New quinoxaline-based VEGFR-2 inhibitors: design, synthesis, and antiproliferative evaluation with in silico docking, ADMET, toxicity, and DFT studies

Mohammed M. Alanazi[a], Hazem Elkady[b*], Nawaf A. Alsaif[a], Ahmad J. Obaidullah[a], Hamad M. Alkahtani[a], Manal M. Alanazi[a], Madhawi A. Alharbi[a], Ibrahim H. Eissa[b*], Mohammed A. Dahab[b*].

[a] Department of Pharmaceutical Chemistry, College of Pharmacy, King Saud University, P.O. Box 2457, Riyadh 11541, Saudi Arabia.

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

*** Corresponding authors:**

Mohammed A. Dahab

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

mohammeddahab@azhar.edu.eg

Ibrahim H. Eissa

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

ibrahimeissa@azhar.edu.eg

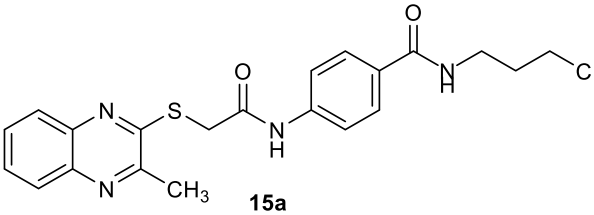
Hazem Elkady

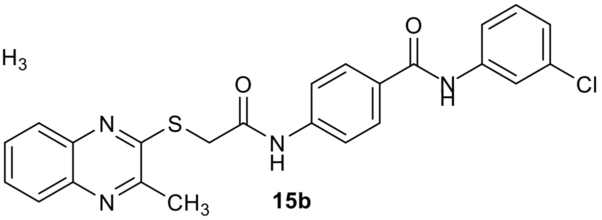
Pharmaceutical Medicinal Chemistry & Drug Design Department, Faculty of Pharmacy (Boys), Al-Azhar University, Cairo11884, Egypt

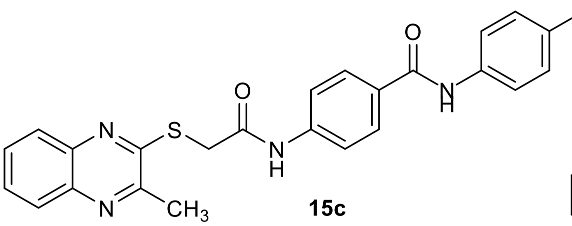
Hazemelkady@azhar.edu.eg

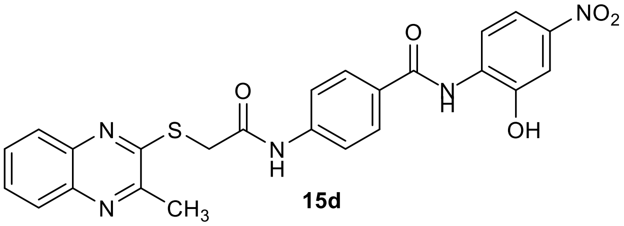
Content

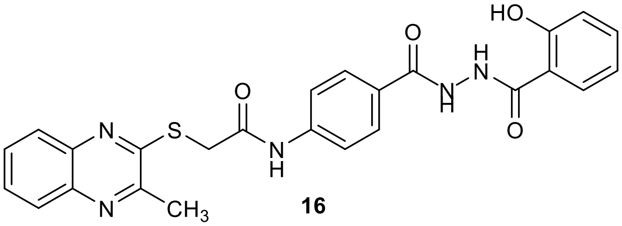
Chemical structures of the target compounds 15a-d, 16, 17a-d, and 18

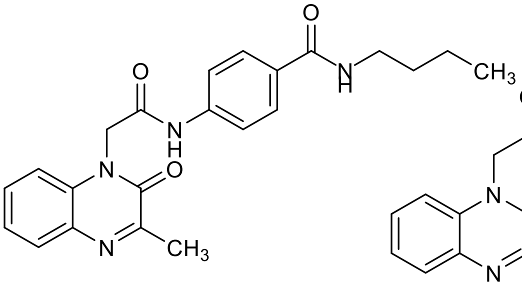

15a

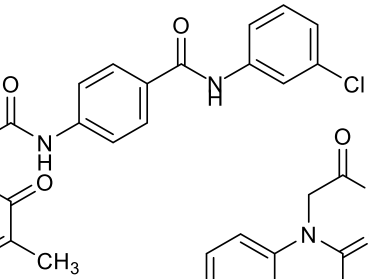

15b

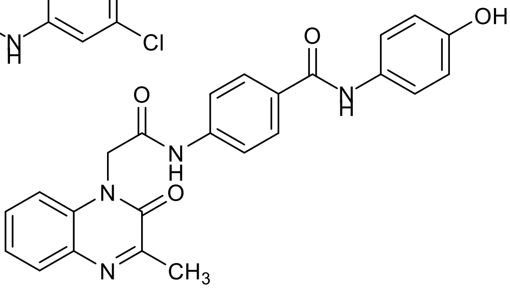

15c

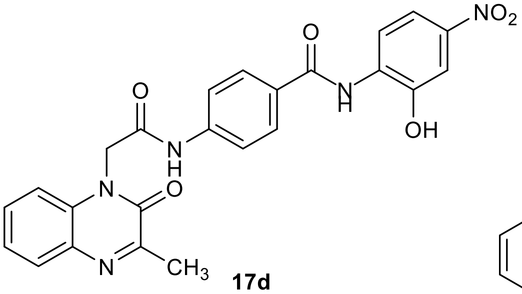

15d

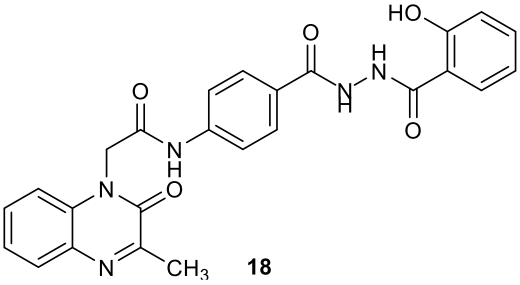

16


17a


17b


17c


17d


18

In silico studies

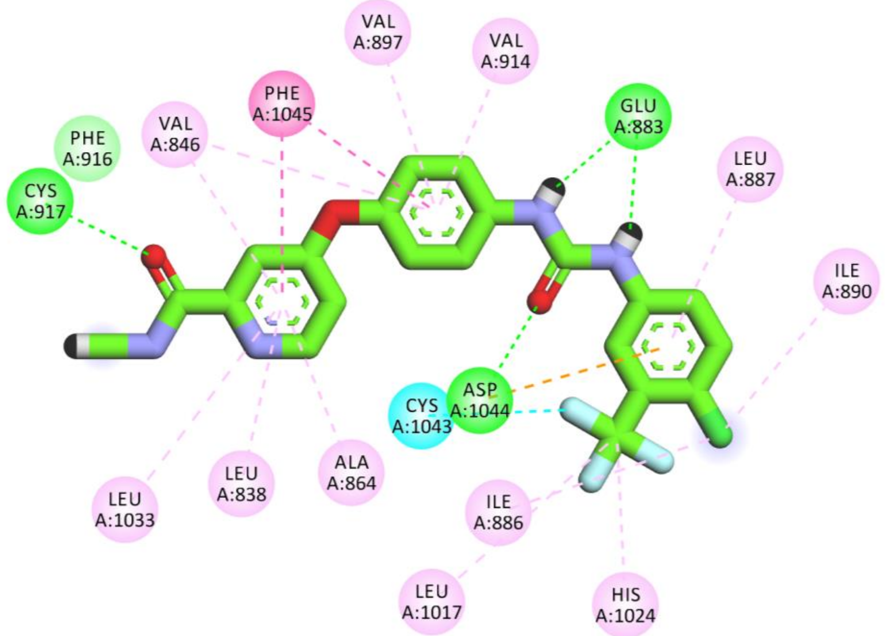


Fig. 1: Interaction of sorafenib with the essential amino acids inside VEGFR-2 active site (**2D**)

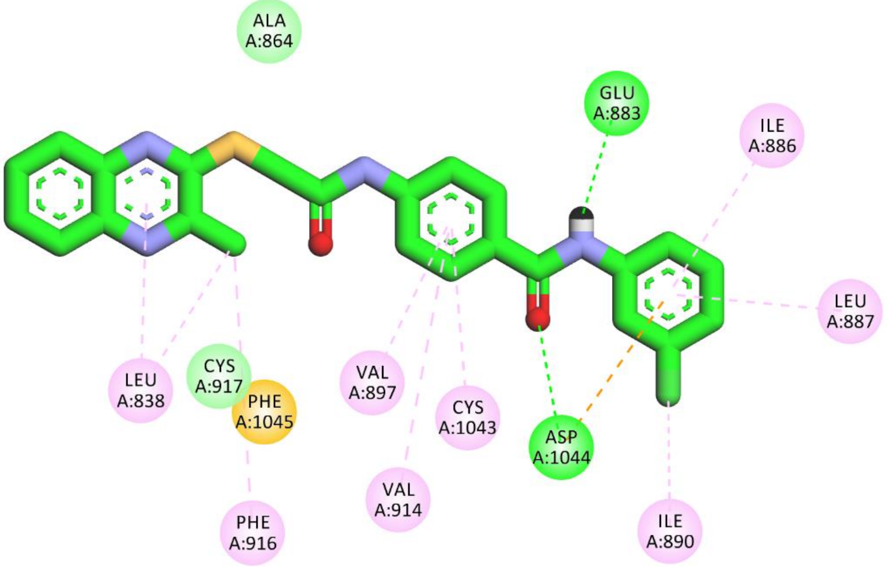


Fig. 2: Mapping surface of **15b.** Green dashed lines = hydrogen bonds, Orange lines = hydrophobic interactions (**2D**)

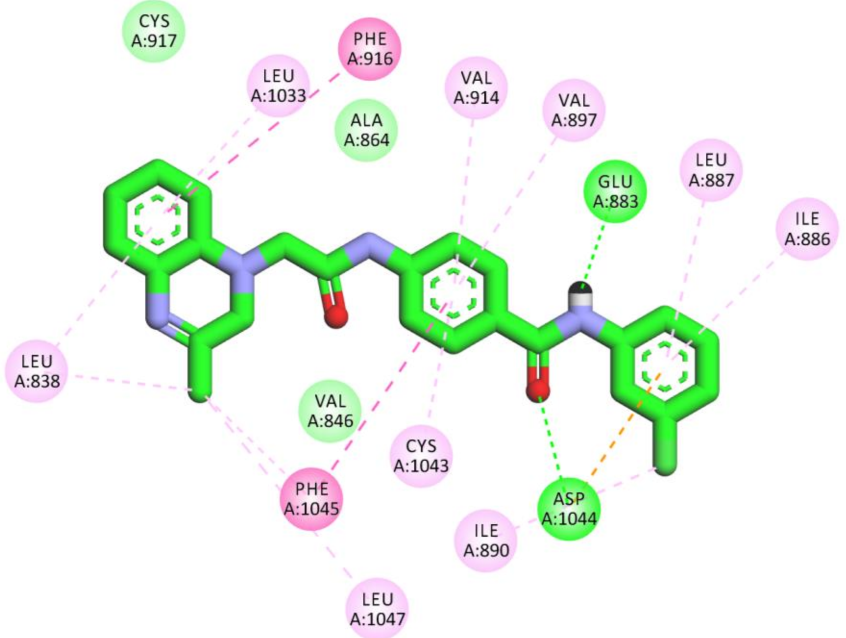


Fig. 3: Mapping surface of **17b**. Green dashed lines = hydrogen bonds, orange lines = hydrophobic interactions (**2D**).

