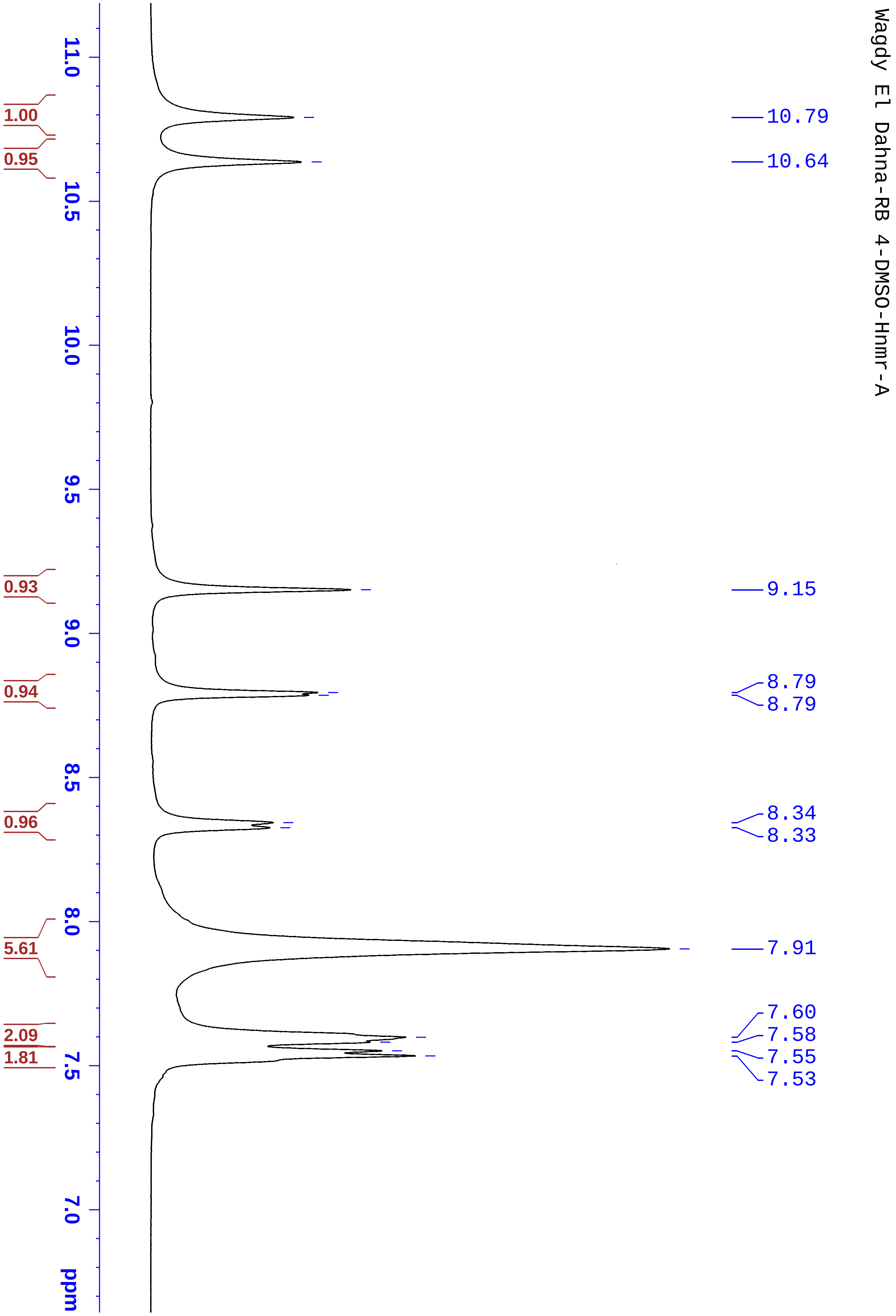
BRUKER

Faculty of Pharmacy

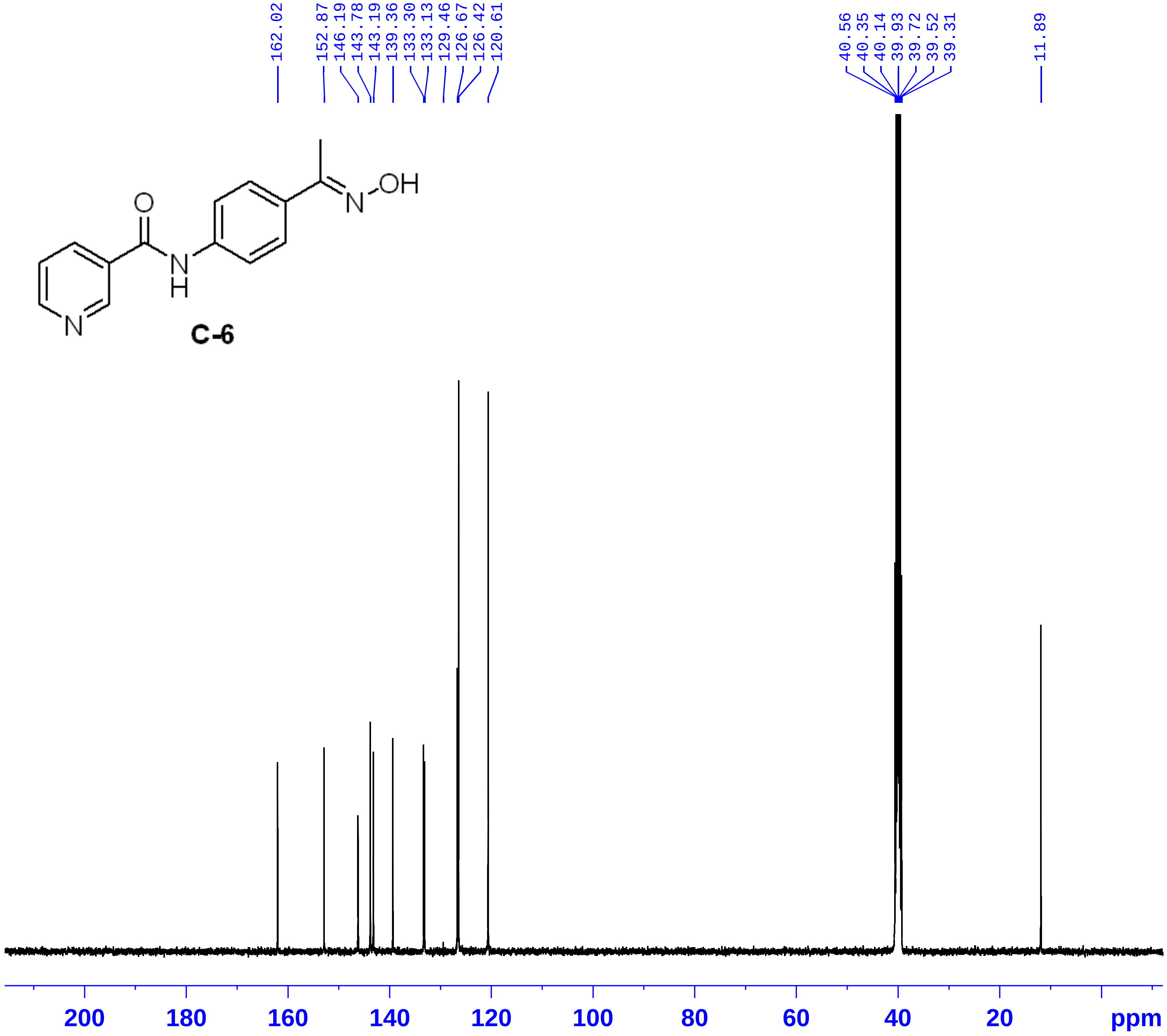
```
Current Data Parameters
NAME       Wagdy El Dahna-RB
EXPNO                     10
PROCNO                     1

F2 - Acquisition Parameters
Date_               20210610
Time                   11.19 h
INSTRUM                spect
PROBHD     Z108618_0945 (
PULPROG                 zg30
TD                     65536
SOLVENT                 DMSO
NS                        16
DS                         2
SWH                 8012.820 Hz
FIDRES              0.244532 Hz
AQ                 4.0894465 se
RG                     88.92
DW                    62.400 use
DE                      6.50 use
TE                     294.5 K
D1                1.00000000 se
TD0                        1
SFO1             400.2024712 MH
NUC1                      1H
P1                     13.50 use
PLW1             13.00000000 W

F2 - Processing parameters
SI                     65536
SF               400.2000000 MH
WDW                       EM
SSB      0
LB                      0.30 Hz
GB       0
PC                      1.00
```

Wagdy El Dahna-RB 4-DMSO-Hnmr-A




Wagdy El Dahna-RB 4-DMSO-Hnmr-A
10.79
10.64
9.15
8.79
8.79
8.34
8.33
7.91
7.60
7.58
7.55
7.53
1.00
0.95
0.93
0.94
0.96
5.61
2.09
1.81
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
ppm