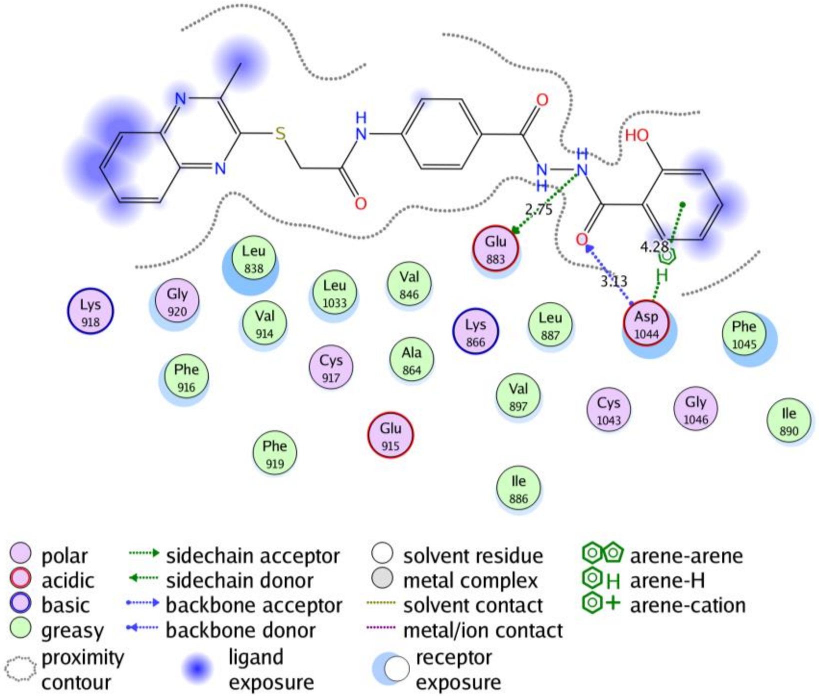


Fig. 4: The predicted binding pattern of compound 16 with the active site of VEGFR-2 (2D).

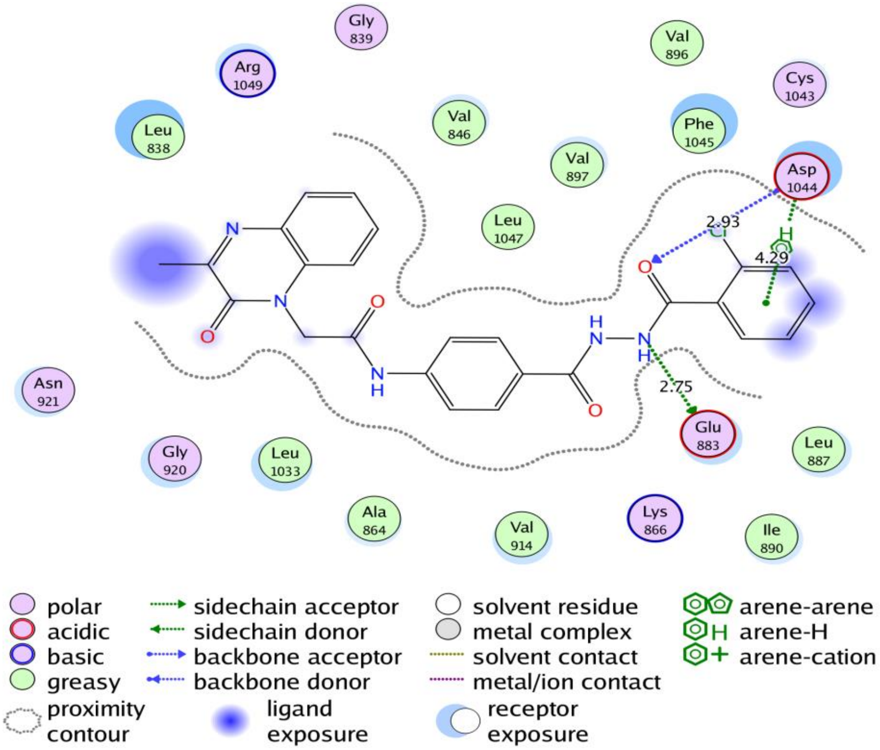


Fig. 5: The predicted binding pattern of compound 18 with the active site of VEGFR-2 (2D).

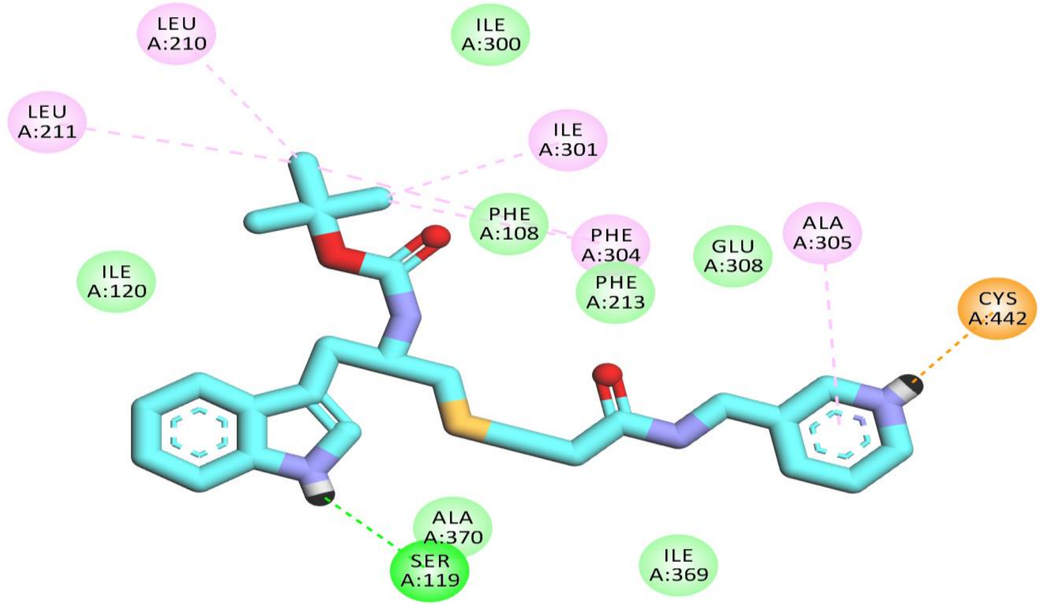


Fig. 6. 2D Structure of co-crystallized ligand (PKT) docked into active pocket of cytochrome P450.

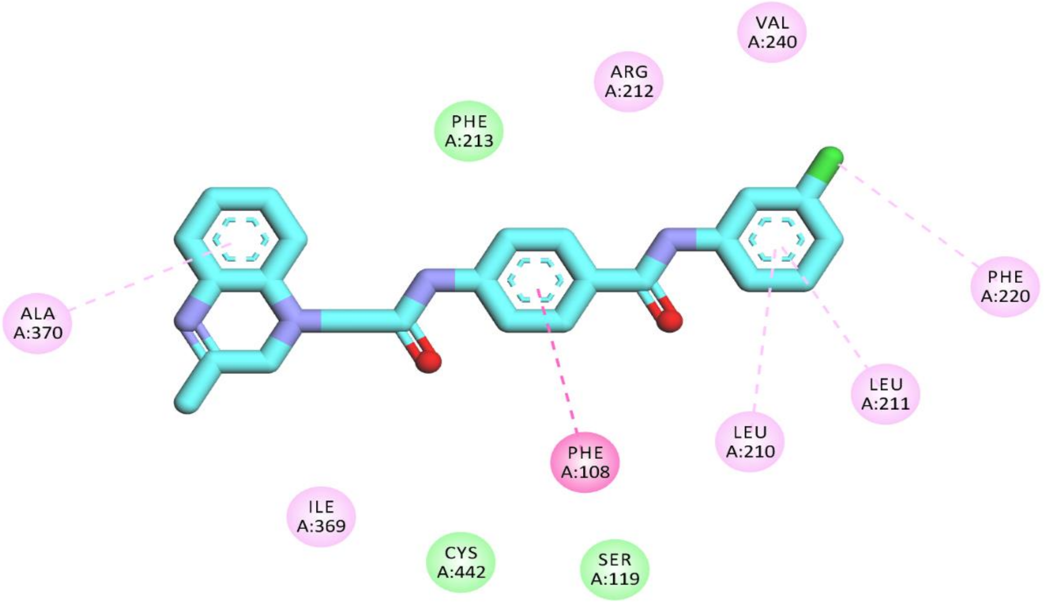


Fig. 7. 2D Structure of compound 17b docked into active pocket of cytochrome P450.

In silico ADMET studies

Blood brain barrier(BBB) penetration, intestinal absorption, aqueous solubility, CYP2D6 binding, and plasma protein binding properties of the synthesized compounds were calculated using Discovery studio 4.0. The BBB penetration levels of the tested compounds were in the low and very low range except compounds **15a** which was showed medium level of BBB penetration. Depending on these results, it may be concluded that there are no CNS side effects associated with these compounds. The tested compounds showed low or good levels of aqueous solubility. In addition, compounds **15a, 15b, 15c, 17a, 17b,** and **17c** showed good levels of intestinal absorption. For cytochrome P450 2D6 (CYP2D6) inhibition, all compounds were predicted as non-inhibitors. Finally, all compounds were expected to bind plasma protein less than 90% except **15b** and **17b** (**Table 1 & Fig. 8**)

Table 1. ADMET parameters for the synthesized compounds.

Comp.	BBB level [a]	Solubility level [b]	Absorption level [c]	CYP2D6 prediction [d]	PPB prediction [e]
15a	++	++	0	F	Less than 90%
15b	++++	++	0	F	More than 90%
15c	++++	++	0	F	Less than 90%
15d	++++	++	++	F	Less than 90%
16	++++	++	++	F	Less than 90%

17a	+++	+++	0	F	Less than 90%
17b	+++	++	0	F	More than 90%
17c	++++	+++	0	F	Less than 90%
17d	++++	+++	+++	F	Less than 90%
18	++++	+++	++	F	Less than 90%
Sorafenib	++++	+	0	F	More than 90%

[a] BBB level, blood brain barrier level, 0 = very high, + = high, ++ = medium, +++ = low, ++++ = very low.

[b] Solubility level, + = very low, ++ = low, +++ = good, ++++ = optimal.

[c] Absorption level, 0 = good, + = moderate, ++ = poor, ++ = very poor.

[d] CYP2D6, cytochrome P2D6, T = inhibitor, F = non inhibitor.

[e] PBB, plasma protein binding (less than 90% or more than 90%)

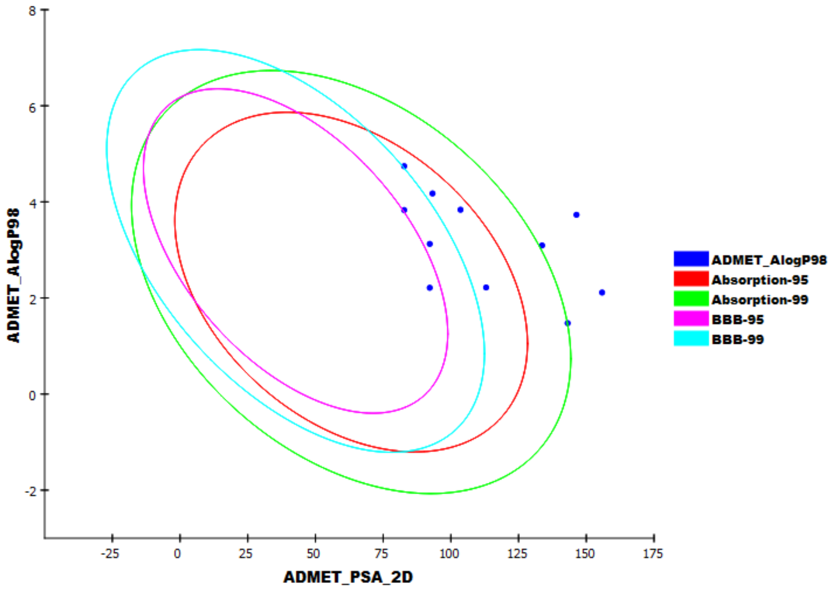


Fig. 8: The expected ADMET study

In silico toxicity studies

Toxicity profile of the synthesized compounds was predicted according to the built-in models of Discovery studio 4.0 software [1, 2]. Seven toxicity parameters were calculated and summarized in **Table 2.**

At first, the carcinogenic potency TD_{50} values (from 9.366 to 142.906 mg/kg body weight/day) of the tested compounds were higher than that of the reference molecule; sorafenib (TD_{50} = 19.236 mg/kg body weight/day), except compounds **15d** and **16**. In addition, the maximum tolerated dose values (from 0.096 to 0.333 g/kg body weight) of all compounds were higher than sorafenib (0.089 g/kg body weight). Furthermore, the tested compounds showed oral LD_{50} values ranging from 4.703 to 12.496 mg/kg body weight/day which were higher than that of sorafenib (0.823 mg/kg body weight/day). For rat chronic lowest observed adverse effect level (LOAEL), the tested molecules showed higher values (from 0.072 to 0.583 g/kg body weight) than sorafenib (0.005 g/kg body weight). Moreover, the tested compound was predicted to be mild irritant against eyes and non-irritant against skin. For aerobic biodegradability model, all compounds were anticipated to be non-degradable except compound **17a.**

Table 2: Toxicity profile of the synthesized compounds

Comp.	Carcinogenic Potency TD_{50} (Mouse) [a]	Rat Maximum Tolerated Dose (Feed) [b]	Rat Oral LD_{50} [b]	Rat Chronic LOAEL [b]	Ocular Irritancy	Skin Irritancy	Aerobic Biodegradability
15a	40.565	0.132	12.496	0.583	Mild	Non-Irritant	Non-Degradable
15b	9.366	0.123	4.969	0.114	Mild	Non-Irritant	Non-Degradable
15c	17.665	0.333	5.932	0.181	Mild	Non-Irritant	Non-Degradable
15d	31.468	0.244	6.811	0.091	Mild	Non-Irritant	Non-Degradable
16	16.061	0.364	4.703	0.382	Mild	Non-Irritant	Non-Degradable
17a	142.906	0.097	18.136	0.334	Mild	Non-Irritant	Degradable
17b	33.153	0.096	7.450	0.072	Mild	Non-Irritant	Non-Degradable
17c	62.436	0.259	7.868	0.132	Mild	Non-Irritant	Non-Degradable
17d	111.609	0.191	5.470	0.103	Mild	Non-Irritant	Non-Degradable
18	56.955	0.284	5.956	0.271	Mild	Non-Irritant	Non-Degradable
Sorafenib	19.236	0.089	0.823	0.005	Mild	Non-Irritant	Non-Degradable

[a] Unit: mg/kg body weight/day

[b] Unit: g/kg body weight

Physico-chemical properties

The A logP values express the degree of lipophilicity of the chemical compound, where the LogD values express the degree of lipophilicity of the chemical compound taking into account the ionization states of the molecule [3]. An increase in these values indicates an increase in the lipophilic character of the tested compound. It is worthwhile to note that the A log P and LogD values for most compounds in acceptable range for oral and intestinal absorption (1.44–4.74) [4].

In addition, the molecular polar surface area (MPSA) is another key property linked to drug bioavailability; the passively absorbed molecules with MPSA >140 have low oral bioavailability [5]. All examined compounds showed acceptable values of MPSA less than 140 except compounds **15d, 15, 17d** and **18**. Moreover, molecular volume (M.V) descriptor determines transport characteristics of molecules, such as intestinal absorption [6]. The drug diffusivity is inversely proportional to the molecular volume. Molecules with lower MV have higher diffusivity [7]. It was observed that the tested compounds exhibited low molecular volume values (from 315.21 to 353.63) when compared with sorafenib (MV = 323.1).

Finally, Lipinski rule of five was applied for the synthesized compounds. It was found that all the compounds have molecular weight less than 500, hydrogen bond acceptor groups less than 10 except compounds **17d** and **18**, and hydrogen bond donor group less than 5. this indicates that most of the synthesized compounds are likely to be orally bioavailable (**Table 3)**.

Table 3: Physico-chemical properties of the synthesized compounds and sorafenib.

Comp.	ALog P [a]	Log D[b]	MPSA[c]	MSA[d]	MV[e]	HBA[f]	HBD[g]	M. WT[h]
15a	3.83	3.83	109.28	426.2	330.3	6	2	408.517
15b	4.74	4.74	109.28	444.39	339.56	6	2	462.951
15c	3.83	3.24	129.51	434.98	336.82	7	3	444.506
15d	3.73	3.48	175.33	472.83	353.28	10	3	489.503
16	3.09	3.05	158.61	471.6	362.2	9	4	487.53
17a	2.21	2.21	90.87	412.18	315.21	7	2	392.451
17b	3.12	3.12	90.87	430.37	328.25	7	2	446.886
17c	2.21	1.62	111.1	420.96	325.5	8	3	428.44
17d	2.11	1.86	156.91	458.8	342.99	11	3	473.438
18	1.47	1.44	140.19	457.57	353.63	10	4	471.465

Sorafenib	4.17	4.17	92.35	434.9	323.1	7	3	464.825

[a] Log of the octanol-water partition coefficient

[b] The octanol-water partition coefficient calculated taking into account the ionization states of the molecule

[c] Molecular surface area: Calculates the total surface area for each molecule using a 2D approximation.

[d] Molecular polar surface area: calculates the polar surface area for each molecule using a 2D approximation.

[e] Molecular volume: calculates the 3D volume for each molecule using the current 3D coordinates.

[f] Hydrogen bond donor atoms

[g] Hydrogen bond acceptor atoms

[h] Molecular weight

DFT studies

Discovery studio software was used to carry out density functional theory (DFT) calculations. Different molecular and atomic properties were calculated including **i**) total energy of the molecules, **ii**) binding energy which describes the interaction energy between all the atoms in the molecule, **iii**) the energy of the highest occupied molecular orbital (HOMO), **iv**) the energy of the lowest unoccupied molecular orbital (LUMO), gap energy which describes the energy difference between LUMO and HOMO, **v**) the magnitude of the dipole moment (μ).

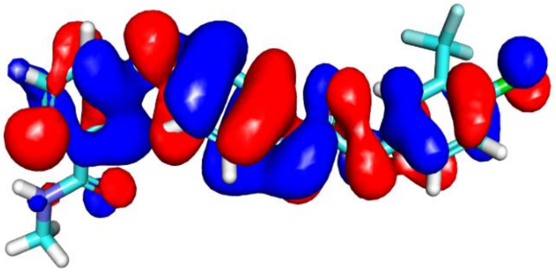
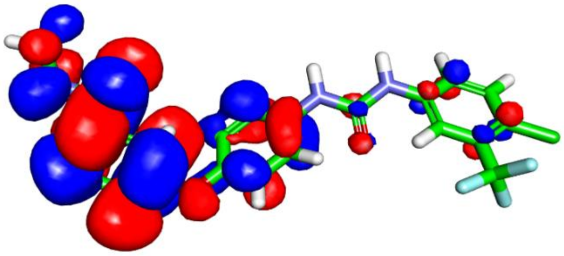
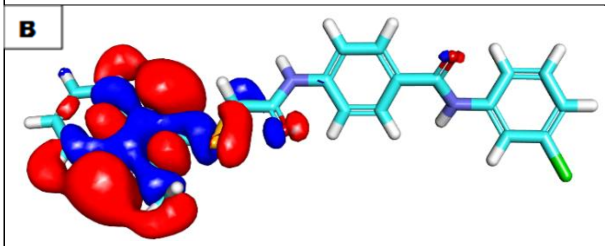
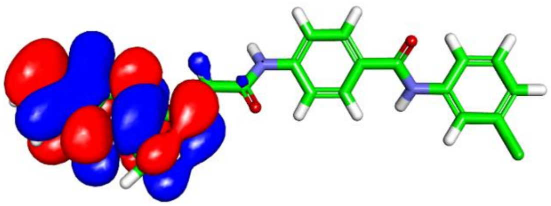
The results (**Table 4**) revealed that the total energies of the tested compounds and sorafenib have negative values which are favorable for spontaneous binding and interaction. In addition, compounds **15b, 15d, 16, 17b**, **17c, 17d,** and **18** have dipole moment values very close to that of sorafenib. The improved dipole moment can enhance hydrogen bond and non-bonded interactions in drug receptor complexes which keep an important role to increase binding affinity. Elevated dipole moment indicated the increased binding affinity with target enzyme during VEGFR-2 inhibitory activities.

Molecular orbital analysis

According to the frontier molecular orbital theory, the energies of HOMO and LUMO play an important role in chemical reactivity [8]. It was evident that compounds **15a, 15b, 17a, 17b, 17c,** and **18** have gap energy values very close to that of sorafenib. **Fig. 9** showed the spatial distribution of molecular orbitals for sorafenib and the most promising members (**15b** and **17b**).

Table 4. Thermodynamic parameters of **the synthesized compounds and sorafenib.**

Name	Total Energy (kcal/mol)	Binding Energy (kcal/mol)	HOMO Energy (kcal/mol)	LUMO Energy (kcal/mol)	Gap Energy	μ
15a	-1608.474	-10.462	-0.200	-0.107	0.093	0.530
15b	-2139.791	-10.576	-0.200	-0.107	0.093	2.365
15c	-1756.361	-10.806	-0.183	-0.106	0.077	1.220
15d	-1959.422	-11.164	-0.205	-0.132	0.074	4.484
16	-1923.712	-11.558	-0.199	-0.122	0.077	2.095
17a	-1286.345	-10.666	-0.190	-0.094	0.097	1.734
17b	-1817.724	-10.717	-0.201	-0.103	0.098	3.061
17c	-1434.226	-11.008	-0.175	-0.093	0.082	2.298
17d	-1637.285	-11.367	-0.196	-0.124	0.072	4.801
18	-1601.579	-11.769	-0.192	-0.095	0.097	2.520
Sorafenib	-2000.377	-9.866	-0.200	-0.091	0.109	3.088

HOMO	LUMO
A 	
B 	
C	

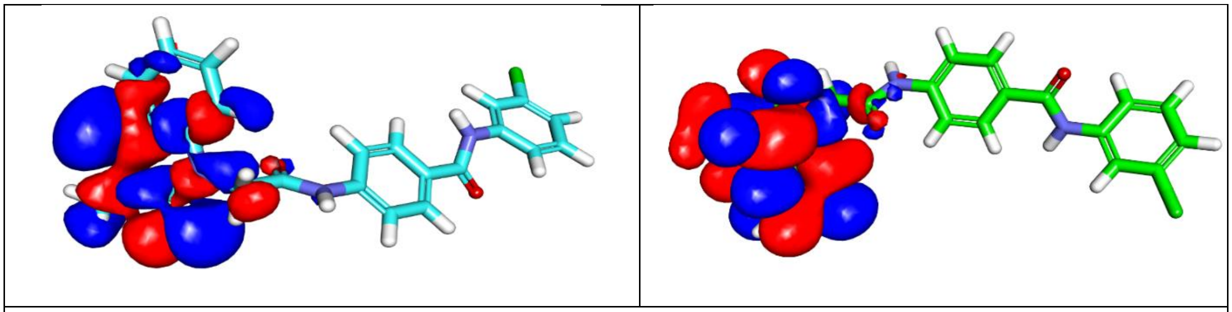
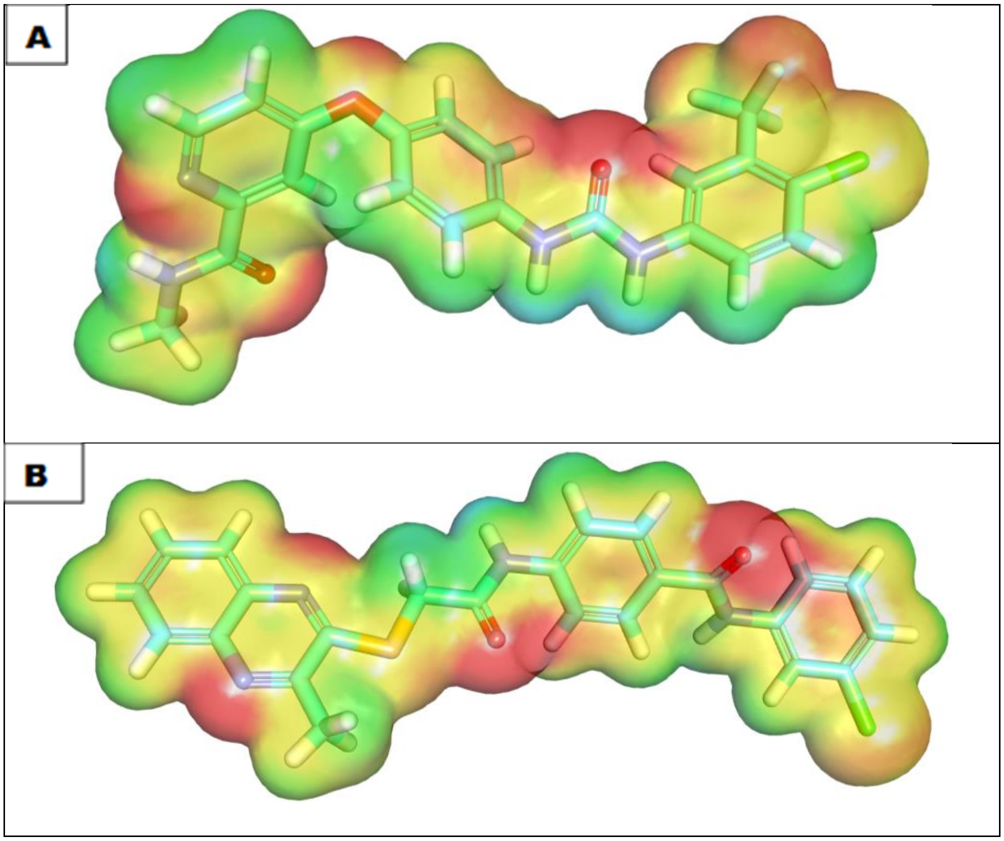


Fig. 9. Spatial distribution of of molecular orbitals for **A**) sorafenib, **B**) compound **15b,** and **C**) compound **17b**.