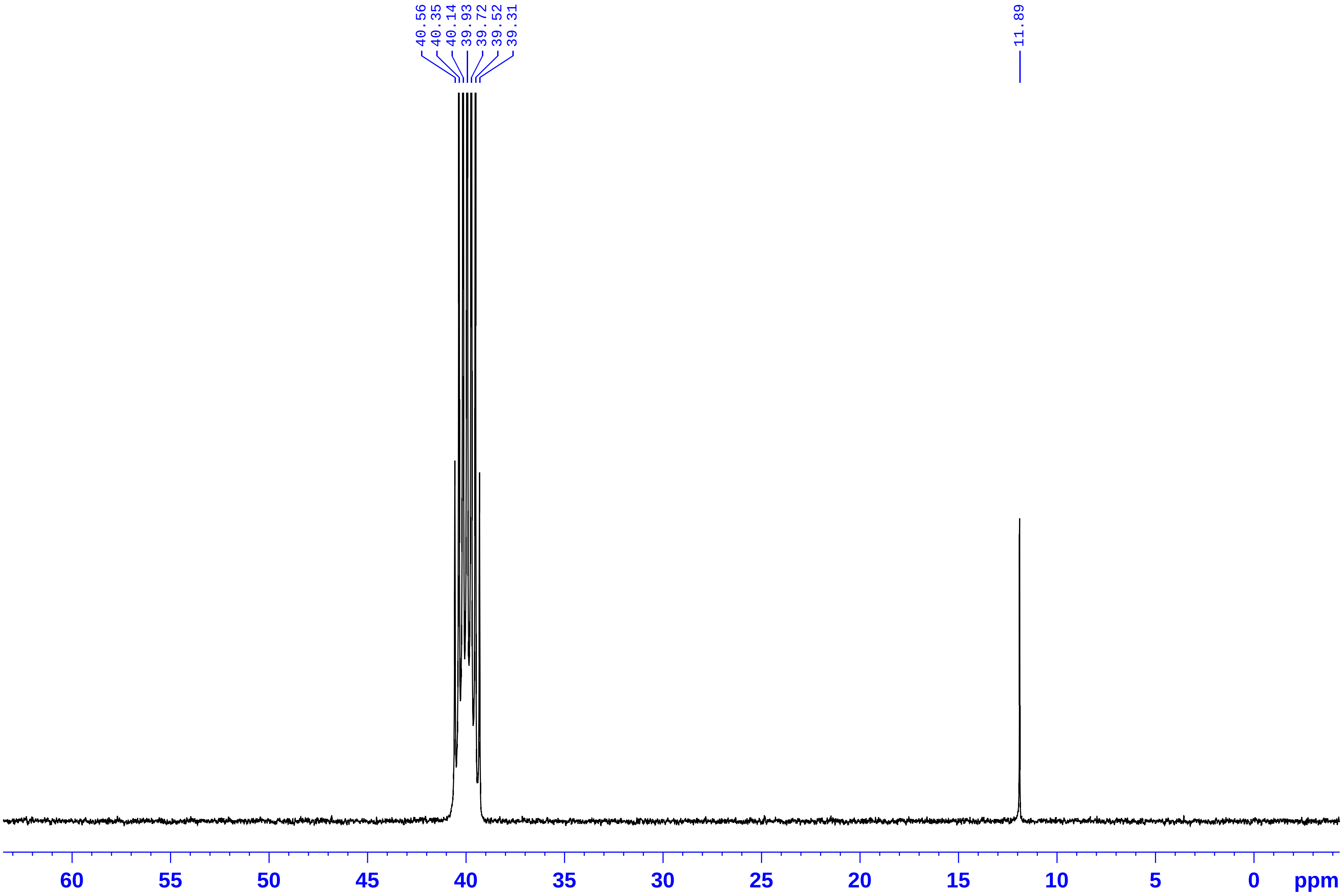
Wagdy eldehna -RB5-carbon-Es



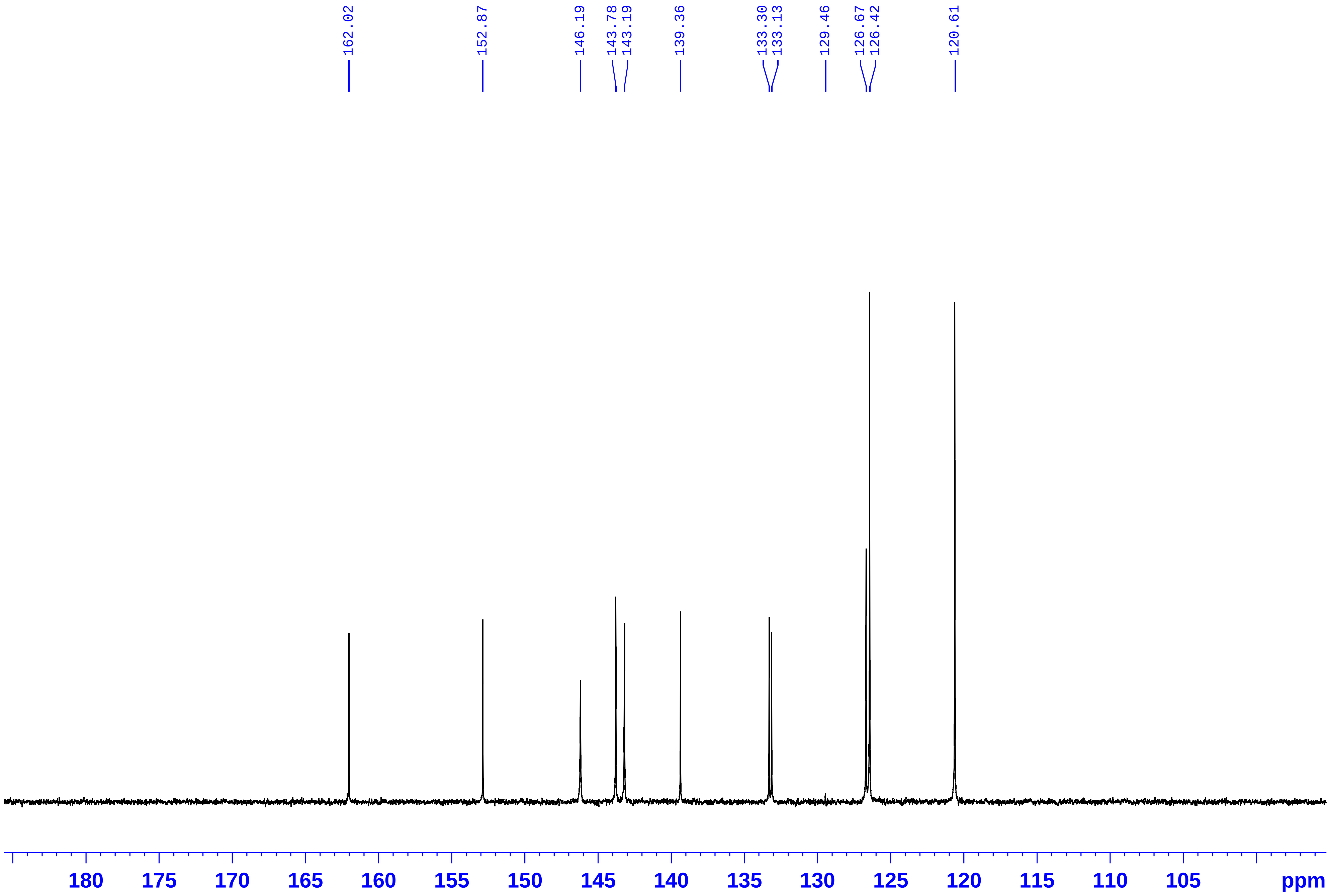
C-6

```
Current Data Parameters
NAME     Wagdy eldehna -RB5-carbon-Es
EXPNO                    10
PROCNO                    1

F2 - Acquisition Parameters
Date_              20210628
Time                   6.44 h
INSTRUM               spect
PROBHD   Z108618_0945 (
PULPROG             zgpg30
TD                    65536
SOLVENT                DMSO
NS                     2200
DS                        4
SWH               24038.461 Hz
FIDRES             0.733596 Hz
AQ                1.3631488 sec
RG                   197.77
DW                   20.800 usec
DE                     6.50 usec
TE                    297.4 K
D1               2.00000000 sec
D11              0.03000000 sec
TD0                       1
SFO1            100.6404331 MHz
NUC1                    13C
P1                    10.00 usec
PLW1            47.00000000 W
SFO2            400.2016008 MHz
NUC2                     1H
CPDPRG[2            waltz16
PCPD2                 90.00 usec
PLW2            13.00000000 W
PLW12            0.29249999 W
PLW13            0.14713000 W

F2 - Processing parameters
SI                    32768
SF              100.6303700 MHz
WDW                      EM
SSB                       0
LB                     1.00 Hz
GB                        0
PC                     1.40
```

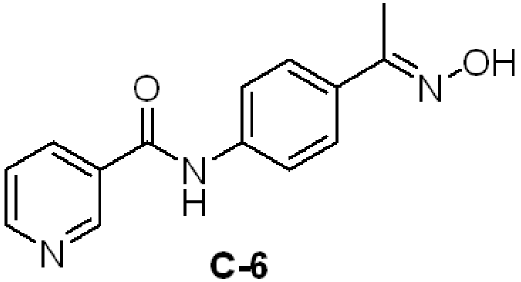
Wagdy eldehna -RB5-carbon-Es



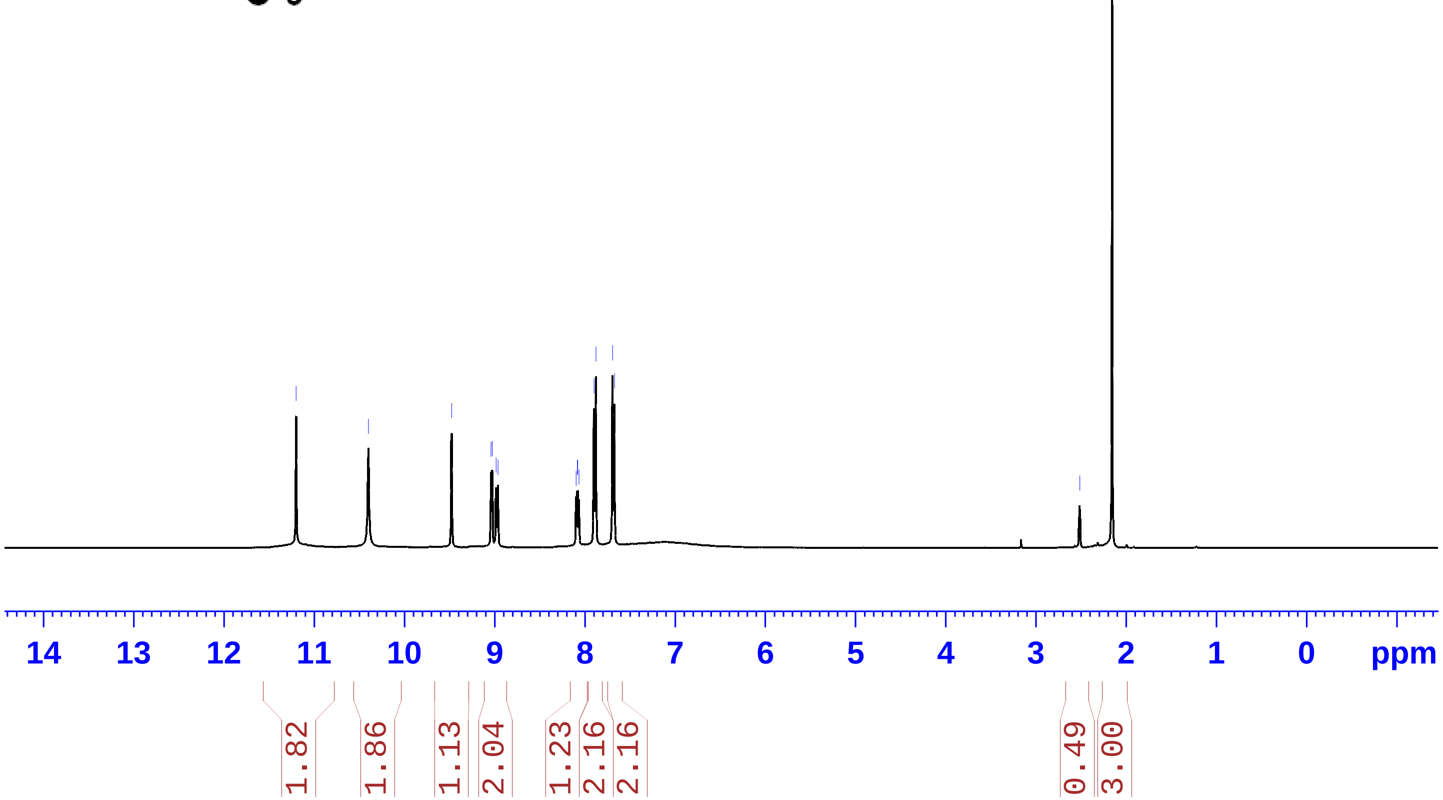
Wagdy eldehna -RB5-carbon-Es



Wagdy El Dahna-RB 5-DMSO-Hnmr-A

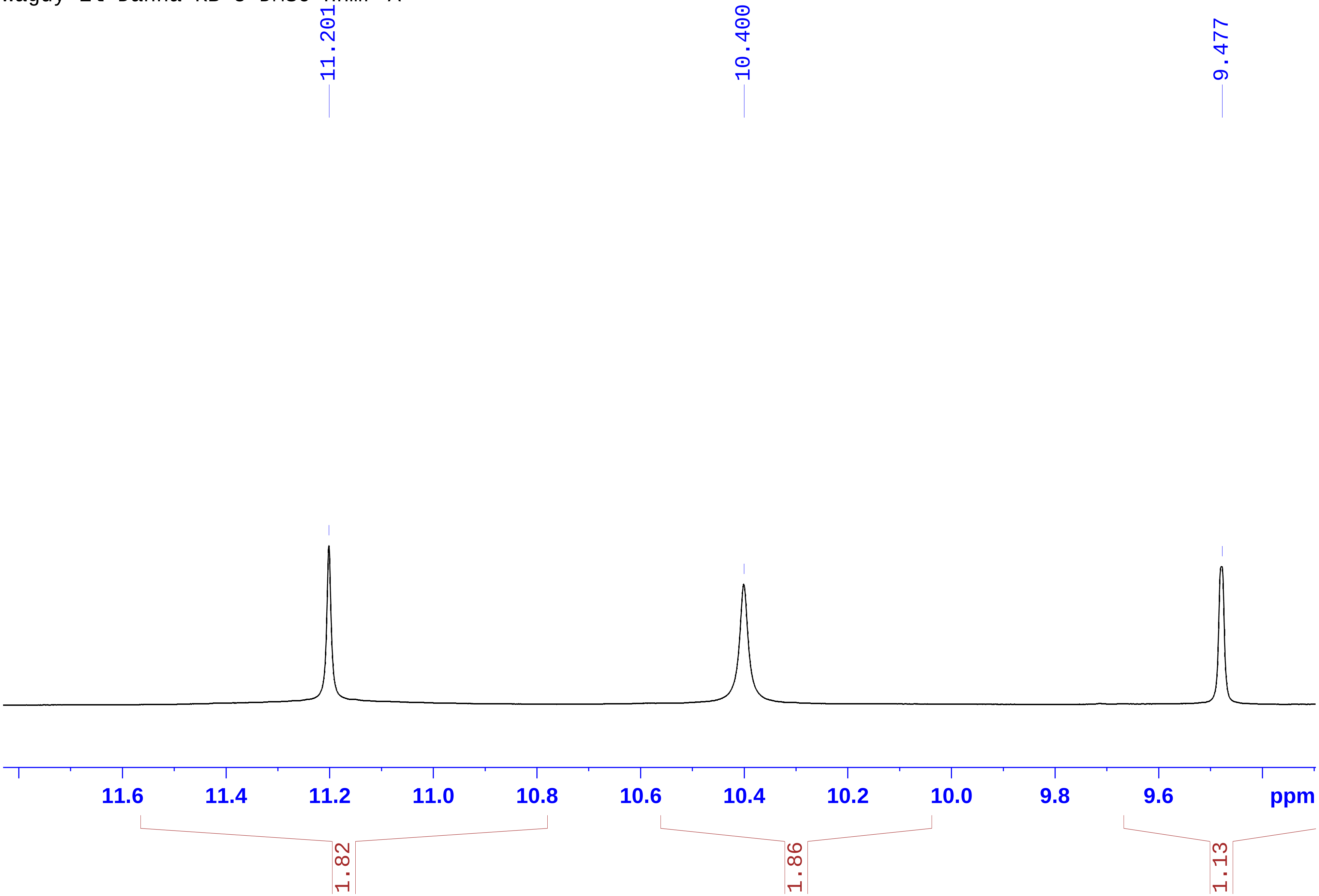


C-6

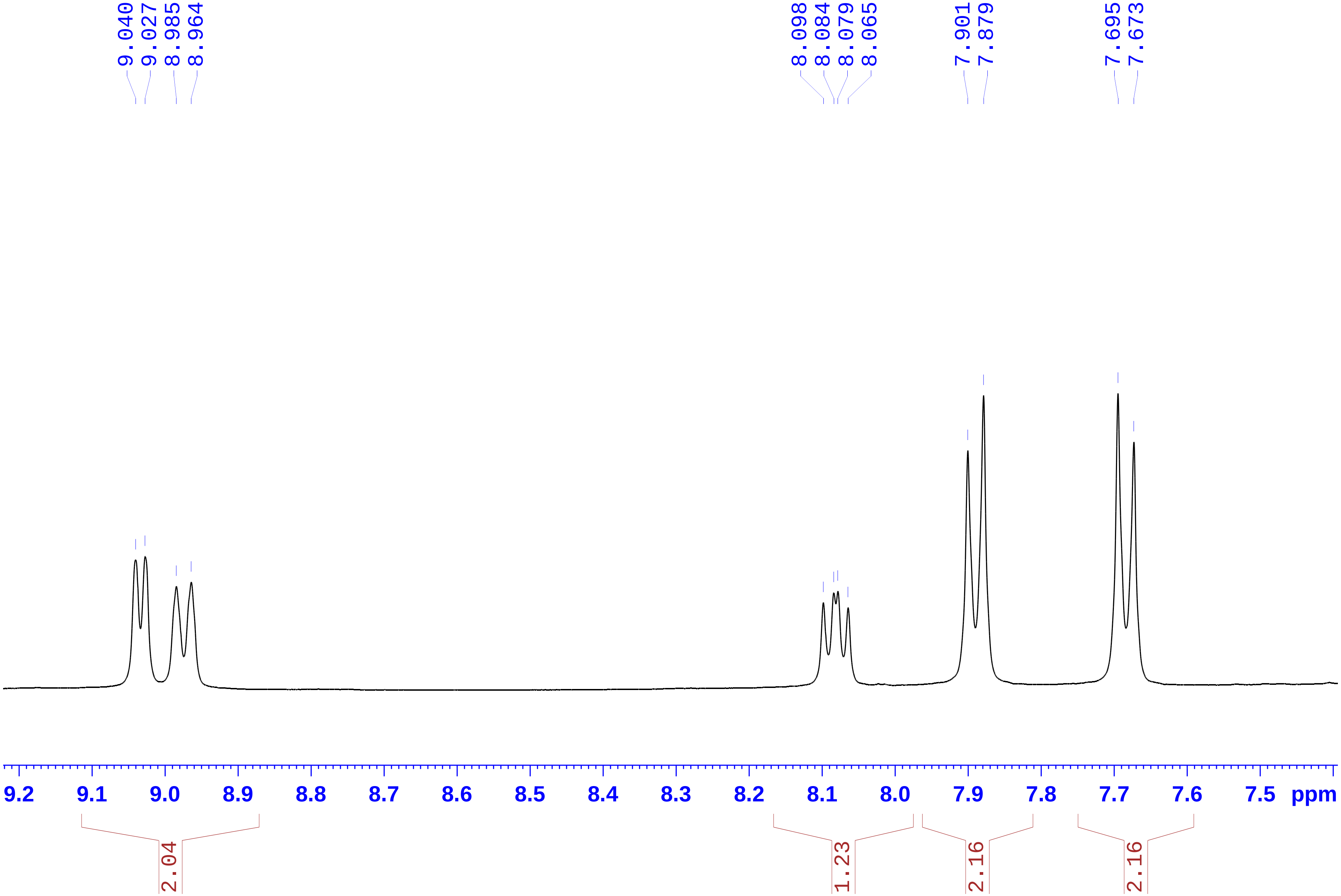


```
Current Data Parameters
NAME       Wagdy El Dahna-RB 5-DMSO-H
EXPNO                    10
PROCNO                    1

F2 - Acquisition Parameters
Date_              20210610
Time                  11.43 h
INSTRUM               spect
PROBHD     Z108618_0945 (
PULPROG                zg30
TD                    65536
SOLVENT                DMSO
NS                       16
DS                        2
SWH                8012.820 Hz
FIDRES             0.244532 Hz
AQ                4.0894465 sec
RG                     99.3
DW                   62.400 usec
DE                     6.50 usec
TE                    294.3 K
D1               1.00000000 sec
TD0                       1
SFO1            400.2024712 MHz
NUC1                     1H
P1                    13.50 usec
PLW1            13.00000000 W

F2 - Processing parameters
SI                    65536
SF              400.2000000 MHz
WDW                      EM
SSB                       0
LB                     0.30 Hz
GB                        0
PC                     1.00
```