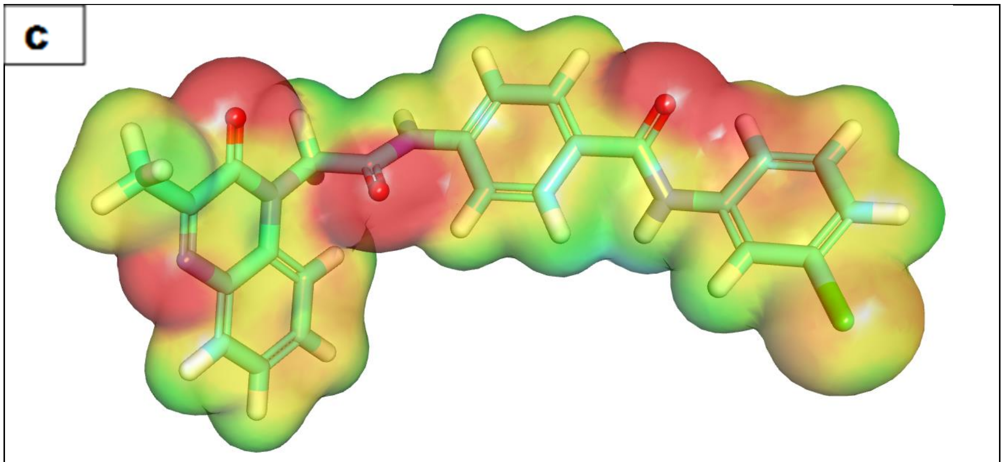


Fig. 10. Molecular electrostatic potential map of (A) sorafenib, (B) compound **15b**, and (C) compound **17b**.

Experimental section

1. Chemistry

All solvents and reagents were commercially available and used without further purification. Reactions progression were monitored by TLC sheets coated with UV fluorescent silica gel (Kieselgel 0.25mm, 60 F254, Merck Germany) and were visualized using UV lamp. The melting points were determined using a Gallen lamp melting point apparatus. Elemental analyses were accomplished using a CHN analyzer. The infrared spectra were recorded on FT/IR-6600typeA spectrophotometer. ^{1}H NMR spectra were recorded at 400 and 700 MHz, while ^{13}C NMR spectra were run at 100 and 176 MHz, on a Bruker Avance NEO-600 equipped with a 1.7 mm TCI CryoProbe. Chemical shifts were expressed in δ (ppm) with reference to TMS and coupling constant (J) in Hertz using DMSO-d_6 and $CDCl_3$-d_6 as solvents. The mass spectra were recorded on an Agilent 6410 triple-quadrupole mass spectrometer equipped with an ESI source.

2. Biological testing

In vitro anti-proliferative activity

Anti-proliferative activity screening of the newly synthesized compounds was carried out against two human cancer cell lines namely, breast cancer (MCF-7) and Hepatocellular carcinoma (HepG-2). The cell lines were obtained from ATCC (American Type Culture Collection) via the Holding company for biological products and vaccines (VACSERA) (Cairo, Egypt). The anti-cancer activity was measured quantitatively using MTT assay protocol. Cell lines were cultured in

RPMI-1640 medium with 10% fetal bovine serum. Antibiotics added were 100 units/ml penicillin and100 µg/ml streptomycin at 37ºC in a 5% CO_2 incubator. The cell lines were seeded in a 96-well plate at a density of 1.0 x 104 cells / well at 37 ºC for 48 h under 5% CO_2. After incubation, the cells were treated with different concentration of synthesized compounds and incubated for 24 h. After 24 h of drug treatment, 20 µl of MTT solution at 5mg/ml was added and incubated for 4 h. Dimethyl sulfoxide (DMSO) in volume of 100 µl was added into each well to dissolve the purple formazan formed. The colorimetric assay was measured and recorded at absorbance of 570 nm using a plate reader (EXL 800, USA). The relative cell viability in percentage was calculated as (A570 of treated samples/A570 of untreated sample) X 100. Results for IC_{50}values of the active compounds were summarized in **Table 1**.

In vitro VEGFR-2 kinase assay

Inhibitory activity of the newly synthesized compounds against VEGFR-2 was evaluated using Human VEGFR-2 ELISA kit (Enzyme-Linked Immunosorbent Assay). A specific antibody for VEGFR-2 was seeded on a 96-well plate and 100 µL of the standard solution or the tested compound was added, all were incubated at room temperature for 2.5 hours. Then washed, 100 µL of the prepared biotin antibody was added, then incubated at room temperature for additional 1 hour. Washed, 100 µL of streptavidin solution was added then incubated for 45 min. at room temperature. Washed again, 100 µL of TMB Substrate reagent was added and incubated for 30 min. at room temperature. 50 µL of the stop solution was added, then read at 450 nm immediately. The standard curve was drawn, concentrations on the X-axis and the absorbance on the Y-axis.

Cell cycle analysis

HepG-2 cells were seeded and incubated for 24 h in six-well plates each one contains 2 x 10^5 cells per well. Fetal bovine serum (FBS, 10%) was added, after that cells were incubated at 37ºC and 5% CO_2. The medium was replaced with (DMSO 1% v/v) containing the 2.5 µM of compound **17b**, then incubated for 48 h, washed with cold phosphate buffered saline (PBS), fixed with 70% ethanol, rinsed with PBS then stained with the DNA fluorochrome PI, kept for 15 min at 37ºC. Then samples were analyzed with a FACS Caliber flow cytometer.

Annexin V-FITC apoptosis assay

Annexin V-FITC/PI apoptosis detection kit was used for this investigation, **HepG-2** cells were stained with Annexin V fluorescein isothiocyanate (FITC) and counterstained with propidium iodide (PI). 2 x 10^5 of HepG-2 cells were incubated with compound **17b** for 48 h, trypsinized, washed with phosphate-buffered saline (PBS), stained for 15 min at 37°C in the dark. Then, analyzed with FACS Caliber flow cytometer.

Induction of apoptosis

In this test HepG-2 cells were treated with compound **17b** (2.3 µM) and analyzed for apoptosis using flowcytometry technique [9]. The obtained results showed that compound **17b** induced early apoptosis (35.80%) by more than 4-fold over the control (8.22%) **Table 3** and **Fig. 4.**

Table 3: Effect of compound **17b** on stages of the cell death (HepG-2).

Sample	**Viable** [a] (Left Bottom)	**Apoptosis** [a]		**Necrosis** [a] (Left Top)
		Early (Right Bottom)	**Late** (Right Top)	
HepG-2	91.44 ± 2.55	8.22 ± 0.61	0.26 ± 0.04	0.11 ± 0.03
17b /HepG-2	63.45 ± 3.18	35.80 ± 3.18**	0.52 ± 0.09	0.20 ± 0.04

[a] Three independent experiments were applied for each value. $^{**}p < 0.01$.

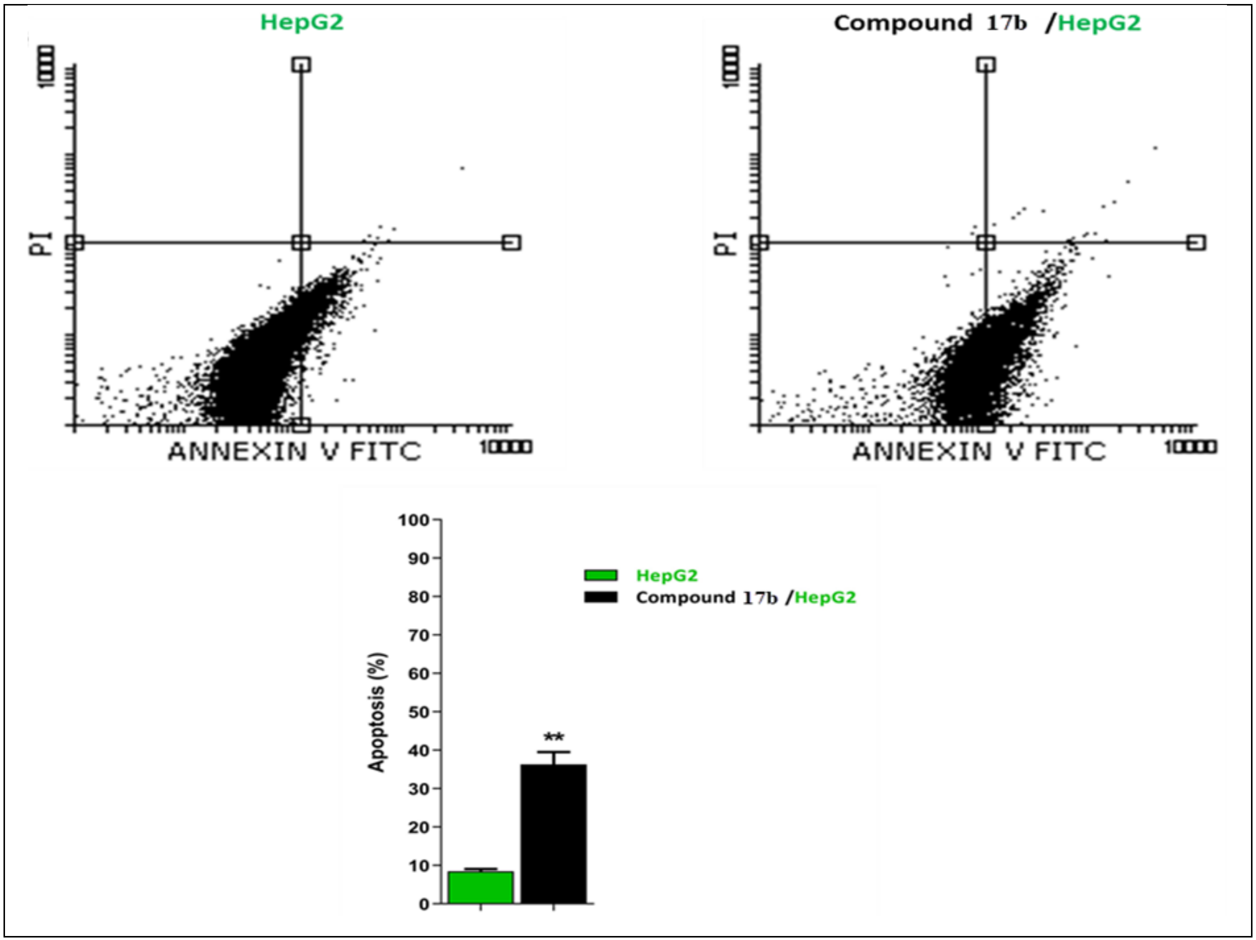


Fig. 11: Apoptosis analysis in HepG-2 cells after the treatment with compound **17b,** $^{**}p < 0.01$.

Caspase-3, and -9 determination

The percentage of caspase-3 activation was determined using the Caspase- Invitrogen Caspase-3 ELISA Kit (KHO1091) following the manufacturer's instructions.

Bax determination (real-time PCR).

The effect of compound **17b** on *BAX* expressions was determined at concentrations of 6.71, 8.68, 46.43, 4.55, and 3.43 µg/mL, respectively after 48 h of treatment on HepG-2 cells, using real-time PCR. The expression level of BAX was determined using a SYBR Green qRT-PCR kit (Bioneer, Daejeon, Korea) according to the manufacturer's protocol. PCR reactions were performed using specific primers for *BAX*. *BAX* forward primer: 5′-GCCCTTTTGCTTCAGGGTTT-3′; *BAX* Reverse primer: 5′ TCCAATGTCCAGCCCATGAT-3′. The reactions were carried out in triplicate using an Eco Real-Time PCR system (San Diego, CA, USA).

3. *In silico* studies

Molecular docking procedures

Molecular Operating Environment MOE, package version 2014.09 software was used for computational analysis. The crystal structure of VEGFR-2 was downloaded from the Protein Data Bank, http://www.rcsb.org/pdb (PDB ID: 2OH4, resolution: 2.05 Å). The crystal structures were imported into MOE, the structure preparation wizard of MOE was used to correct all the issues in protein structures. The water molecules were removed from downloaded protein then subjected to 3D protonation and energy minimization up to 0.01 gradient. Protein was subjected to energy minimization by applying MMFF94 force field, the binding site of the protein was defined and prepared for docking. The final optimized structures were saved in the working directory. Sorafenib and the designed compounds 2D structures were sketched using ChemBioDraw Ultra 14.0 and saved in MDL Molfile format. Next, the Molfile was opened, 3D structures were protonated, and energy minimized by applying MMFF94 force field then prepared for docking by optimizing the parameters, using ASE as score function. A maximum of 10 conformers was considered for each molecule in the docking analysis. After that, the docking scores (S) of the best -fitted conformation of each of the docked molecules with the amino acids at the VEGFR-2 binding pocket were recorded. The obtained compound–receptor complexes were then used to study the predicted ligand-receptor attachments at the target sites and their binding energies.

The same procedure was carried out for the docking studies of the most active compound 17b against cytochrome P450 (PDB ID: 4D7D).

ADMET studies

ADMET descriptors (absorption, distribution, metabolism, excretion and toxicity) of the compounds were determined using Discovery studio 4.0. At first, the CHARMM force field was applied then the tested compounds were prepared and minimized according to the preparation of small molecule protocol. Then ADMET descriptors protocol was applied to carry out these studies.

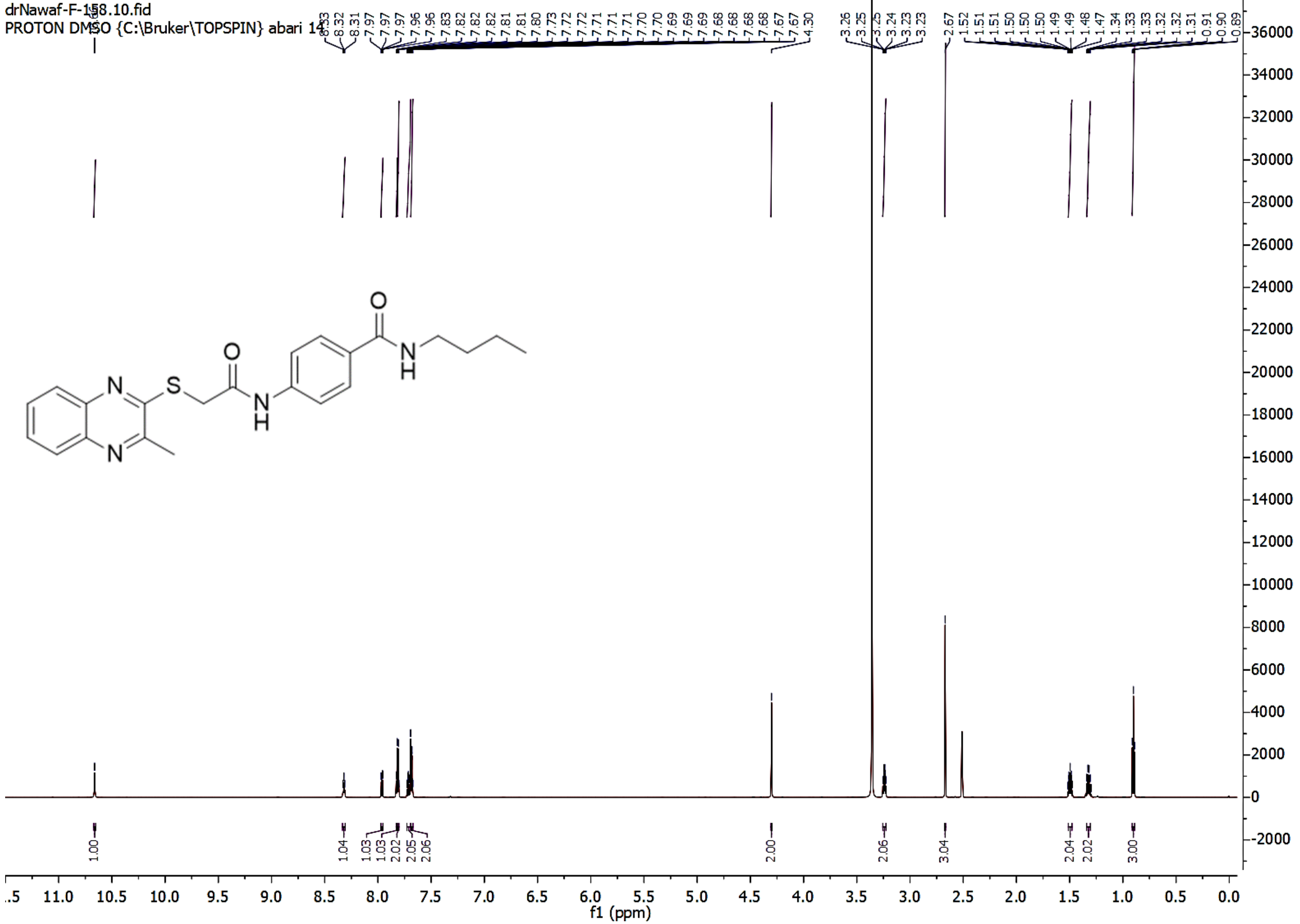
Toxicity studies

The toxicity parameters of the synthesized compounds were calculated using Discovery studio 4.0. Sorafenib was used as a reference drug. At first, the CHARMM force field was applied then the compounds were prepared and minimized according to the preparation of small molecule protocol. Then different parameters were calculated from toxicity prediction (extensible) protocol.

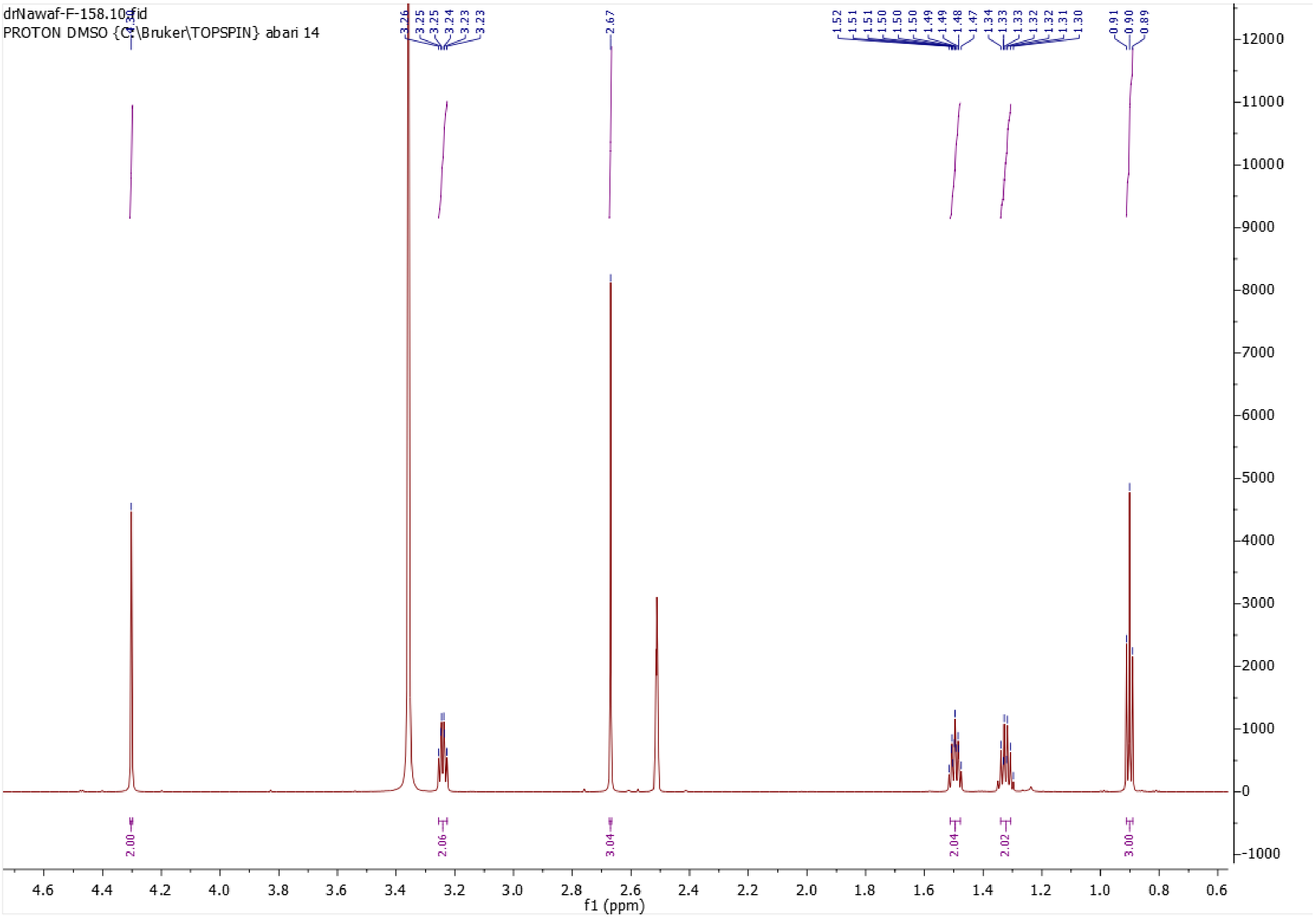
DFT studies:

The DFT parameters (total energy, binding energy, HOMO, LUMO, gap energy, dipole moment, and electrostatic potential) were calculated using Discovery studio software. the tested compounds were prepared using prepare ligand protocol. Then, the prepared compounds were subjected to DFT calculation protocol using the default option.