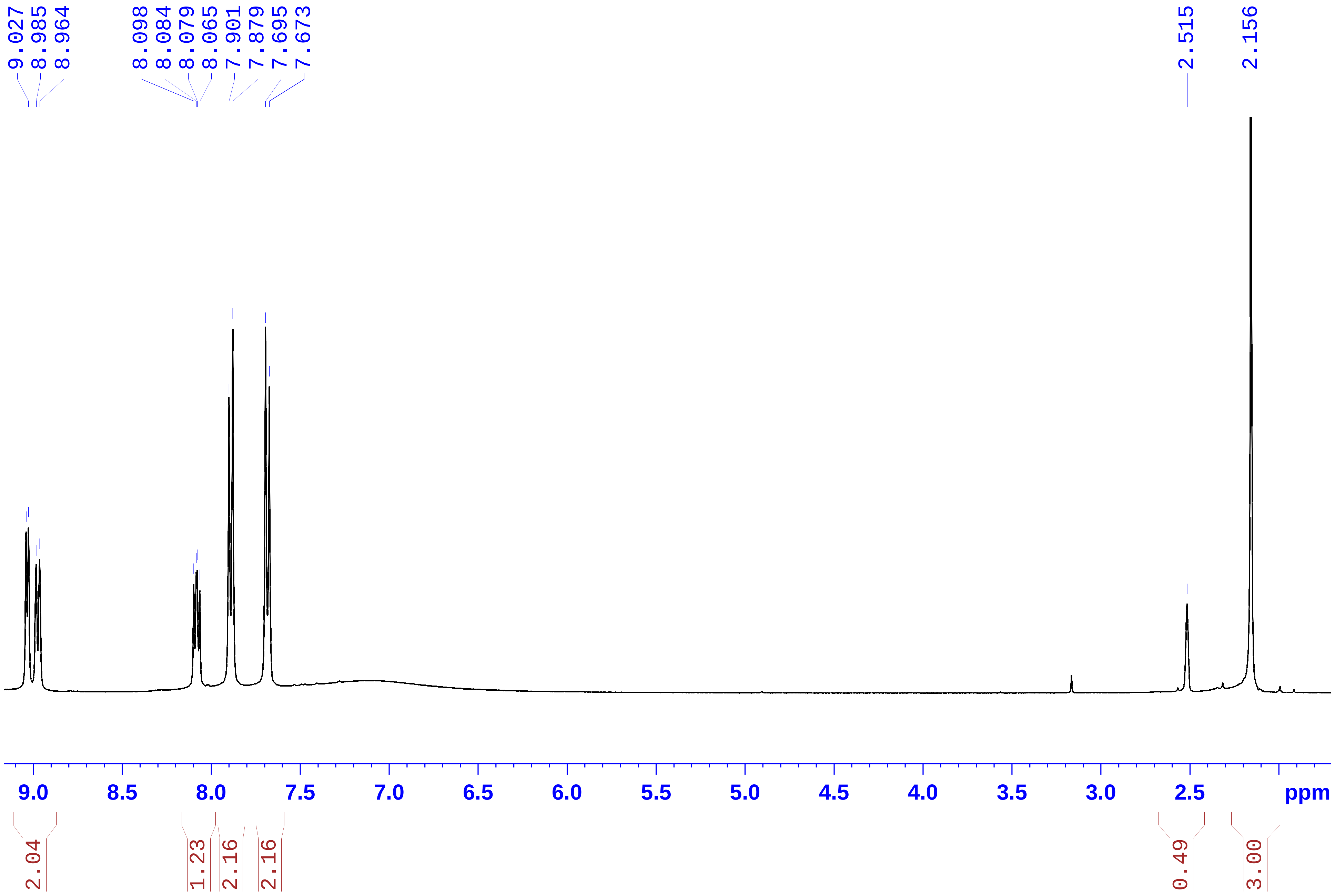
Wagdy El Dahna-RB 5-DMSO-Hnmr-A



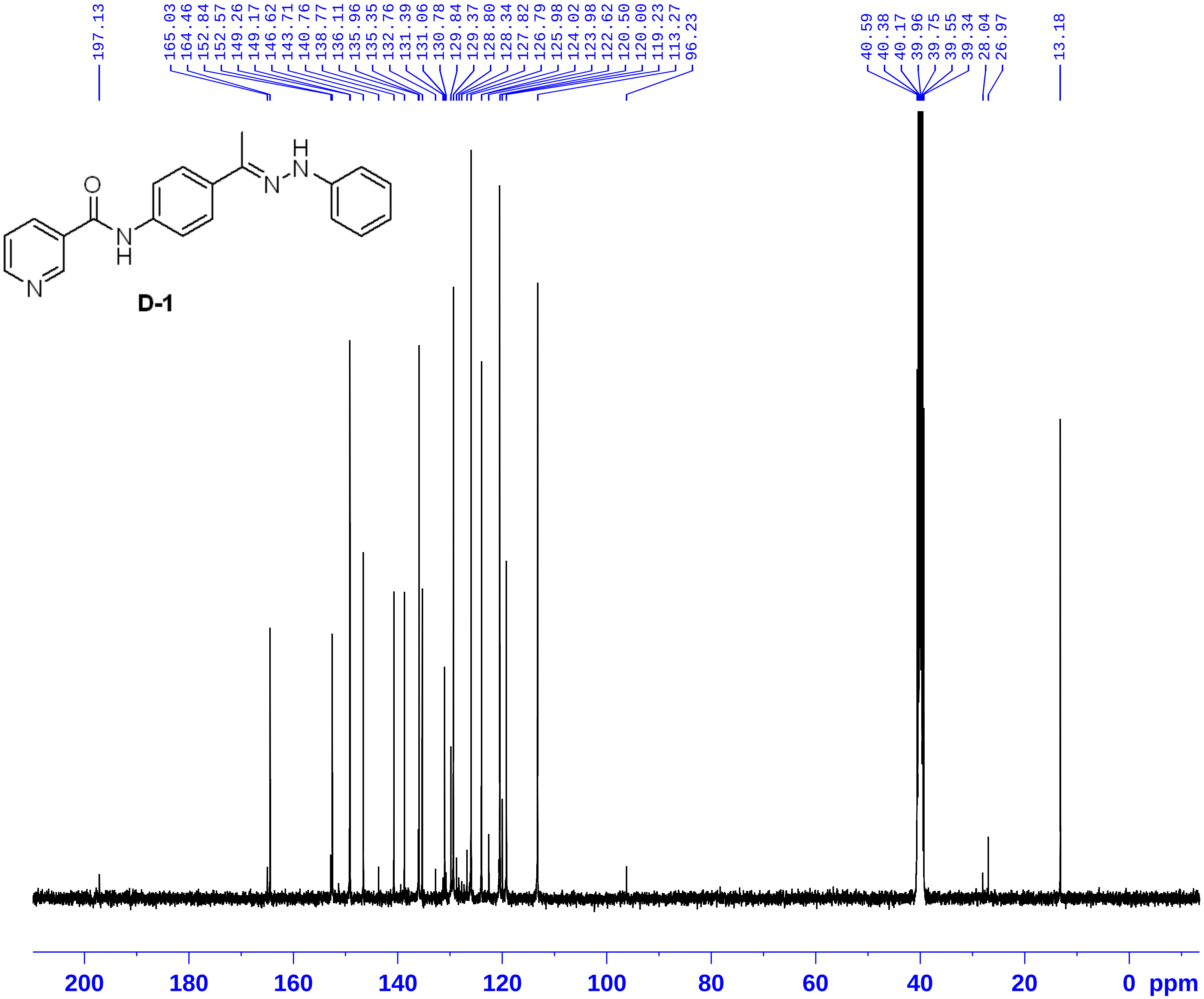
Wagdy El Dahna-RB 5-DMSO-Hnmr-A



Wagdy El Dahna-RB 5-DMSO-Hnmr-A



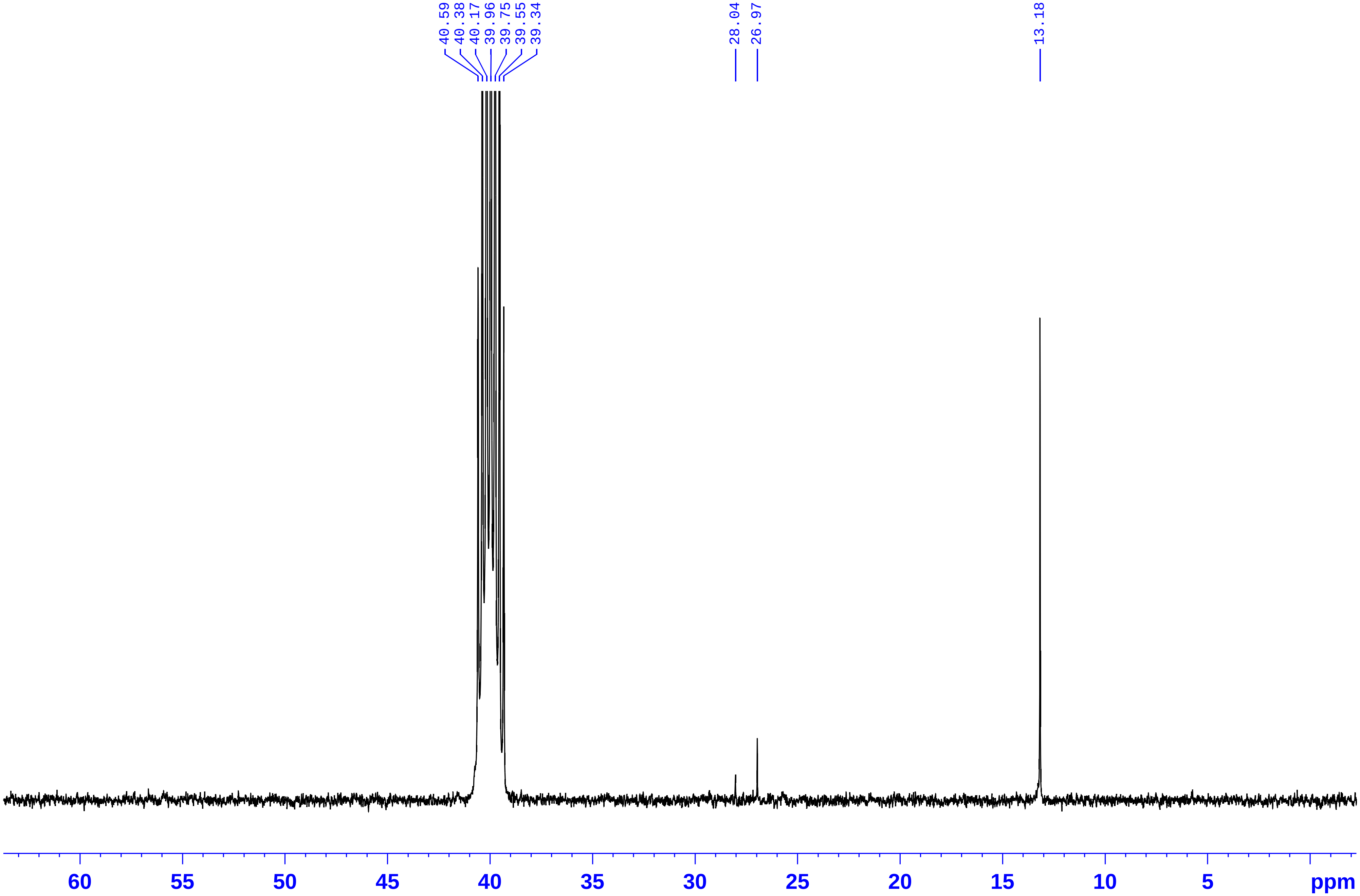
Wagdy eldehna -RB8-carbon-Es

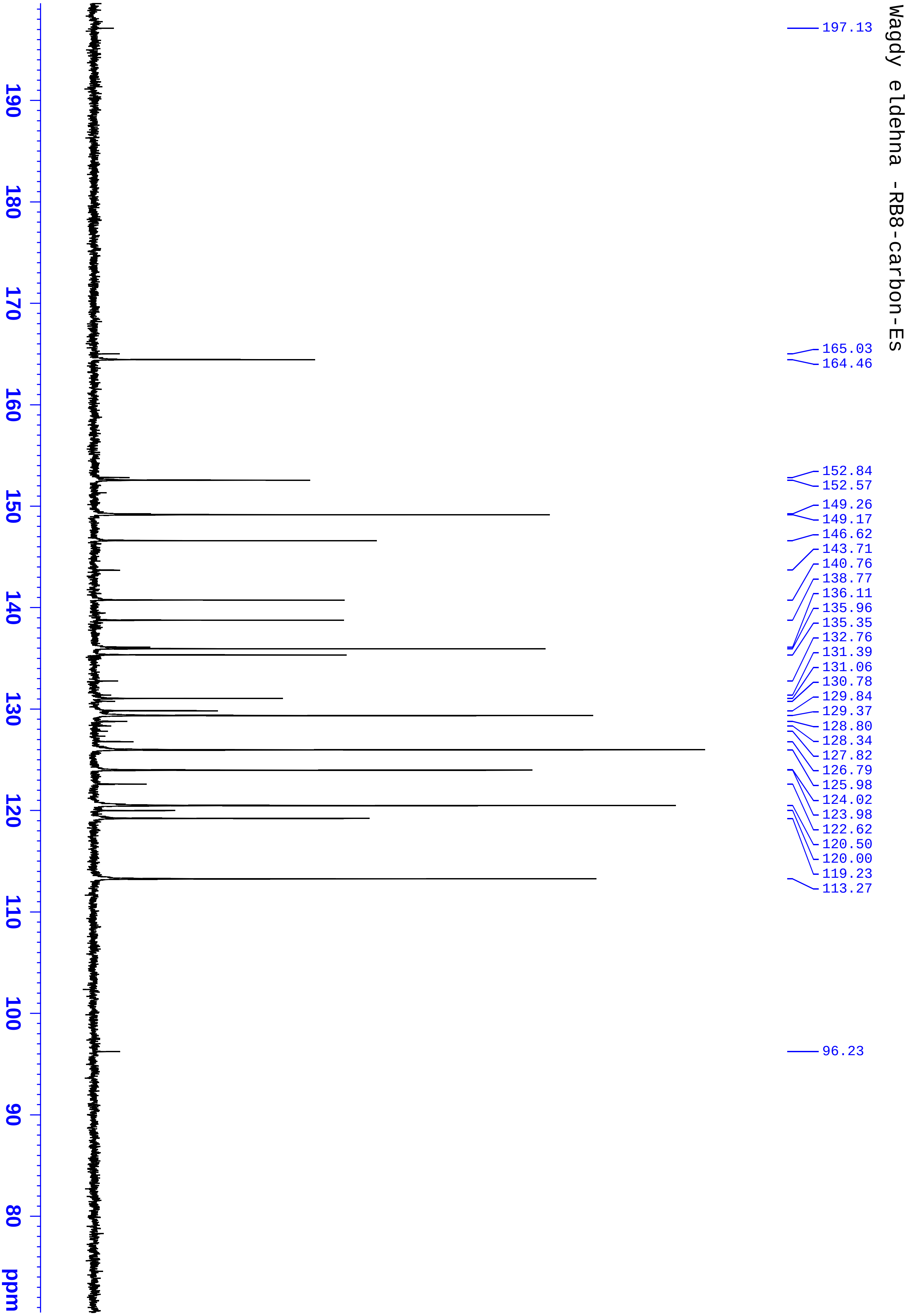


Faculty of Pharmacy

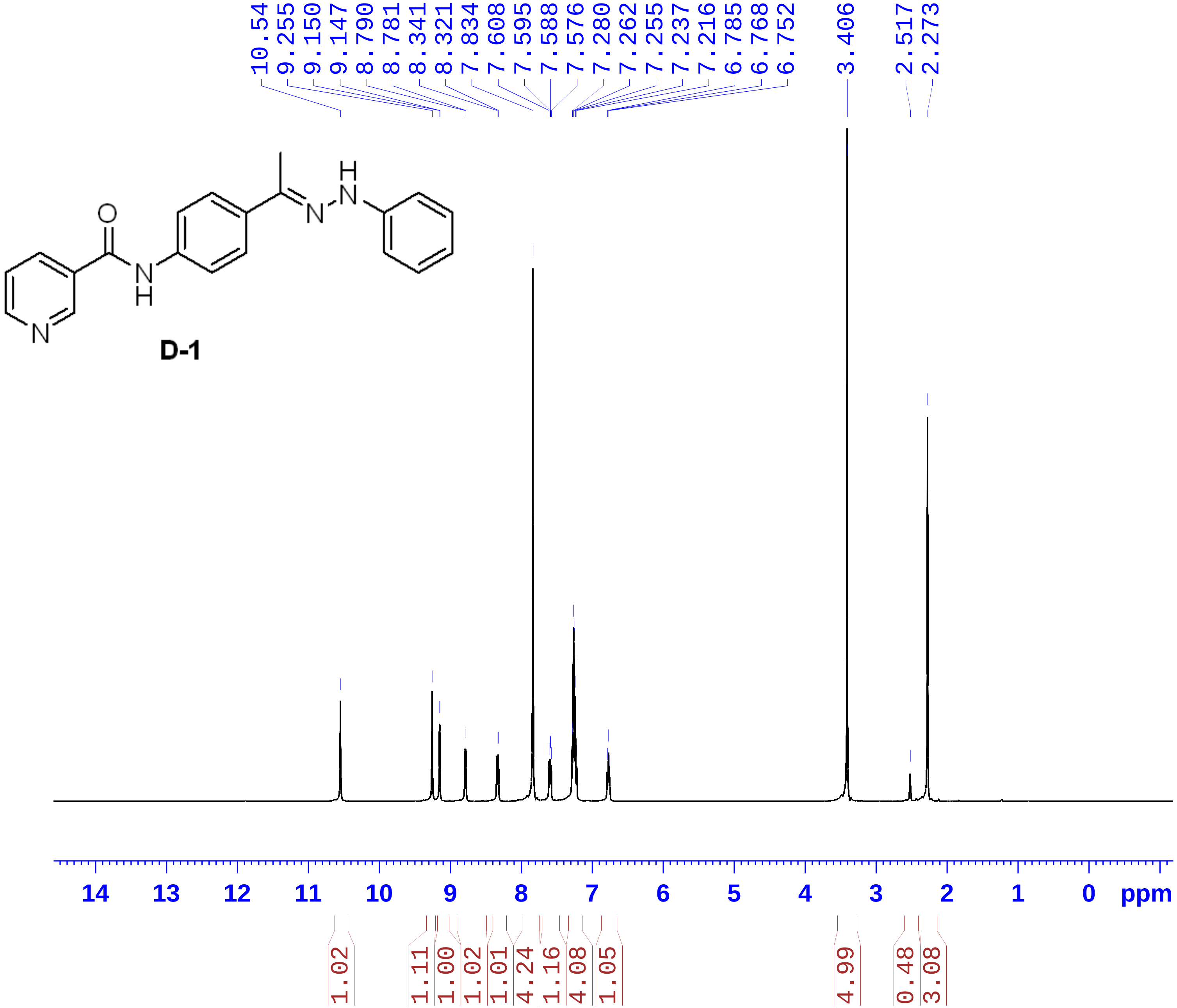
```
Current Data Parameters
NAME     Wagdy eldehna -RB8-carbon-Es
EXPNO                     10
PROCNO                     1

F2 - Acquisition Parameters
Date_               20210627
Time                   22.07 h
INSTRUM                spect
PROBHD   Z108618_0945 (
PULPROG              zgpg30
TD                    65536
SOLVENT                DMSO
NS                     2200
DS                        4
SWH               24038.461 Hz
FIDRES             0.733596 Hz
AQ                1.3631488 sec
RG                   197.77
DW                   20.800 usec
DE                     6.50 usec
TE                    297.3 K
D1               2.00000000 sec
D11              0.03000000 sec
TD0                       1
SFO1            100.6404331 MHz
NUC1                    13C
P1                    10.00 usec
PLW1            47.00000000 W
SFO2            400.2016008 MHz
NUC2                     1H
CPDPRG[2            waltz16
PCPD2                 90.00 usec
PLW2            13.00000000 W
PLW12            0.29249999 W
PLW13            0.14713000 W

F2 - Processing parameters
SI                    32768
SF              100.6303700 MHz
WDW                      EM
SSB                       0
LB                     1.00 Hz