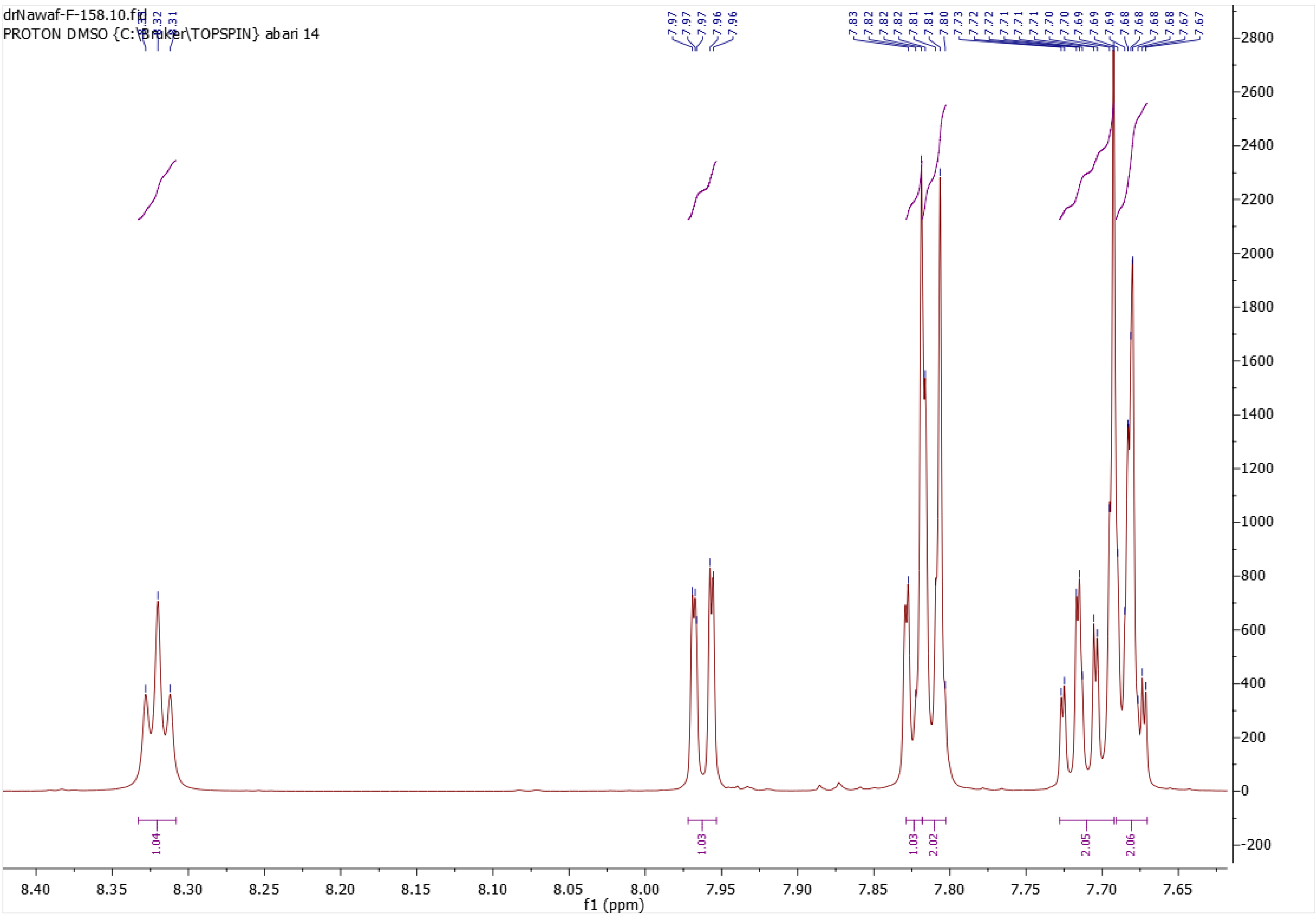
15a


drNawaf-F-158.10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 14
8.33 8.32 8.31 7.97 7.97 7.97 7.96 7.96 7.83 7.82 7.82 7.82 7.81 7.81 7.80 7.73 7.72 7.72 7.71 7.71 7.71 7.70 7.70 7.69 7.69 7.69 7.68 7.68 7.68 7.68 7.67 7.67 4.30
3.26 3.25 3.25 3.24 3.23 3.23
2.67 1.52 1.51 1.51 1.50 1.50 1.50 1.49 1.49 1.48 1.47 1.34 1.33 1.33 1.32 1.32 1.31 0.91 0.90 0.89
N
N
S
O
N
H
O
N
H
1.00
1.04
1.03
1.03
2.02
2.05
2.06
2.00
2.06
3.04
2.04
2.02
3.00
.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 0.0
f1 (ppm)
36000 34000 32000 30000 28000 26000 24000 22000 20000 18000 16000 14000 12000 10000 8000 6000 4000 2000 0 -2000

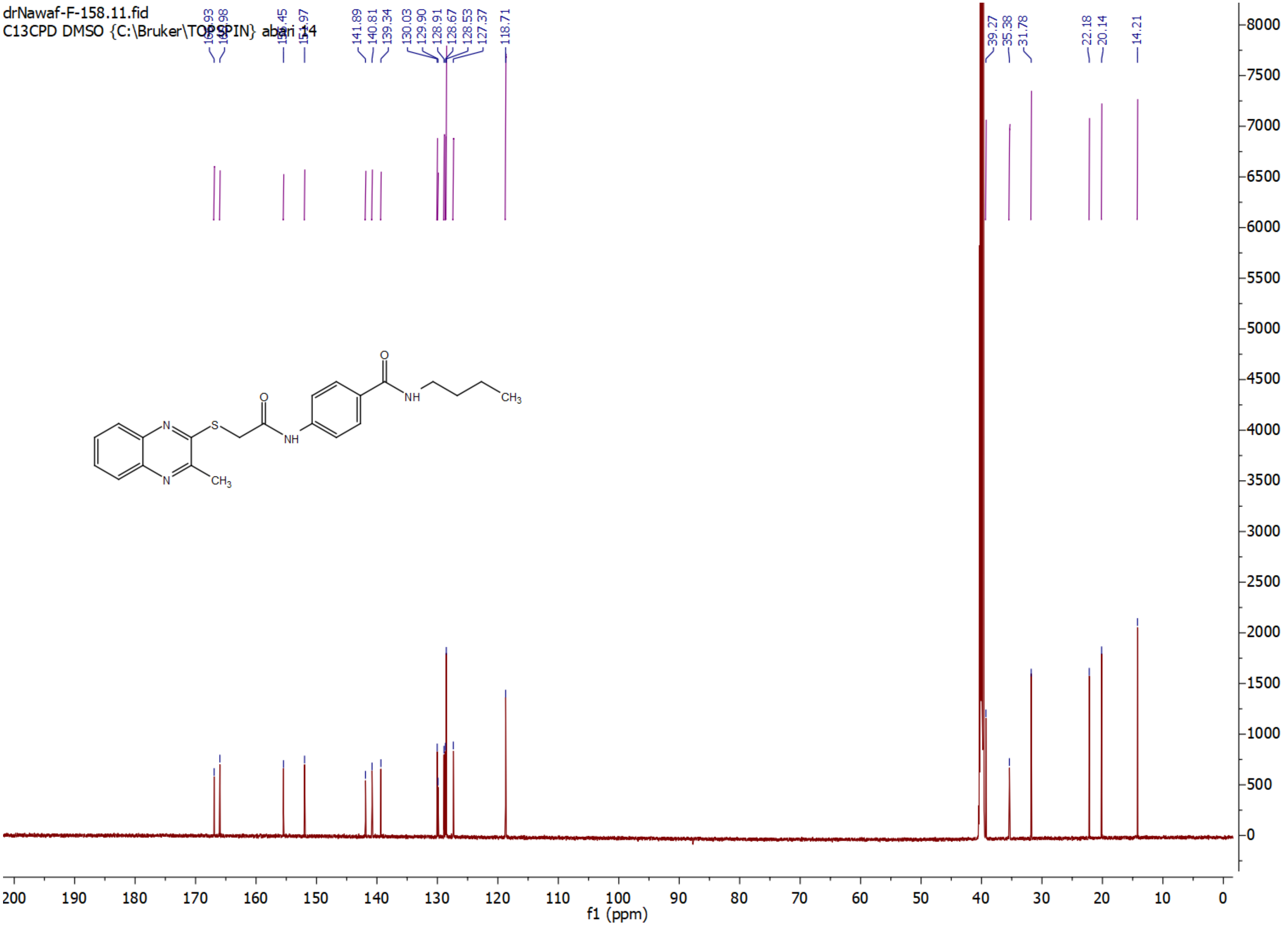
15a


drNawaf-F-158.10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 14
4.30
3.26
3.25
3.25
3.24
3.23
3.23
2.67
1.52
1.51
1.51
1.50
1.50
1.50
1.49
1.49
1.48
1.47
1.34
1.33
1.33
1.32
1.32
1.31
1.30
0.91
0.90
0.89
2.00
2.06
3.04
2.04
2.02
3.00
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000
4.6
4.4
4.2
4.0
3.8
3.6
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
f1 (ppm)

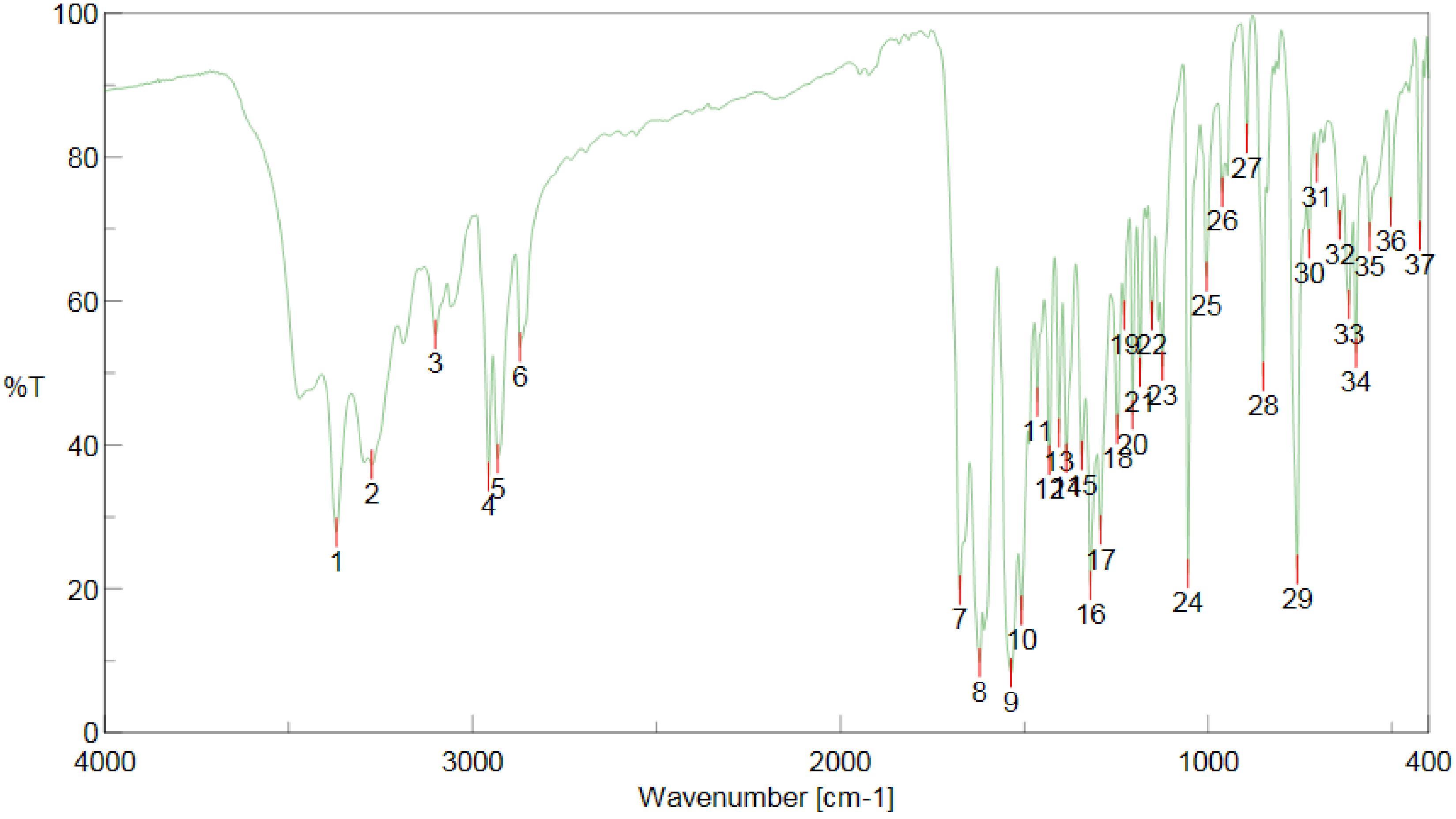
15a


drNawaf-F-158.10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 14
8.33
8.32
8.31
7.97
7.97
7.97
7.96
7.96
7.83
7.82
7.82
7.82
7.81
7.81
7.80
7.73
7.72
7.72
7.71
7.71
7.71
7.70
7.70
7.69
7.69
7.69
7.68
7.68
7.68
7.68
7.67
7.67
1.04
1.03
1.03
2.02
2.05
2.06
8.40
8.35
8.30
8.25
8.20
8.15
8.10
8.05
8.00
7.95
7.90
7.85
7.80
7.75
7.70
7.65
f1 (ppm)
2800
2600
2400
2200
2000
1800
1600
1400
1200
1000
800
600
400
200
0
-200