GB                        0
PC                     1.40
```

Wagdy eldehna -RB8-carbon-Es




Wagdy eldehna -RB8-carbon-Es
197.13
165.03
164.46
152.84
152.57
149.26
149.17
146.62
143.71
140.76
138.77
136.11
135.96
135.35
132.76
131.39
131.06
130.78
129.84
129.37
128.80
128.34
127.82
126.79
125.98
124.02
123.98
122.62
120.50
120.00
119.23
113.27
96.23
190
180
170
160
150
140
130
120
110
100
90
80
ppm

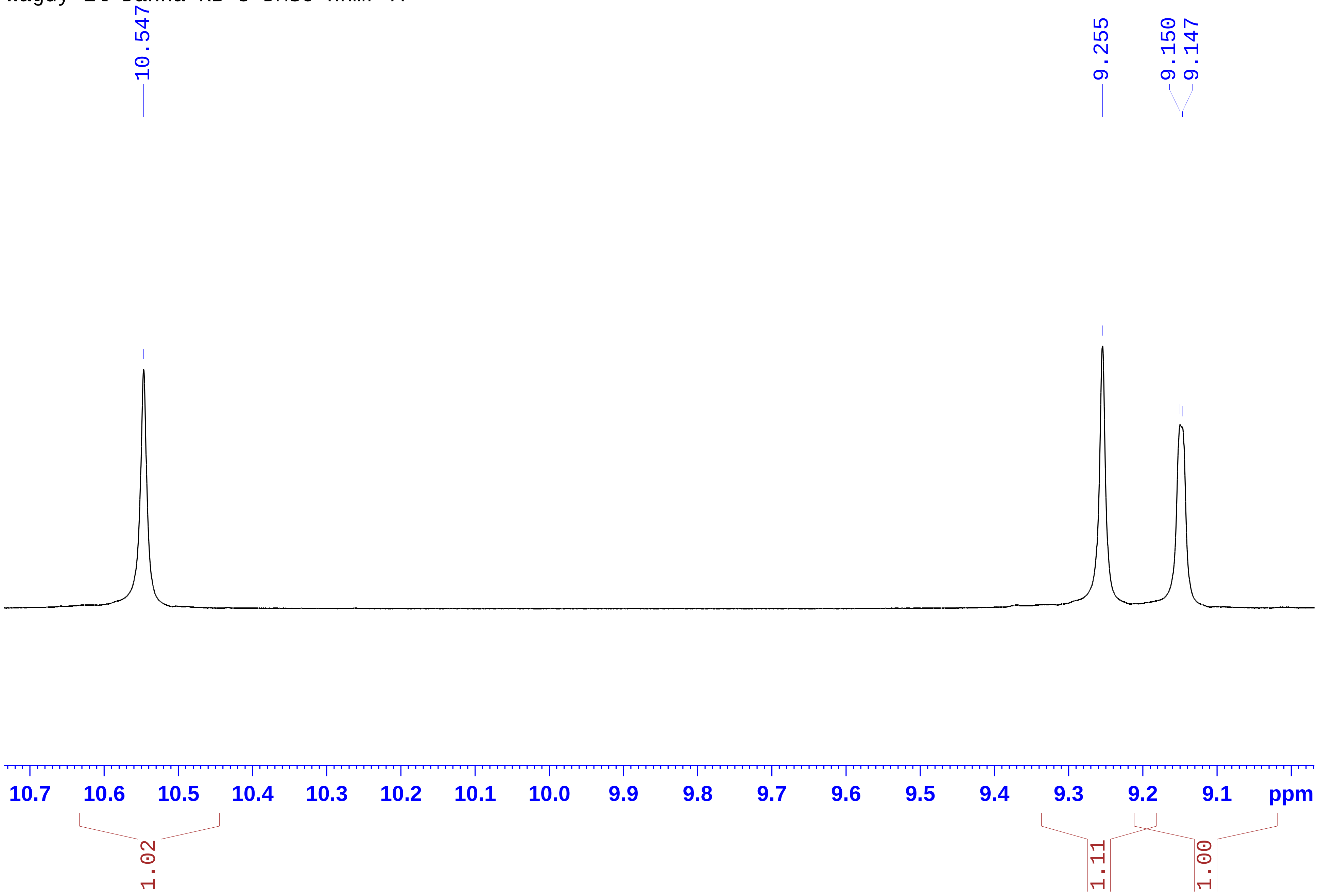
Wagdy El Dahna-RB 8-DMSO-Hnmr-A



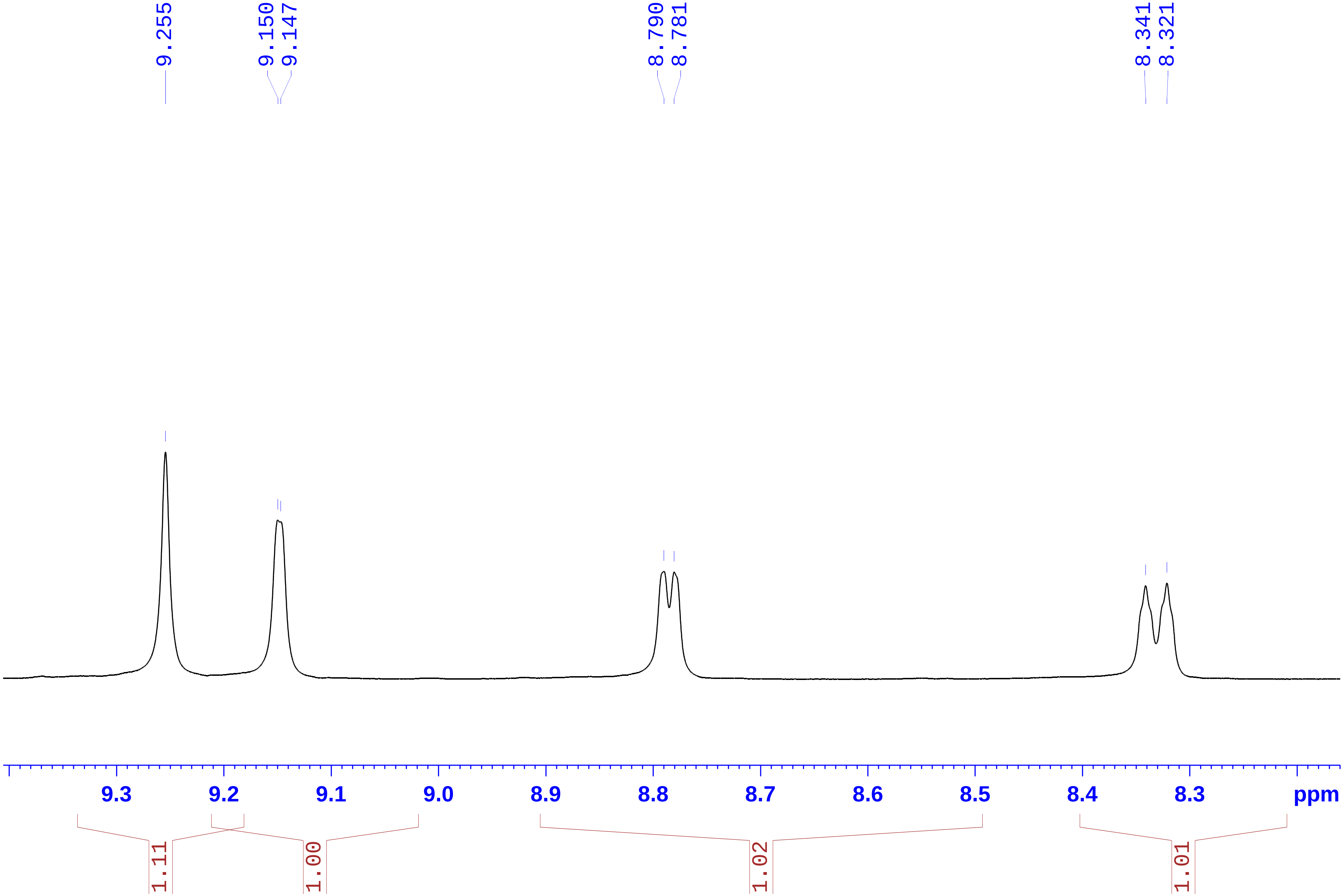
D-1

```
Current Data Parameters
NAME      Wagdy El Dahna-RB 8-DMSO-H
EXPNO                     10
PROCNO                     1

F2 - Acquisition Parameters
Date_               20210610
Time                   11.24 h
INSTRUM                spect
PROBHD   Z108618_0945 (
PULPROG                 zg30
TD                     65536
SOLVENT                 DMSO
NS                        16
DS                         2
SWH                 8012.820 Hz
FIDRES              0.244532 Hz
AQ                 4.0894465 sec
RG                      99.3
DW                    62.400 usec
DE                      6.50 usec
TE                     294.4 K
D1                1.00000000 sec
TD0                        1
SFO1             400.2024712 MHz
NUC1                      1H
P1                     13.50 usec
PLW1             13.00000000 W

F2 - Processing parameters
SI                     65536
SF               400.2000000 MHz
WDW                       EM
SSB                        0
LB                      0.30 Hz
GB                         0
PC                      1.00
```

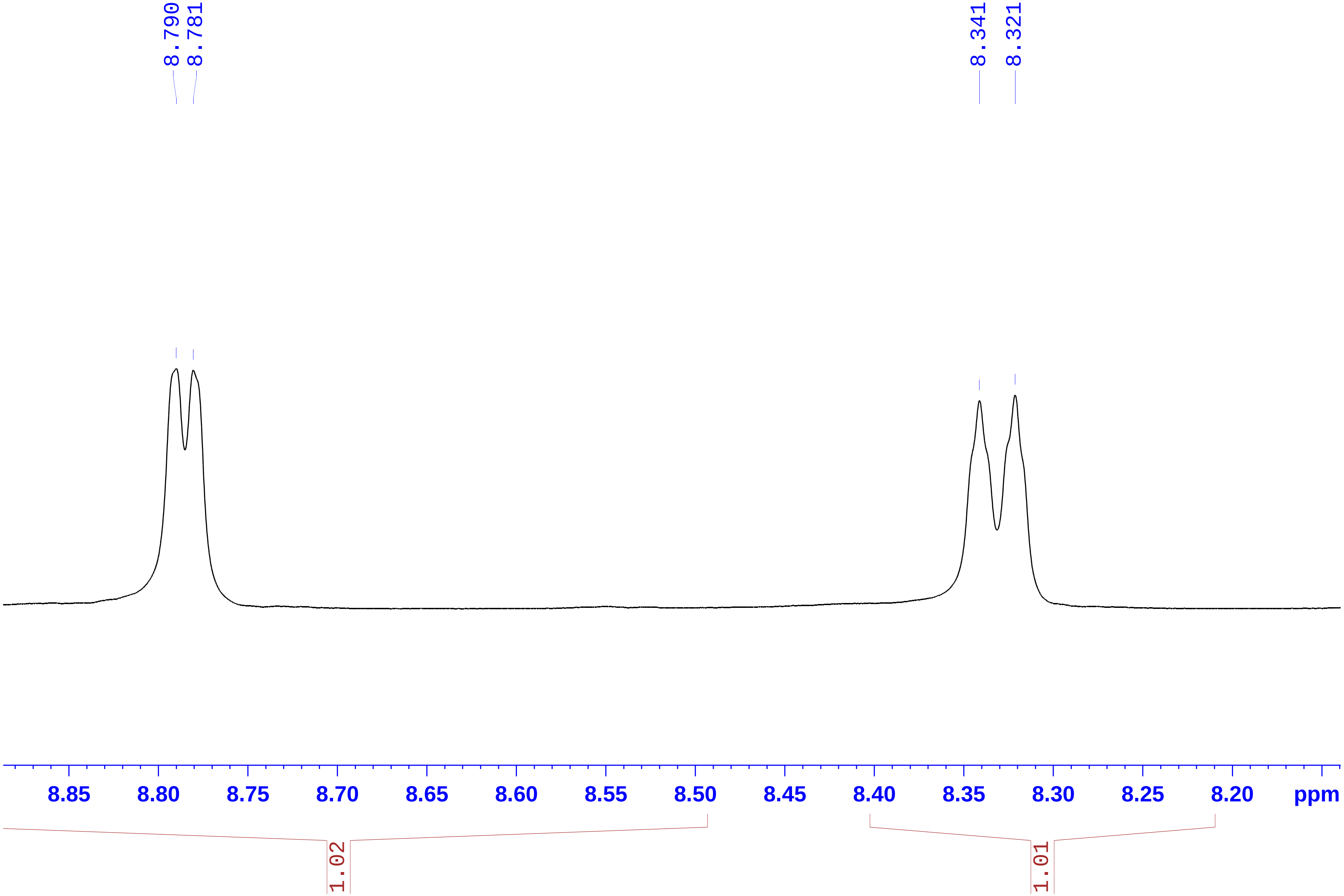
Wagdy El Dahna-RB 8-DMSO-Hnmr-A



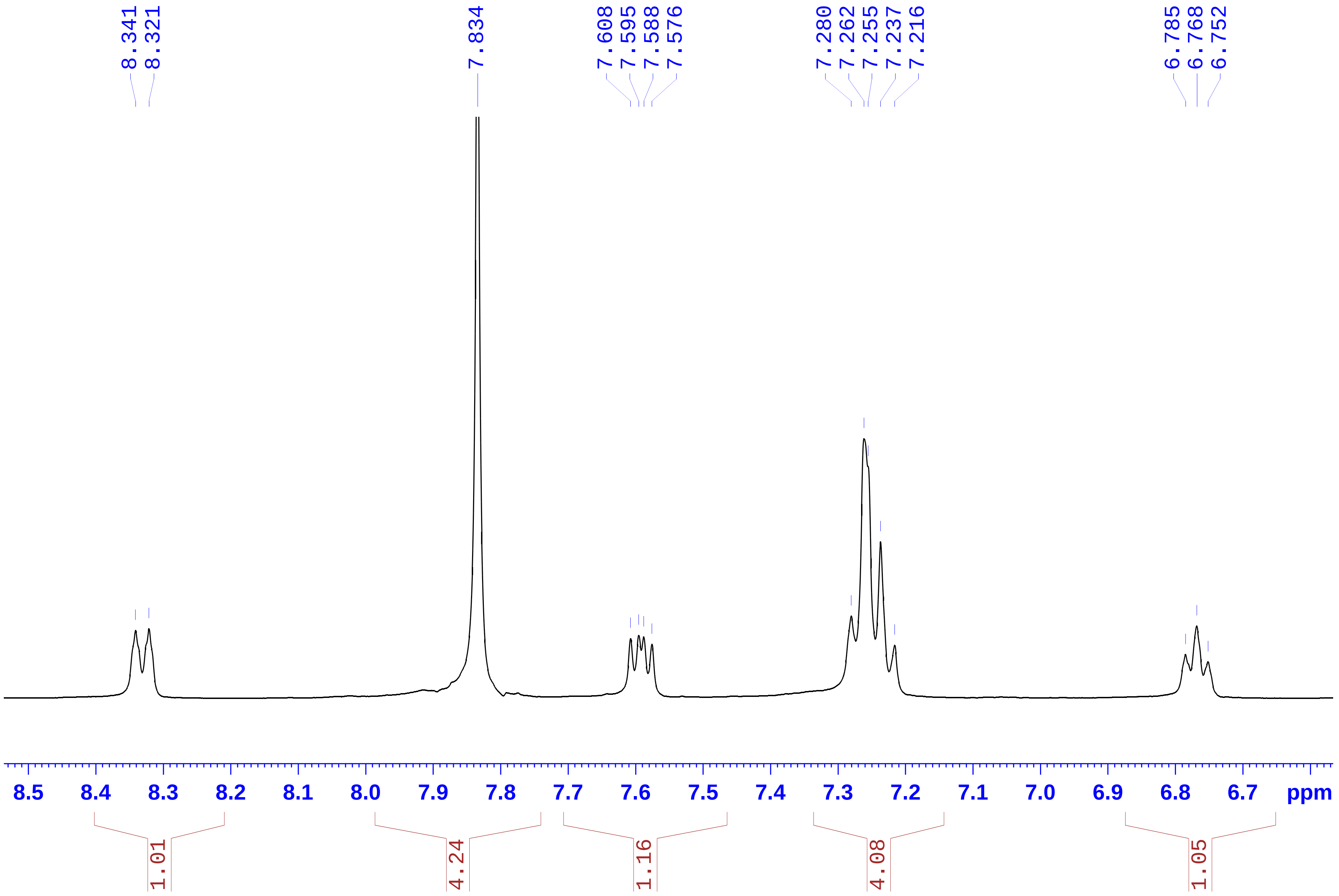
Wagdy El Dahna-RB 8-DMSO-Hnmr-A



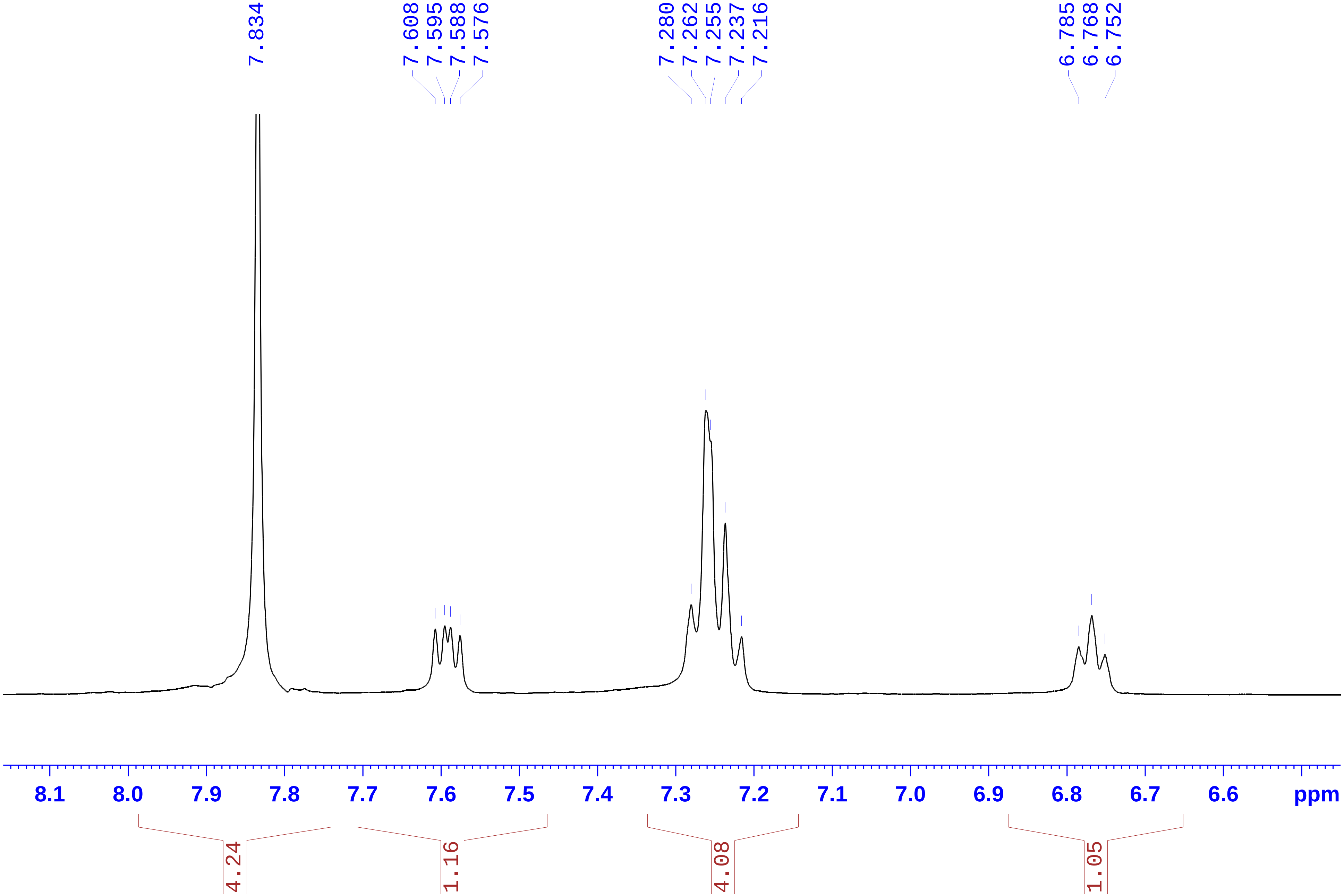
Wagdy El Dahna-RB 8-DMSO-Hnmr-A

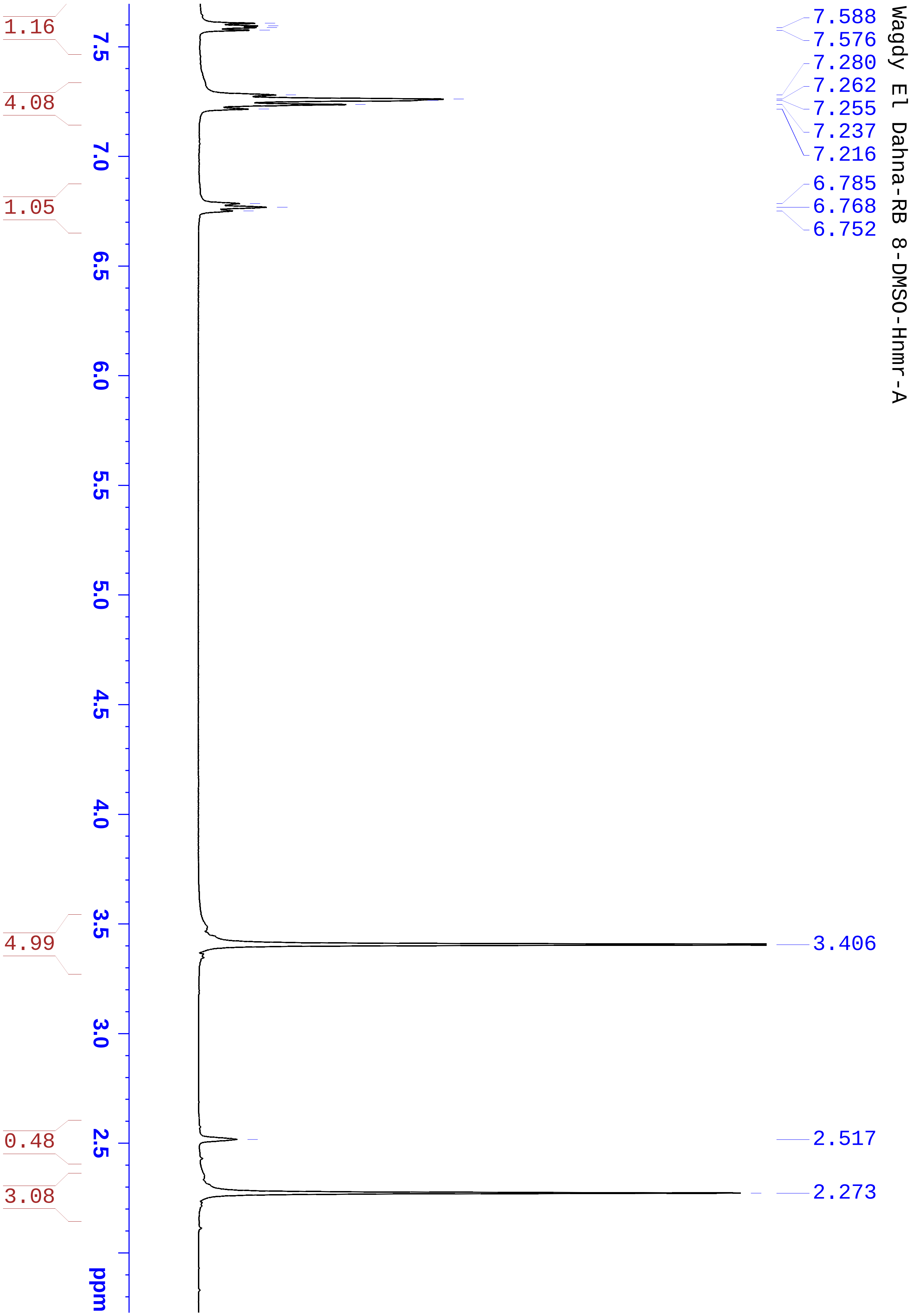


Wagdy El Dahna-RB 8-DMSO-Hnmr-A

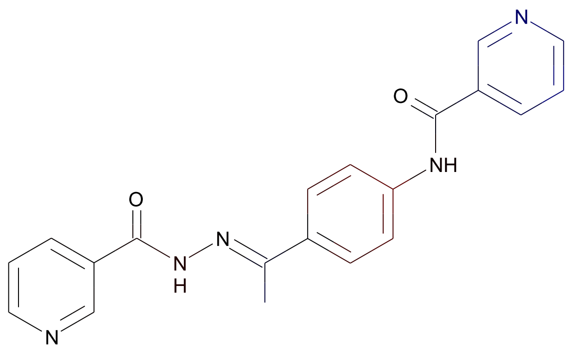


Wagdy El Dahna-RB 8-DMSO-Hnmr-A




Wagdy El Dahna-RB 8-DMSO-Hnmr-A
7.588
7.576
7.280
7.262
7.255
7.237
7.216
6.785
6.768
6.752
3.406
2.517
2.273
1.16
4.08
1.05
4.99
0.48
3.08
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
ppm