15a

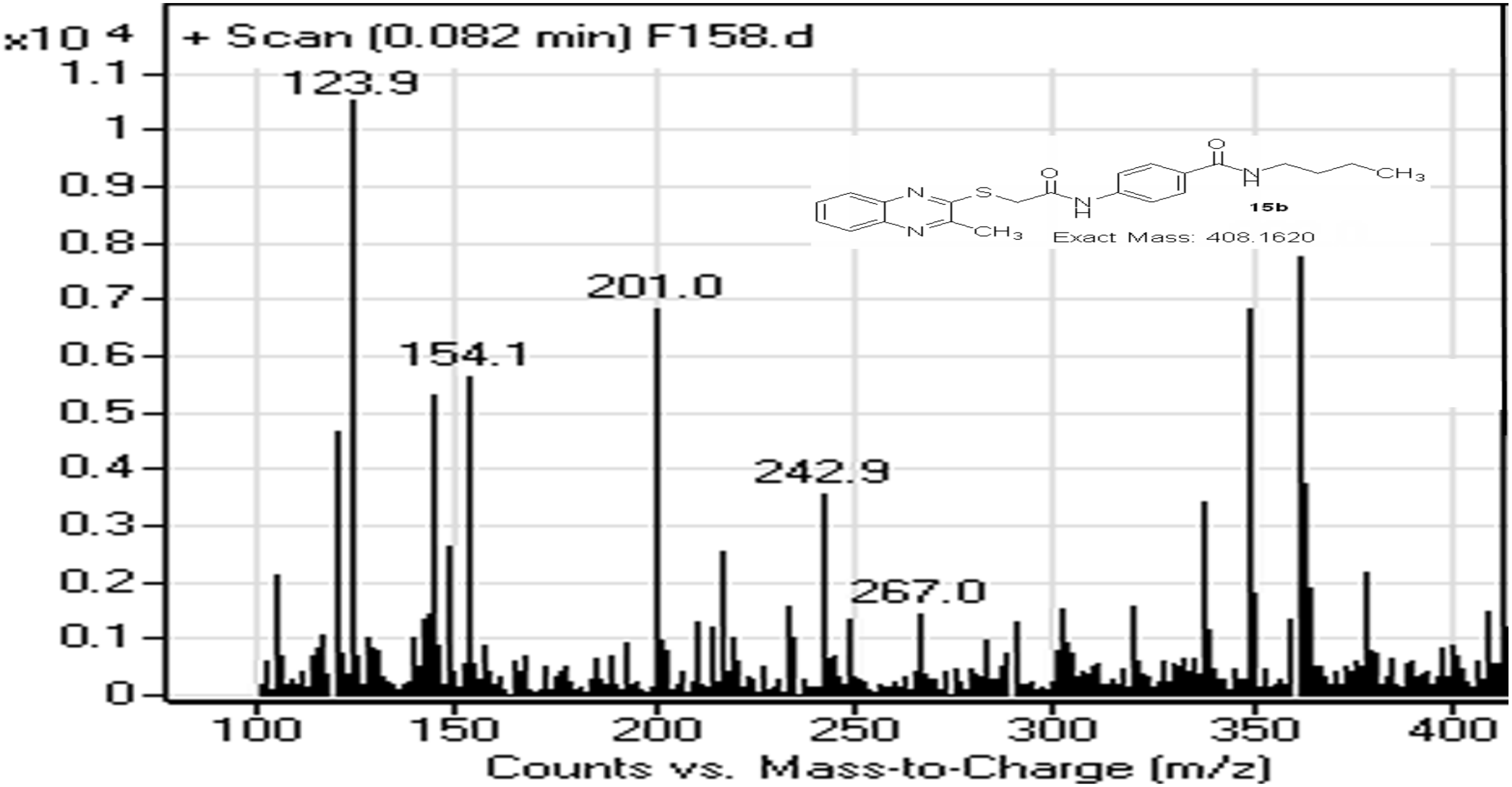

drNawaf-F-158.11.fid
C13CPD DMSO {C:\Bruker\TOPSPIN} abari 14
166.93
165.98
155.45
151.97
141.89
140.81
139.34
130.03
129.90
128.91
128.67
128.53
127.37
118.71
39.27
35.38
31.78
22.18
20.14
14.21
N
N
S
O
NH
O
NH
CH3
CH3
8000
7500
7000
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

15a

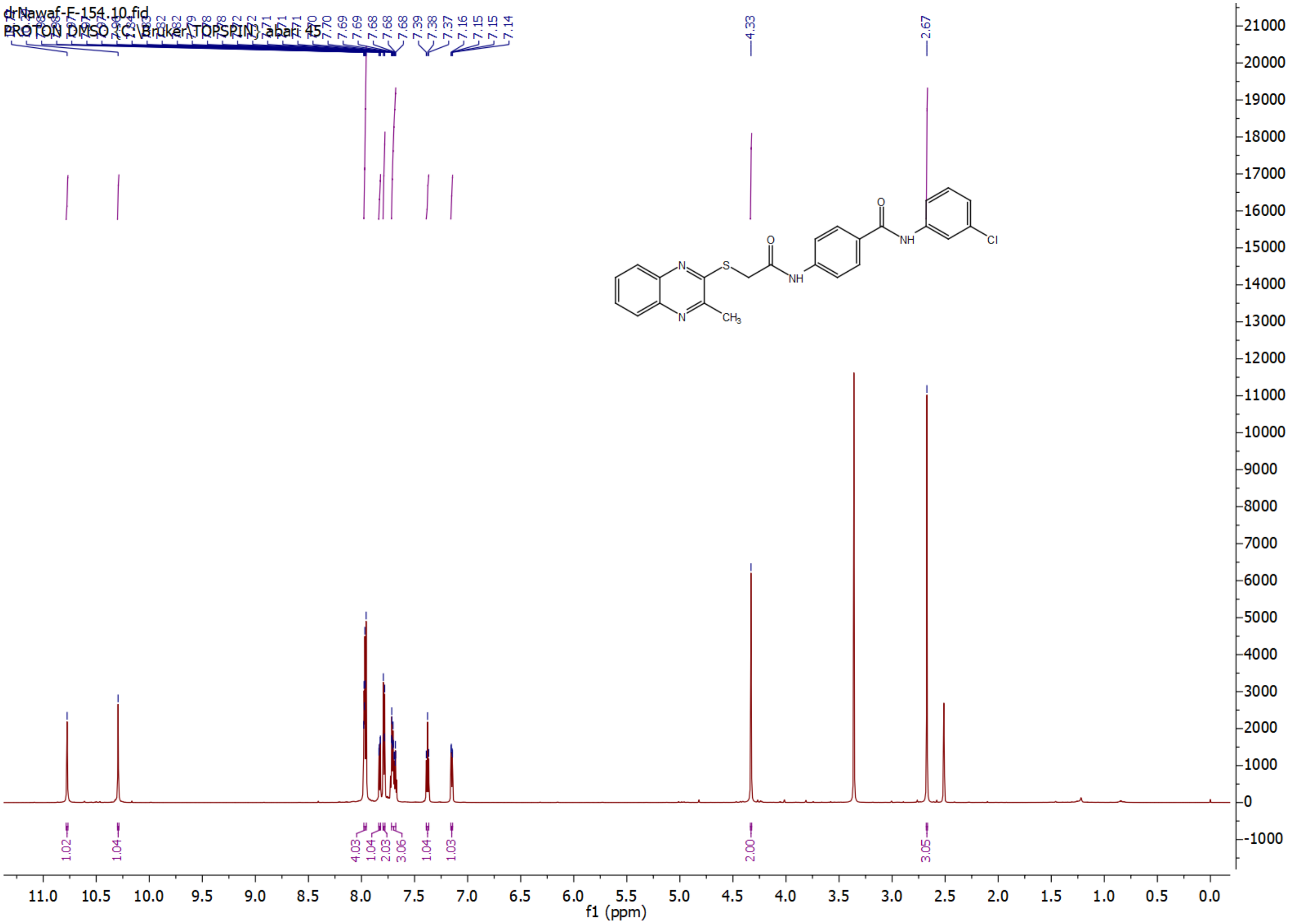


100
80
60
40
20
0
%T
4000
3000
2000
1000
400
Wavenumber [cm-1]
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

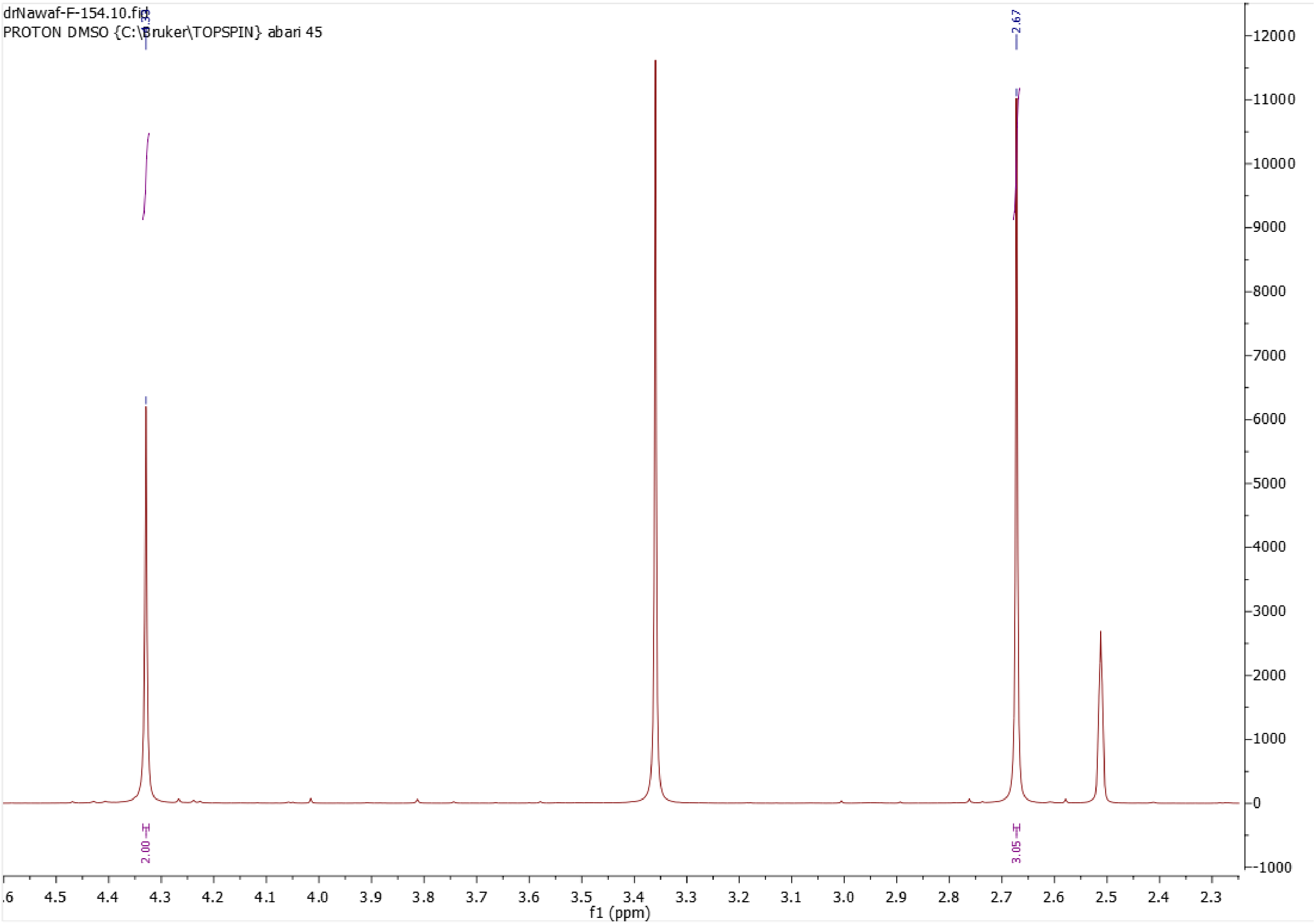
15a


x10 4
+ Scan (0.082 min) F158.d
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
123.9
154.1
201.0
242.9
267.0
100
150
200
250
300
350
400
Counts vs. Mass-to-Charge (m/z)
N
N
S
CH3
O
N
H
O
N
H
CH3
15b
Exact Mass: 408.1620