A-1 TOPKAT_Ames_Mutagenicity



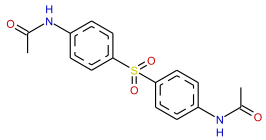
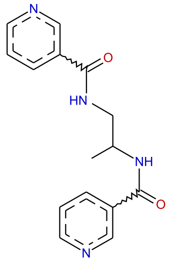
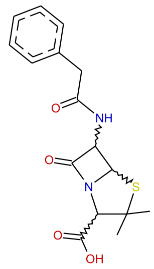
$C_{20}H_{17}N_5O_2$
Molecular Weight: 359.38128
ALogP: 1.147
Rotatable Bonds: 5
Acceptors: 5
Donors: 2

Model Prediction

Prediction: Non-Mutagen
Probability: 0.716
Enrichment: 1.28
Bayesian Score: -1.11
Mahalanobis Distance: 7.98
Mahalanobis Distance p-value: 0.993

Prediction: Positive if the Bayesian score is above the estimated best cutoff value from minimizing the false positive and false negative rate.
Probability: The esimated probability that the sample is in the positive category. This assumes that the Bayesian score follows a normal distribution and is different from the prediction using a cutoff.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Mahalanobis Distance: The Mahalanobis distance (MD) is the distance to the center of the training data. The larger the MD, the less trustworthy the prediction.
Mahalanobis Distance p-value: The p-value gives the fraction of training data with an MD greater than or equal to the one for the given sample, assuming normally distributed data. The smaller the p-value, the less trustworthy the prediciton. For highly non-normal X properties (e.g., fingerprints), the MD p-value is wildly inaccurate.

Structural Similar Compounds

Name	**77-46-3**	**79455-30-4**	**PENICILLIN G POTASSIUM**
Structure			
Actual Endpoint	Non-Mutagen	Non-Mutagen	Non-Mutagen
Predicted Endpoint	Non-Mutagen	Non-Mutagen	Non-Mutagen
Distance	0.549	0.570	0.581
Reference	Kazius et. al., J. Med. Chem. (2005) 48, 312-320	Kazius et. al., J. Med. Chem. (2005) 48, 312-320	EMIC

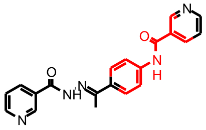
Model Applicability

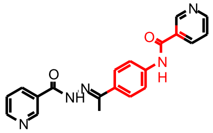
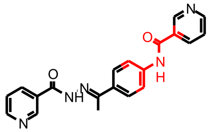
Unknown features are fingerprint features in the query molecule, but not found or appearing too infrequently in the training set.

1. All properties and OPS components are within expected ranges.

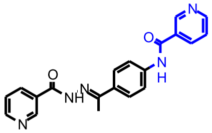
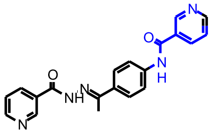
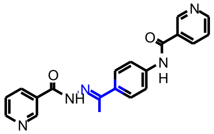
Feature Contribution

Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Mutagen in training set**
SCFP_12	555539852	 [*]:[cH]:[c](:[cH]:[*])C(=O)N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	0.447	22 out of 24

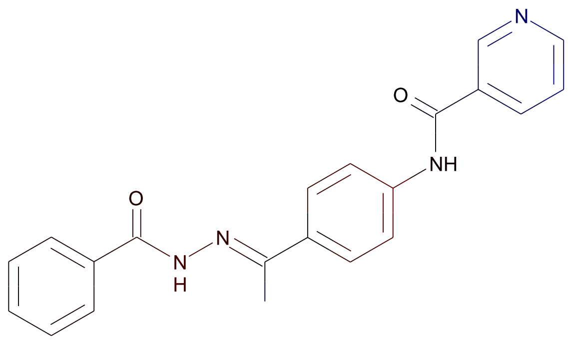
SCFP_12	818445224	 [*][c]1:[cH]:[cH]:[c](NC(=O)[c](:[*]):[*]):[cH]:[cH]:1	0.434	12 out of 13
SCFP_12	2096901122	 [*]:[cH]:[c](NC(=O)[c](:[*]):[*]):[cH]:[*]	0.429	33 out of 37

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Mutagen in training set
SCFP_12	1165971455	 [*]NC(=O)[c]1:[cH]:[cH]:[cH]:n:[cH]:1	-0.762	0 out of 2
SCFP_12	903335088	 [*]NC(=O)[c]1:[cH]:[cH]:[*]:n:[cH]:1	-0.762	0 out of 2
SCFP_12	-331724199	 [*]\N=C(/C)\[c](:[*]):[*]	-0.762	0 out of 2

B-1

TOPKAT_Ames_Mutagenicity



$C_{21}H_{18}N_4O_2$

Molecular Weight: 358.39322

ALogP: 2.297

Rotatable Bonds: 5

Acceptors: 4

Donors: 2

Model Prediction

Prediction: Non-Mutagen

Probability: 0.724

Enrichment: 1.3

Bayesian Score: -0.747

Mahalanobis Distance: 7.86

Mahalanobis Distance p-value: 0.996

Prediction: Positive if the Bayesian score is above the estimated best cutoff value from minimizing the false positive and false negative rate.

Probability: The esimated probability that the sample is in the positive category. This assumes that the Bayesian score follows a normal distribution and is different from the prediction using a cutoff.

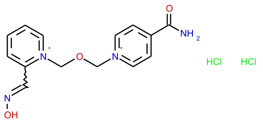
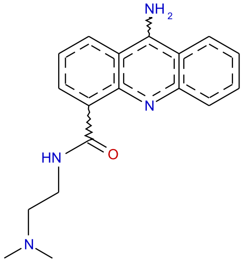
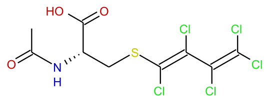
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.

Bayesian Score: The standard Laplacian-modified Bayesian score.

Mahalanobis Distance: The Mahalanobis distance (MD) is the distance to the center of the training data. The larger the MD, the less trustworthy the prediction.

Mahalanobis Distance p-value: The p-value gives the fraction of training data with an MD greater than or equal to the one for the given sample, assuming normally distributed data. The smaller the p-value, the less trustworthy the prediciton. For highly non-normal X properties (e.g., fingerprints), the MD p-value is wildly inaccurate.

Structural Similar Compounds

Name	**34433-31-3**	**89459-43-8**	**89784-39-4**
Structure			
Actual Endpoint	Non-Mutagen	Mutagen	Mutagen
Predicted Endpoint	Non-Mutagen	Mutagen	Mutagen
Distance	0.499	0.548	0.562
Reference	Kazius et. al., J. Med. Chem. (2005) 48, 312-320	Kazius et. al., J. Med. Chem. (2005) 48, 312-320	Kazius et. al., J. Med. Chem. (2005) 48, 312-320

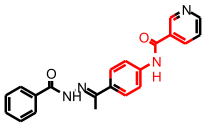
Model Applicability

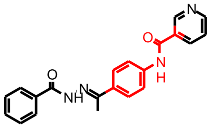
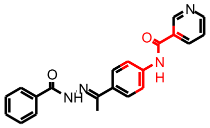
Unknown features are fingerprint features in the query molecule, but not found or appearing too infrequently in the training set.

1. All properties and OPS components are within expected ranges.

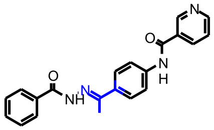
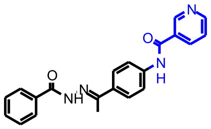
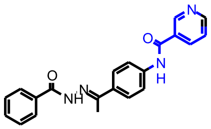
Feature Contribution

Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Mutagen in training set**
SCFP_12	555539852	 [*]:[cH]:[c](:[cH]:[*])C(=O)N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	0.447	22 out of 24

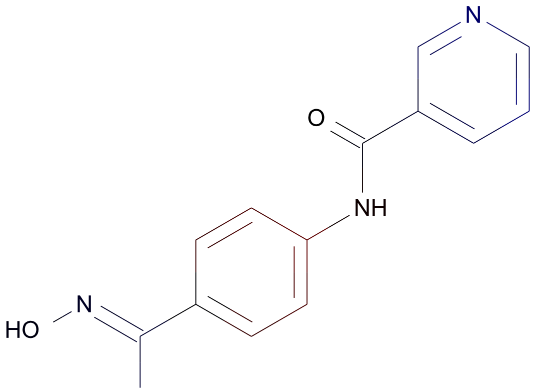
SCFP_12	818445224	 [*][c]1:[cH]:[cH]:[c](NC(=O)[c](:[*]):[*]):[cH]:[cH]:1	0.434	12 out of 13
SCFP_12	2096901122	 [*]:[cH]:[c](NC(=O)[c](:[*]):[*]):[cH]:[*]	0.429	33 out of 37

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Mutagen in training set**
SCFP_12	-331724199	 [*]\N=C(/C)\[c](:[*]):[*]	-0.762	0 out of 2
SCFP_12	1165971455	 [*]NC(=O)[c]1:[cH]:[cH]:[cH]:n:[cH]:1	-0.762	0 out of 2
SCFP_12	903335088	 [*]NC(=O)[c]1:[cH]:[cH]:[*]:n:[cH]:1	-0.762	0 out of 2

C-6

TOPKAT_Ames_Mutagenicity



$C_{14}H_{13}N_3O_2$

Molecular Weight: 255.27191

ALogP: 1.271

Rotatable Bonds: 3

Acceptors: 4

Donors: 2

Model Prediction

Prediction: Non-Mutagen

Probability: 0.718

Enrichment: 1.29

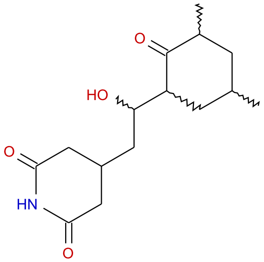
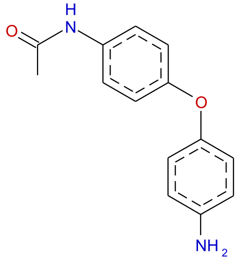
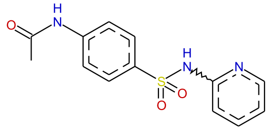
Bayesian Score: -1.02

Mahalanobis Distance: 7.73

Mahalanobis Distance p-value: 0.998

Prediction: Positive if the Bayesian score is above the estimated best cutoff value from minimizing the false positive and false negative rate.
Probability: The esimated probability that the sample is in the positive category. This assumes that the Bayesian score follows a normal distribution and is different from the prediction using a cutoff.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Mahalanobis Distance: The Mahalanobis distance (MD) is the distance to the center of the training data. The larger the MD, the less trustworthy the prediction.
Mahalanobis Distance p-value: The p-value gives the fraction of training data with an MD greater than or equal to the one for the given sample, assuming normally distributed data. The smaller the p-value, the less trustworthy the prediciton. For highly non-normal X properties (e.g., fingerprints), the MD p-value is wildly inaccurate.

Structural Similar Compounds

Name	**Cycloheximide**	**2687-41-4**	**19077-98-6**
Structure			
Actual Endpoint	Non-Mutagen	Mutagen	Non-Mutagen
Predicted Endpoint	Non-Mutagen	Mutagen	Non-Mutagen
Distance	0.518	0.518	0.526
Reference	Environ. Mol. Mut. 19(21):1992	Kazius et. al., J. Med. Chem. (2005) 48, 312-320	Kazius et. al., J. Med. Chem. (2005) 48, 312-320

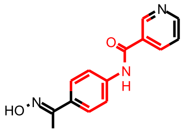
Model Applicability

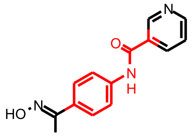
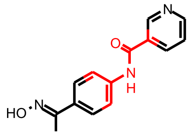
Unknown features are fingerprint features in the query molecule, but not found or appearing too infrequently in the training set.

1. All properties and OPS components are within expected ranges.

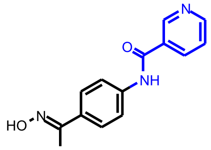
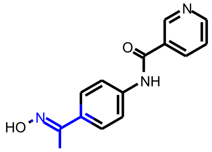
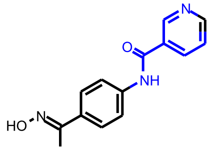
Feature Contribution

Top features for positive contribution

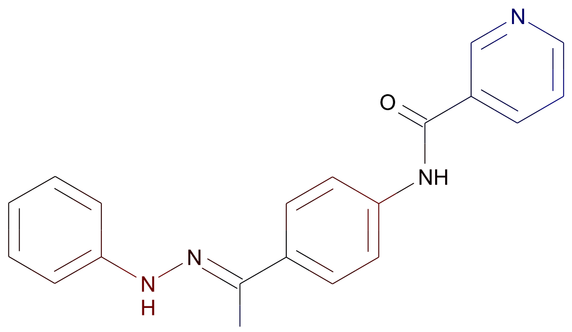
Fingerprint	Bit/Smiles	Feature Structure	Score	Mutagen in training set
SCFP_12	555539852	 [*]:[cH]:[c](:[cH]:[*])C(=O)N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	0.447	22 out of 24

SCFP_12	818445224	 [*][c]1:[cH]:[cH]:[c](NC(=O)[c](:[*]):[*]):[cH]:[cH]:1	0.434	12 out of 13
SCFP_12	2096901122	 [*]:[cH]:[c](NC(=O)[c](:[*]):[*]):[cH]:[*]	0.429	33 out of 37

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Mutagen in training set**
SCFP_12	1165971455	 [*]NC(=O)[c]1:[cH]:[cH]:[cH]:n:[cH]:1	-0.762	0 out of 2
SCFP_12	-331724199	 [*]\N=C(/C)\[c](:[*]):[*]	-0.762	0 out of 2
SCFP_12	903335088	 [*]NC(=O)[c]1:[cH]:[cH]:[*]:n:[cH]:1	-0.762	0 out of 2

D-1 TOPKAT_Ames_Mutagenicity



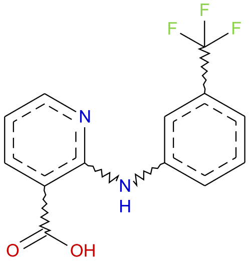
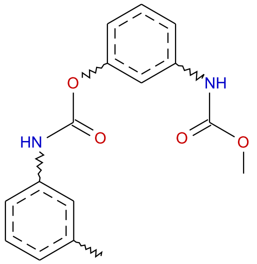
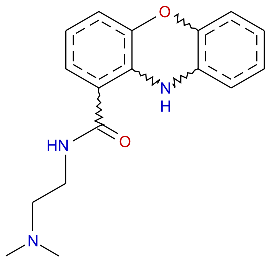
$C_{20}H_{18}N_4O$
Molecular Weight: 330.38312
ALogP: 3.063
Rotatable Bonds: 5
Acceptors: 4
Donors: 2

Model Prediction

Prediction: Mutagen
Probability: 0.75
Enrichment: 1.34
Bayesian Score: 0.586
Mahalanobis Distance: 8.09
Mahalanobis Distance p-value: 0.989

Prediction: Positive if the Bayesian score is above the estimated best cutoff value from minimizing the false positive and false negative rate.
Probability: The esimated probability that the sample is in the positive category. This assumes that the Bayesian score follows a normal distribution and is different from the prediction using a cutoff.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Mahalanobis Distance: The Mahalanobis distance (MD) is the distance to the center of the training data. The larger the MD, the less trustworthy the prediction.
Mahalanobis Distance p-value: The p-value gives the fraction of training data with an MD greater than or equal to the one for the given sample, assuming normally distributed data. The smaller the p-value, the less trustworthy the prediciton. For highly non-normal X properties (e.g., fingerprints), the MD p-value is wildly inaccurate.

Structural Similar Compounds

Name	4394-00-7	Phenmedipham	112022-13-6
Structure			
Actual Endpoint	Non-Mutagen	Mutagen	Mutagen
Predicted Endpoint	Non-Mutagen	Mutagen	Mutagen
Distance	0.528	0.535	0.538
Reference	Kazius et. al., J. Med. Chem. (2005) 48, 312-320	Environ. Mol. Mut. 19(21):1992	Kazius et. al., J. Med. Chem. (2005) 48, 312-320

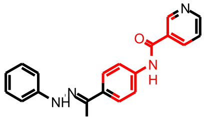
Model Applicability

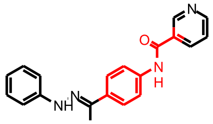
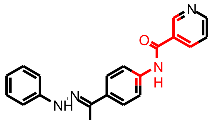
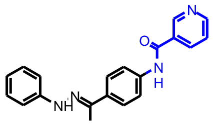
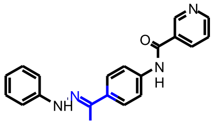
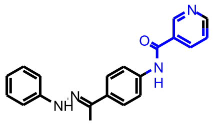
Unknown features are fingerprint features in the query molecule, but not found or appearing too infrequently in the training set.

1. All properties and OPS components are within expected ranges.

Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Mutagen in training set
SCFP_12	555539852	 [*]:[cH]:[c](:[cH]:[*])C(=O)N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	0.447	22 out of 24

SCFP_12	818445224	 [*][c]1:[cH]:[cH]:[c](NC(=O)[c](:[*]):[*]):[cH]:[cH]:1	0.434	12 out of 13
SCFP_12	124026986	 [*]:[cH]:[c](:[cH]:[*])C(=O)N[c](:[*]):[*]	0.429	33 out of 37
Top Features for negative contribution				
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Mutagen in training set**
SCFP_12	1165971455	 [*]NC(=O)[c]1:[cH]:[cH]:[cH]:n:[cH]:1	-0.762	0 out of 2
SCFP_12	-331724199	 [*]\N=C(/C)\[c](:[*]):[*]	-0.762	0 out of 2
SCFP_12	903335088	 [*]NC(=O)[c]1:[cH]:[cH]:[*]:n:[cH]:1	-0.762	0 out of 2

Sorafenib

TOPKAT_Ames_Mutagenicity

$C_{21}H_{16}ClF_3N_4O_3$

Molecular Weight: 464.82494

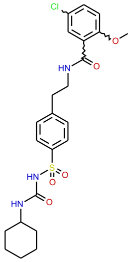
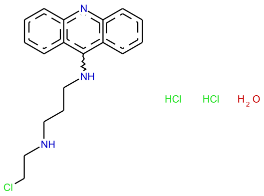
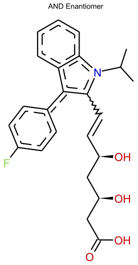
ALogP: 4.175

Rotatable Bonds: 6

Acceptors: 4

Donors: 3

Structural Similar Compounds

Name	GLYBURIDE	38914-96-4	93957-54-1
Structure			
Actual Endpoint	Non-Mutagen	Mutagen	Non-Mutagen
Predicted Endpoint	Non-Mutagen	Mutagen	Non-Mutagen
Distance	0.590	0.592	0.600
Reference	PDR 1994	Kazius et. al., J. Med. Chem. (2005) 48, 312-320	US Environmental Protection Agency at http://www.epa.gov/NCCT/dsstox/sdf_isscan_external.html

Model Prediction

Prediction: Non-Mutagen

Probability: 0.0531

Enrichment: 0.0951

Bayesian Score: -19.7

Mahalanobis Distance: 13.1

Mahalanobis Distance p-value: 2.73e-006

Prediction: Positive if the Bayesian score is above the estimated best cutoff value from minimizing the false positive and false negative rate.

Probability: The esimated probability that the sample is in the positive category. This assumes that the Bayesian score follows a normal distribution and is different from the prediction using a cutoff.

Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.

Bayesian Score: The standard Laplacian-modified Bayesian score.

Mahalanobis Distance: The Mahalanobis distance (MD) is the distance to the center of the training data. The larger the MD, the less trustworthy the prediction.

Mahalanobis Distance p-value: The p-value gives the fraction of training data with an MD greater than or equal to the one for the given sample, assuming normally distributed data. The smaller the p-value, the less trustworthy the prediciton. For highly non-normal X properties (e.g., fingerprints), the MD p-value is wildly inaccurate.

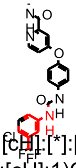
Model Applicability

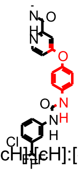
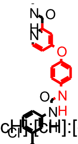
Unknown features are fingerprint features in the query molecule, but not found or appearing too infrequently in the training set.

1. All properties and OPS components are within expected ranges.

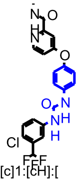
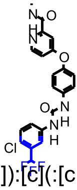
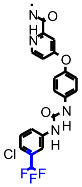
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Mutagen in training set
SCFP_12	-347281112	 [*]N[c]1:[cH]:[*]:[c]([*]):[c](:[cH]:1)C([*])([*])[*]	0.337	18 out of 22

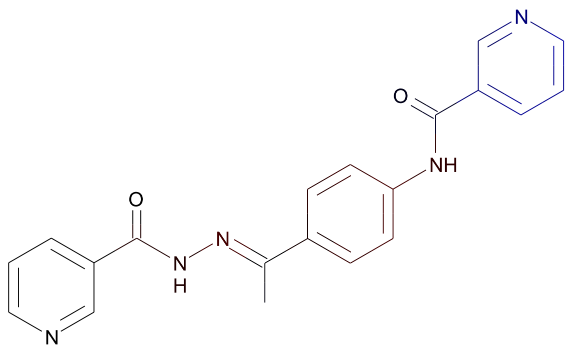
SCFP_12	1208843554	 [*]N[c]1:[cH]:[cH]:[c](O[c](:[*]):[*]):[cH]:[cH]:1	0.337	6 out of 7
SCFP_12	-1943080297	 [*]N[c]1:[cH]:[cH]:[c](O[c]2:[cH]:[cH]:[*]:[c]([*]):[cH]:2):[cH]:[cH]:1	0.304	5 out of 6

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Mutagen in training set
SCFP_12	816802409	 [*]NC(=O)N[c]1:[cH]:[cH]:[c]([*]):[cH]:[cH]:1	-1.82	0 out of 9
SCFP_12	-1903175541	 [*][c](:[*]):[c](:[cH]:[*])C(F)(F)F	-1.51	3 out of 30
SCFP_12	-300280774	 [*]:[c](:[*])C(F)(F)F	-1.51	3 out of 30

A-1

TOPKAT_Mouse_Female_FDA_None_vs_Carcinogen



$C_{20}H_{17}N_5O_2$

Molecular Weight: 359.38128

ALogP: 1.147

Rotatable Bonds: 5

Acceptors: 5

Donors: 2

Model Prediction

Prediction: Non-Carcinogen

Probability: 0.242

Enrichment: 0.755

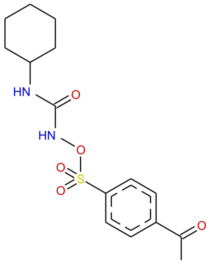
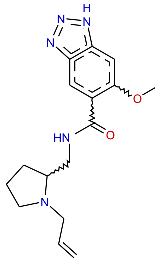
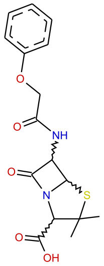
Bayesian Score: -1.2

Mahalanobis Distance: 11

Mahalanobis Distance p-value: 0.0949

Prediction: Positive if the Bayesian score is above the estimated best cutoff value from minimizing the false positive and false negative rate.
Probability: The esimated probability that the sample is in the positive category. This assumes that the Bayesian score follows a normal distribution and is different from the prediction using a cutoff.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Mahalanobis Distance: The Mahalanobis distance (MD) is the distance to the center of the training data. The larger the MD, the less trustworthy the prediction.
Mahalanobis Distance p-value: The p-value gives the fraction of training data with an MD greater than or equal to the one for the given sample, assuming normally distributed data. The smaller the p-value, the less trustworthy the prediciton. For highly non-normal X properties (e.g., fingerprints), the MD p-value is wildly inaccurate.

Structural Similar Compounds

Name	Acetohexamide	Alizapride	Penicillin
Structure			
Actual Endpoint	Non-Carcinogen	Carcinogen	Non-Carcinogen
Predicted Endpoint	Non-Carcinogen	Carcinogen	Non-Carcinogen
Distance	0.578	0.581	0.603
Reference	US FDA (Centre for Drug Eval.& Res./Off. Testing & Res.) Sept. 1997	US FDA (Centre for Drug Eval.& Res./Off. Testing & Res.) Sept. 1997	US FDA (Centre for Drug Eval.& Res./Off. Testing & Res.) Sept. 1997

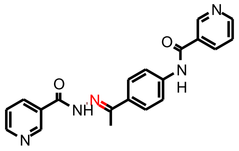
Model Applicability

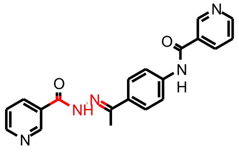
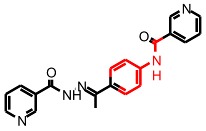
Unknown features are fingerprint features in the query molecule, but not found or appearing too infrequently in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_2 feature: 128986386: [*]\N=C(/C)\[c](:[*]):[*]
3. Unknown ECFP_2 feature: 560380707: [*]NN=C([*])[*]

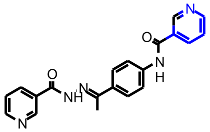
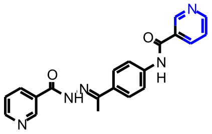
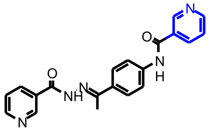
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Carcinogen in training set
ECFP_6	-1087070950	 [*]N=[*]	0.724	10 out of 14

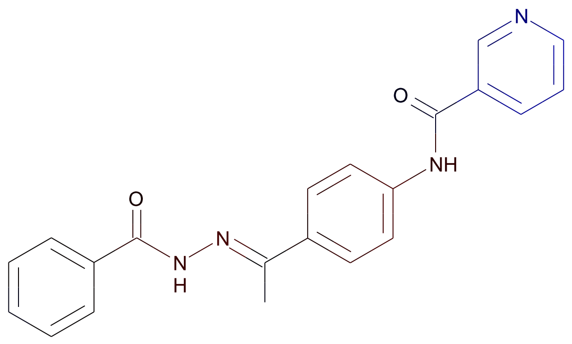
ECFP_6	544048674	 [*]C(=[*])NN=[*]	0.617	2 out of 2
ECFP_6	738938915	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	0.617	2 out of 2

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Carcinogen in training set
ECFP_6	2013347047	 [*][c]1:[*]:n:[cH]:[cH]:[cH]:1	-0.805	0 out of 4
ECFP_6	764951226	 [*]1:[cH]:[cH]:[cH]:n:[cH]:1	-0.482	0 out of 2
ECFP_6	-1818873508	 [*][c]1:[cH]:[cH]:[cH]:n:[cH]:1	-0.482	0 out of 2

B-1

TOPKAT_Mouse_Female_FDA_None_vs_Carcinogen



$C_{21}H_{18}N_4O_2$

Molecular Weight: 358.39322

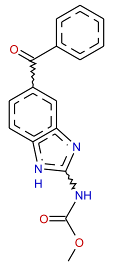
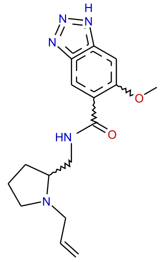
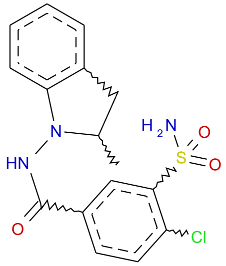
ALogP: 2.297

Rotatable Bonds: 5

Acceptors: 4

Donors: 2

Structural Similar Compounds

Name	Mebendazole	Alizapride	Indapamide
Structure			
Actual Endpoint	Non-Carcinogen	Carcinogen	Non-Carcinogen
Predicted Endpoint	Non-Carcinogen	Carcinogen	Non-Carcinogen
Distance	0.570	0.601	0.601
Reference	US FDA (Centre for Drug Eval.& Res./Off. Testing & Res.) Sept. 1997	US FDA (Centre for Drug Eval.& Res./Off. Testing & Res.) Sept. 1997	US FDA (Centre for Drug Eval.& Res./Off. Testing & Res.) Sept. 1997

Model Prediction

Prediction: Non-Carcinogen

Probability: 0.231

Enrichment: 0.721

Bayesian Score: -1.96

Mahalanobis Distance: 11.7

Mahalanobis Distance p-value: 0.0236

Prediction: Positive if the Bayesian score is above the estimated best cutoff value from minimizing the false positive and false negative rate.
Probability: The esimated probability that the sample is in the positive category. This assumes that the Bayesian score follows a normal distribution and is different from the prediction using a cutoff.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Mahalanobis Distance: The Mahalanobis distance (MD) is the distance to the center of the training data. The larger the MD, the less trustworthy the prediction.
Mahalanobis Distance p-value: The p-value gives the fraction of training data with an MD greater than or equal to the one for the given sample, assuming normally distributed data. The smaller the p-value, the less trustworthy the prediciton. For highly non-normal X properties (e.g., fingerprints), the MD p-value is wildly inaccurate.

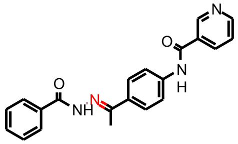
Model Applicability

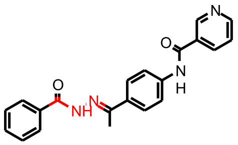
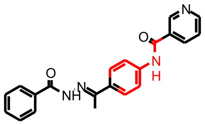
Unknown features are fingerprint features in the query molecule, but not found or appearing too infrequently in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_2 feature: 128986386: [*]\N=C(/C)\[c](:[*]):[*]
3. Unknown ECFP_2 feature: 560380707: [*]NN=C([*])[*]

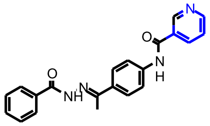
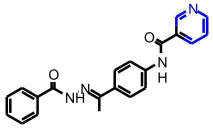
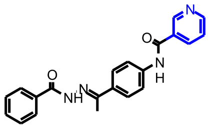
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Carcinogen in training set
ECFP_6	-1087070950	 [*]N=[*]	0.724	10 out of 14

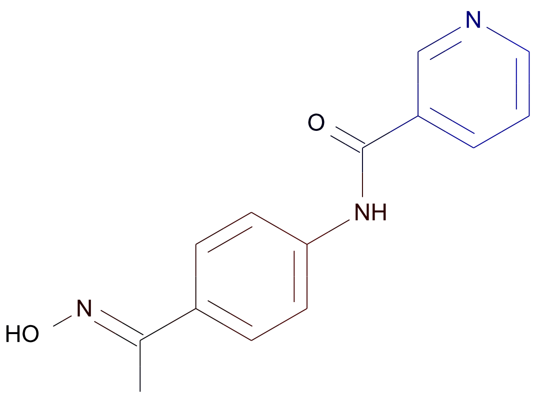
ECFP_6	544048674	 [*]C(=[*])NN=[*]	0.617	2 out of 2
ECFP_6	738938915	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	0.617	2 out of 2

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Carcinogen in training set
ECFP_6	2013347047	 [*][c]1:[*]:n:[cH]:[cH]:[cH]:1	-0.805	0 out of 4
ECFP_6	764951226	 [*]1:[cH]:[cH]:[cH]:n:[cH]:1	-0.482	0 out of 2
ECFP_6	-1818873508	 [*][c]1:[cH]:[cH]:[cH]:n:[cH]:1	-0.482	0 out of 2

C-6

TOPKAT_Mouse_Female_FDA_None_vs_Carcinogen



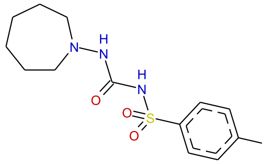
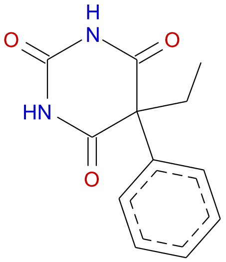
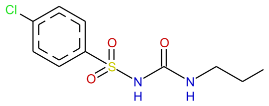
$C_{14}H_{13}N_3O_2$
Molecular Weight: 255.27191
ALogP: 1.271
Rotatable Bonds: 3
Acceptors: 4
Donors: 2

Model Prediction

Prediction: Non-Carcinogen
Probability: 0.233
Enrichment: 0.726
Bayesian Score: -1.83
Mahalanobis Distance: 10.9
Mahalanobis Distance p-value: 0.121

Prediction: Positive if the Bayesian score is above the estimated best cutoff value from minimizing the false positive and false negative rate.
Probability: The esimated probability that the sample is in the positive category. This assumes that the Bayesian score follows a normal distribution and is different from the prediction using a cutoff.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Mahalanobis Distance: The Mahalanobis distance (MD) is the distance to the center of the training data. The larger the MD, the less trustworthy the prediction.
Mahalanobis Distance p-value: The p-value gives the fraction of training data with an MD greater than or equal to the one for the given sample, assuming normally distributed data. The smaller the p-value, the less trustworthy the prediciton. For highly non-normal X properties (e.g., fingerprints), the MD p-value is wildly inaccurate.

Structural Similar Compounds

Name	Tolazamide	Phenobarbital	Chlorpropamide
Structure			
Actual Endpoint	Non-Carcinogen	Carcinogen	Non-Carcinogen
Predicted Endpoint	Non-Carcinogen	Carcinogen	Non-Carcinogen
Distance	0.576	0.586	0.591
Reference	US FDA (Centre for Drug Eval.& Res./Off. Testing & Res.) Sept. 1997	US FDA (Centre for Drug Eval.& Res./Off. Testing & Res.) Sept. 1997	US FDA (Centre for Drug Eval.& Res./Off. Testing & Res.) Sept. 1997

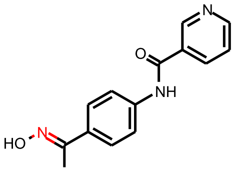
Model Applicability

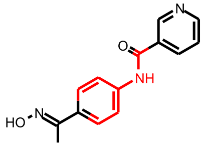
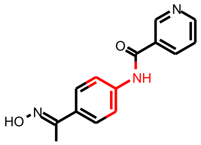
Unknown features are fingerprint features in the query molecule, but not found or appearing too infrequently in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_2 feature: 128986386: [*]\N=C(/C)\[c](:[*]):[*]
3. Unknown ECFP_2 feature: 559672449: [*]C(=NO)[*]
4. Unknown ECFP_2 feature: 2025254696: [*]=NO

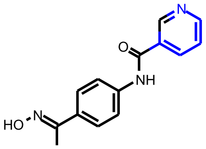
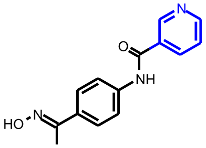
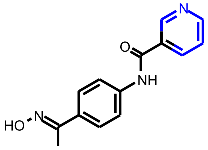
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Carcinogen in training set
ECFP_6	-1087070950	 [*]N=[*]	0.724	10 out of 14

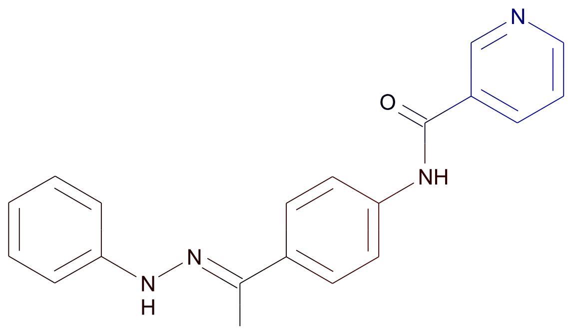
ECFP_6	738938915	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	0.617	2 out of 2
ECFP_6	-177077903	 [*]N[c](:[cH]:[*]):[cH]:[*]	0.279	4 out of 9

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Carcinogen in training set**
ECFP_6	2013347047	 [*][c]1:[*]:n:[cH]:[cH]:[cH]:1	-0.805	0 out of 4
ECFP_6	-1818873508	 [*][c]1:[cH]:[cH]:[cH]:n:[cH]:1	-0.482	0 out of 2
ECFP_6	764951226	 [*]1:[cH]:[cH]:[cH]:n:[cH]:1	-0.482	0 out of 2

D-1

TOPKAT_Mouse_Female_FDA_None_vs_Carcinogen



$C_{20}H_{18}N_4O$

Molecular Weight: 330.38312

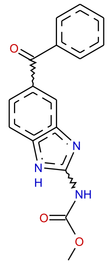
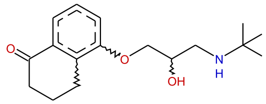
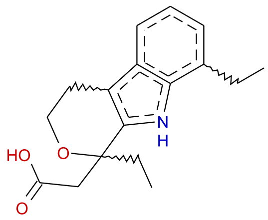
ALogP: 3.063

Rotatable Bonds: 5

Acceptors: 4

Donors: 2

Structural Similar Compounds

Name	Mebendazole	Bunolol	Etodolac
Structure			
Actual Endpoint	Non-Carcinogen	Carcinogen	Non-Carcinogen
Predicted Endpoint	Non-Carcinogen	Carcinogen	Non-Carcinogen
Distance	0.559	0.582	0.606
Reference	US FDA (Centre for Drug Eval.& Res./Off. Testing & Res.) Sept. 1997	US FDA (Centre for Drug Eval.& Res./Off. Testing & Res.) Sept. 1997	US FDA (Centre for Drug Eval.& Res./Off. Testing & Res.) Sept. 1997

Model Applicability

Unknown features are fingerprint features in the query molecule, but not found or appearing too infrequently in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_2 feature: 128986386: [*]\N=C(/C)\[c](:[*]):[*]
3. Unknown ECFP_2 feature: 560380707: [*]NN=C([*])[*]
4. Unknown ECFP_2 feature: -1236714312: [*]=NN[c](:[*]):[*]

Model Prediction

Prediction: Non-Carcinogen

Probability: 0.234

Enrichment: 0.729

Bayesian Score: -1.77

Mahalanobis Distance: 11.7

Mahalanobis Distance p-value: 0.0253

Prediction: Positive if the Bayesian score is above the estimated best cutoff value from minimizing the false positive and false negative rate.

Probability: The esimated probability that the sample is in the positive category. This assumes that the Bayesian score follows a normal distribution and is different from the prediction using a cutoff.

Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.

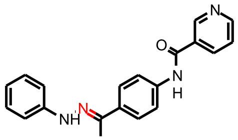
Bayesian Score: The standard Laplacian-modified Bayesian score.

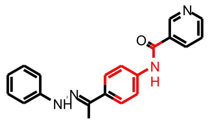
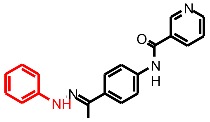
Mahalanobis Distance: The Mahalanobis distance (MD) is the distance to the center of the training data. The larger the MD, the less trustworthy the prediction.

Mahalanobis Distance p-value: The p-value gives the fraction of training data with an MD greater than or equal to the one for the given sample, assuming normally distributed data. The smaller the p-value, the less trustworthy the prediciton. For highly non-normal X properties (e.g., fingerprints), the MD p-value is wildly inaccurate.

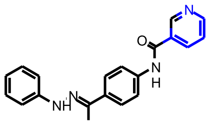
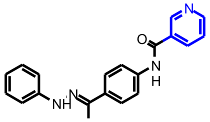
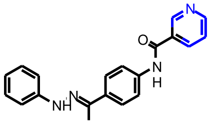
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Carcinogen in training set
ECFP_6	-1087070950	 [*]N=[*]	0.724	10 out of 14

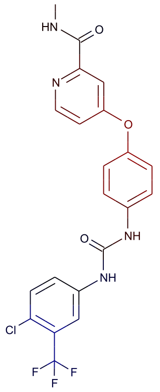
ECFP_6	738938915	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	0.617	2 out of 2
ECFP_6	547788473	 [*]N[c]1:[cH]:[cH]:[cH]:[cH]:[cH]:1	0.424	1 out of 1

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Carcinogen in training set**
ECFP_6	2013347047	 [*][c]1:[*]:n:[cH]:[cH]:[cH]:1	-0.805	0 out of 4
ECFP_6	-1818873508	 [*][c]1:[cH]:[cH]:[cH]:n:[cH]:1	-0.482	0 out of 2
ECFP_6	764951226	 [*]1:[cH]:[cH]:[cH]:n:[cH]:1	-0.482	0 out of 2

Sorafenib

TOPKAT_Mouse_Female_FDA_None_vs_Carcinogen



$C_{21}H_{16}ClF_3N_4O_3$

Molecular Weight: 464.82494

ALogP: 4.175

Rotatable Bonds: 6

Acceptors: 4

Donors: 3

Model Prediction

Prediction: Carcinogen

Probability: 0.257

Enrichment: 0.801

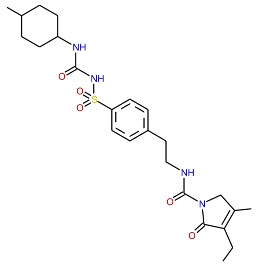
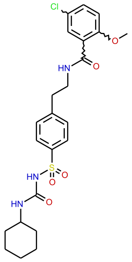
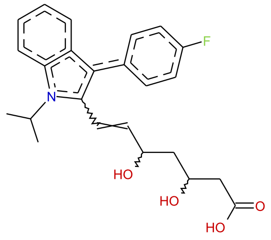
Bayesian Score: -0.321

Mahalanobis Distance: 14.9

Mahalanobis Distance p-value: 4.21e-007

Prediction: Positive if the Bayesian score is above the estimated best cutoff value from minimizing the false positive and false negative rate.
Probability: The esimated probability that the sample is in the positive category. This assumes that the Bayesian score follows a normal distribution and is different from the prediction using a cutoff.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Mahalanobis Distance: The Mahalanobis distance (MD) is the distance to the center of the training data. The larger the MD, the less trustworthy the prediction.
Mahalanobis Distance p-value: The p-value gives the fraction of training data with an MD greater than or equal to the one for the given sample, assuming normally distributed data. The smaller the p-value, the less trustworthy the prediciton. For highly non-normal X properties (e.g., fingerprints), the MD p-value is wildly inaccurate.

Structural Similar Compounds

Name	Glimepride	Glyburide	Fluvastatin
Structure			
Actual Endpoint	Carcinogen	Non-Carcinogen	Non-Carcinogen
Predicted Endpoint	Carcinogen	Non-Carcinogen	Non-Carcinogen
Distance	0.605	0.615	0.625
Reference	US FDA (Centre for Drug Eval.& Res./Off. Testing & Res.) Sept. 1997	US FDA (Centre for Drug Eval.& Res./Off. Testing & Res.) Sept. 1997	US FDA (Centre for Drug Eval.& Res./Off. Testing & Res.) Sept. 1997

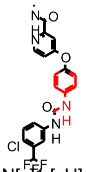
Model Applicability

Unknown features are fingerprint features in the query molecule, but not found or appearing too infrequently in the training set.

1. OPS PC20 out of range. Value: -3.3309. Training min, max, SD, explained variance: -3.1862, 4.4571, 1.28, 0.0167.

Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Carcinogen in training set
ECFP_6	738938915	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	0.617	2 out of 2