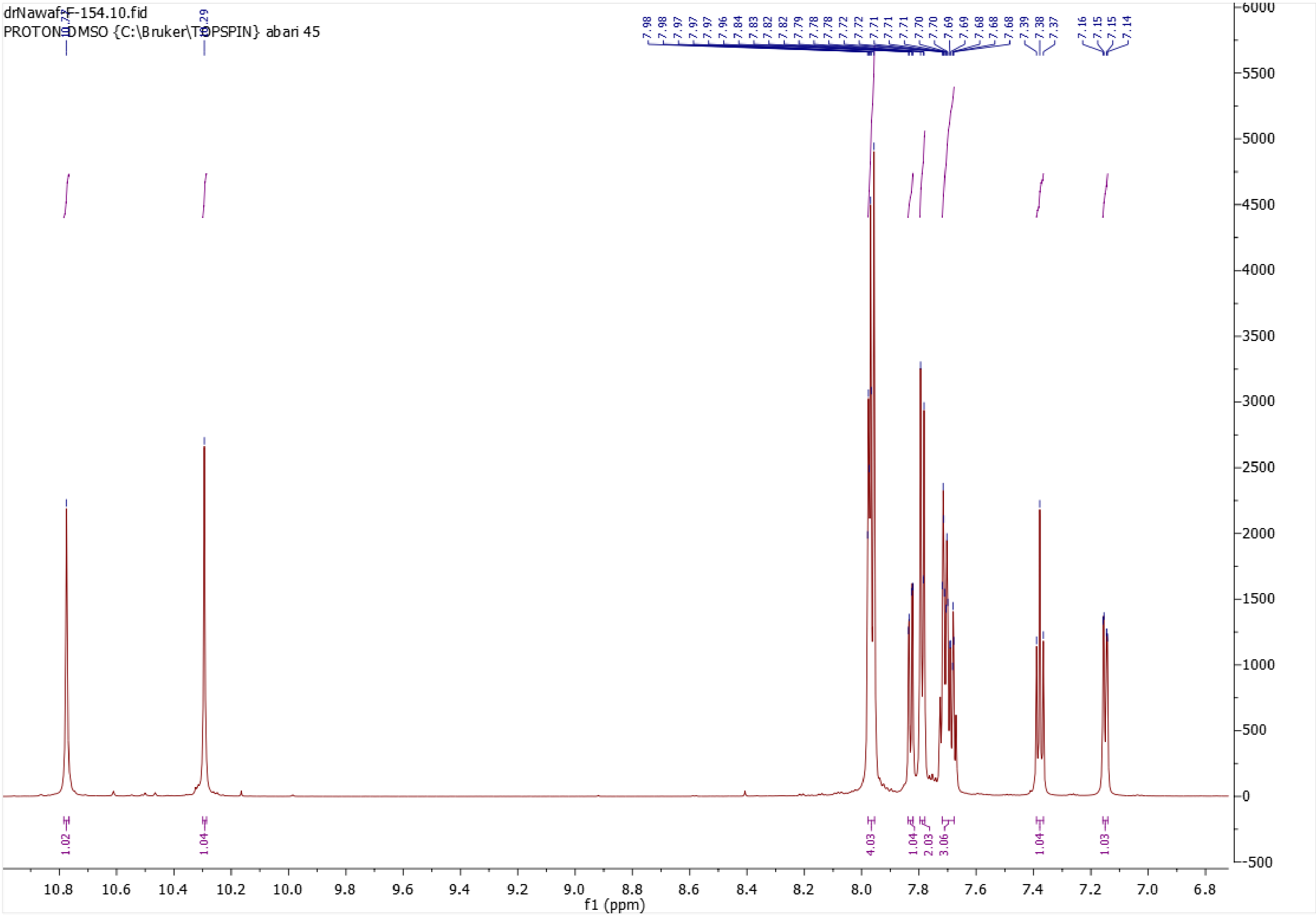
15b


drNawaf-F-154.10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 45
7.82
7.82
7.79
7.78
7.78
7.72
7.72
7.71
7.71
7.71
7.70
7.70
7.69
7.69
7.68
7.68
7.68
7.39
7.38
7.37
7.16
7.15
7.15
7.14
4.33
2.67
N
N
S
O
NH
O
NH
Cl
CH3
1.02
1.04
4.03
1.04
2.03
3.06
1.04
1.03
2.00
3.05
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
0.0
f1 (ppm)
21000
20000
19000
18000
17000
16000
15000
14000
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000

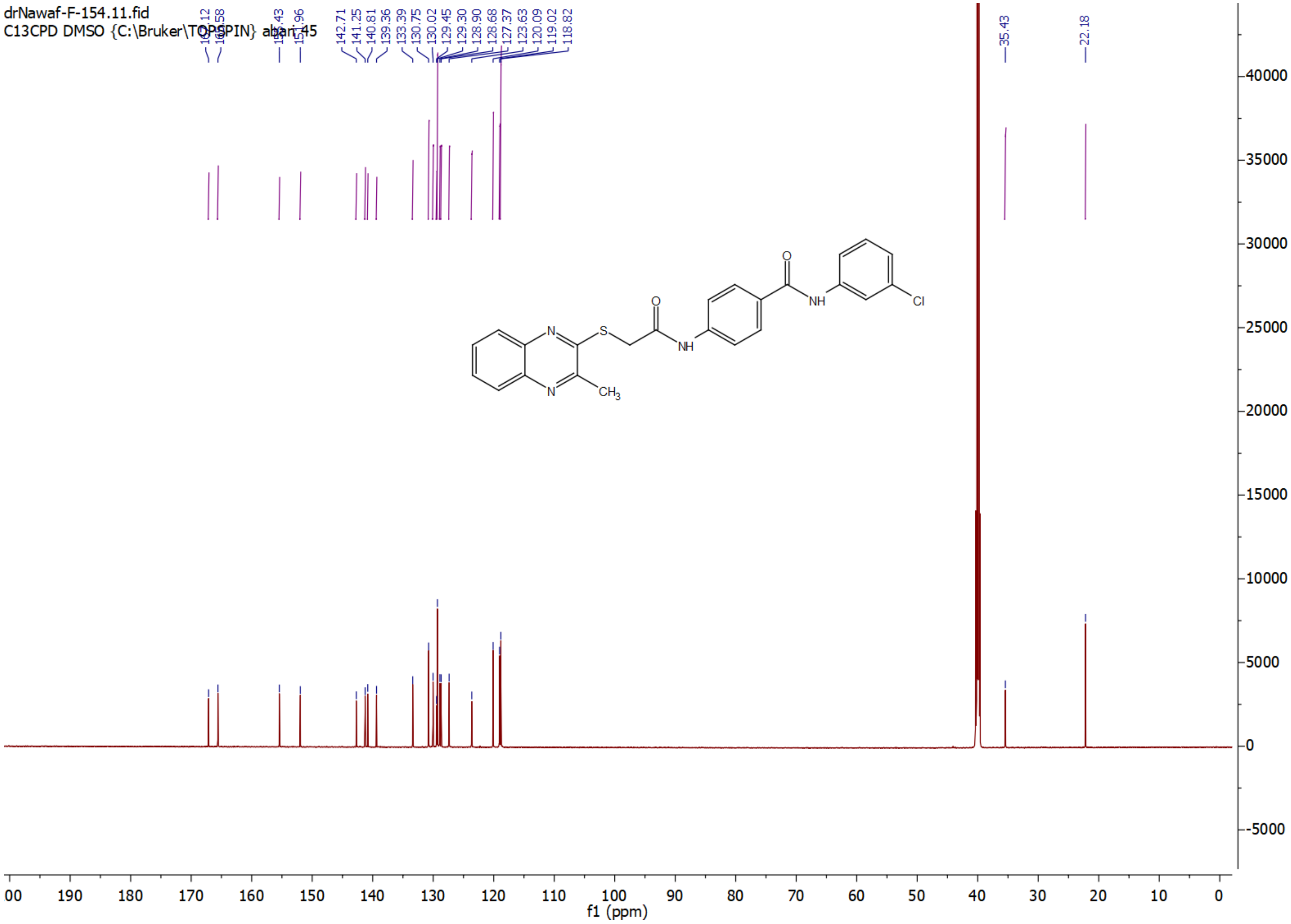
15b


drNawaf-F-154.10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 45
4.33
2.67
2.00
3.05
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000
.6
4.5
4.4
4.3
4.2
4.1
4.0
3.9
3.8
3.7
3.6
3.5
3.4
3.3
3.2
3.1
3.0
2.9
2.8
2.7
2.6
2.5
2.4
2.3
f1 (ppm)

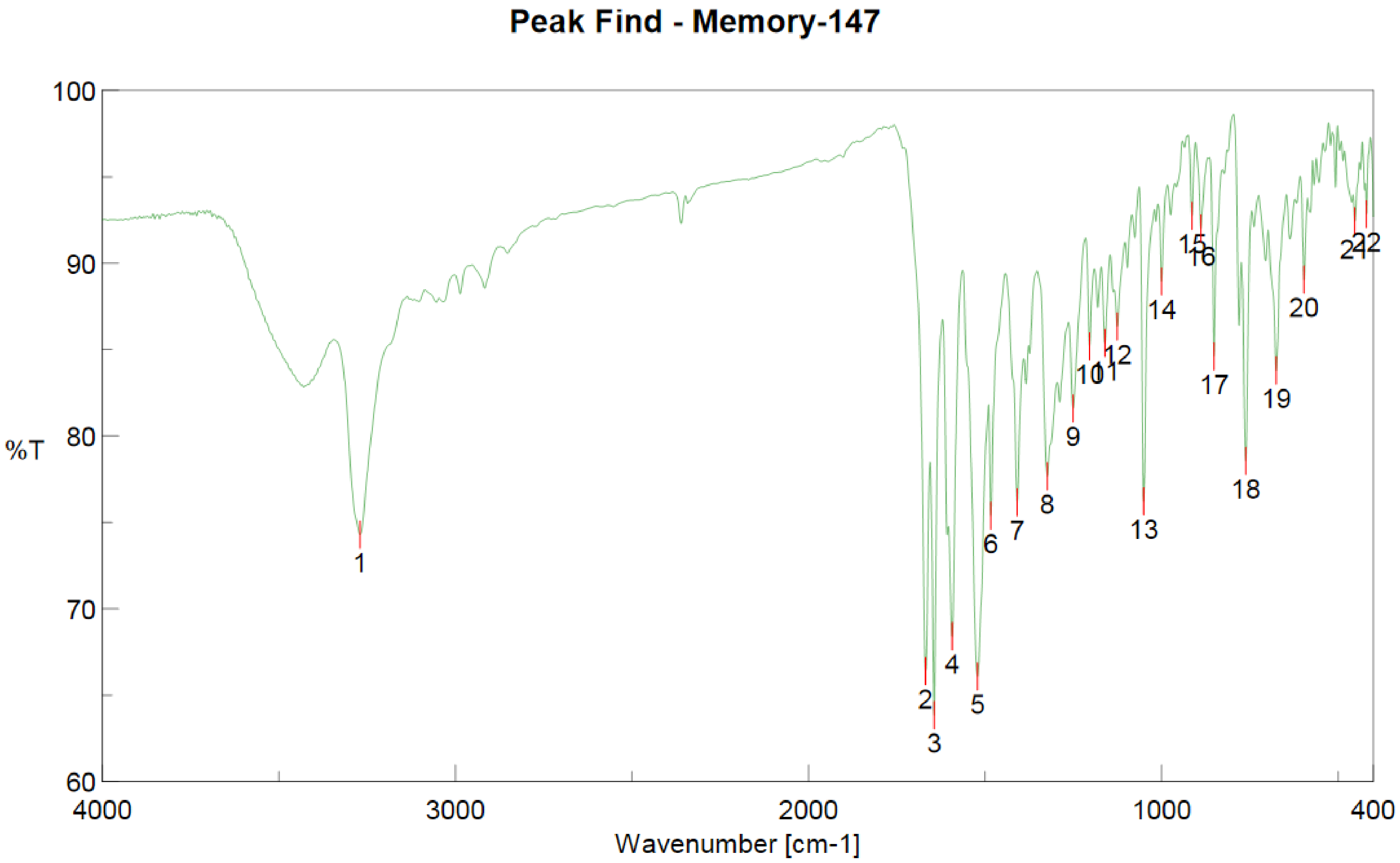
15b


drNawaf-F-154.10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 45
10.77
10.29
7.98
7.98
7.97
7.97
7.96
7.84
7.83
7.82
7.82
7.79
7.78
7.78
7.72
7.72
7.71
7.71
7.71
7.70
7.70
7.69
7.69
7.68
7.68
7.68
7.39
7.38
7.37
7.16
7.15
7.15
7.14
1.02
1.04
4.03
1.04
2.03
3.06
1.04
1.03
10.8
10.6
10.4
10.2
10.0
9.8
9.6
9.4
9.2
9.0
8.8
8.6
8.4
8.2
8.0
7.8
7.6
7.4
7.2
7.0
6.8
f1 (ppm)
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500

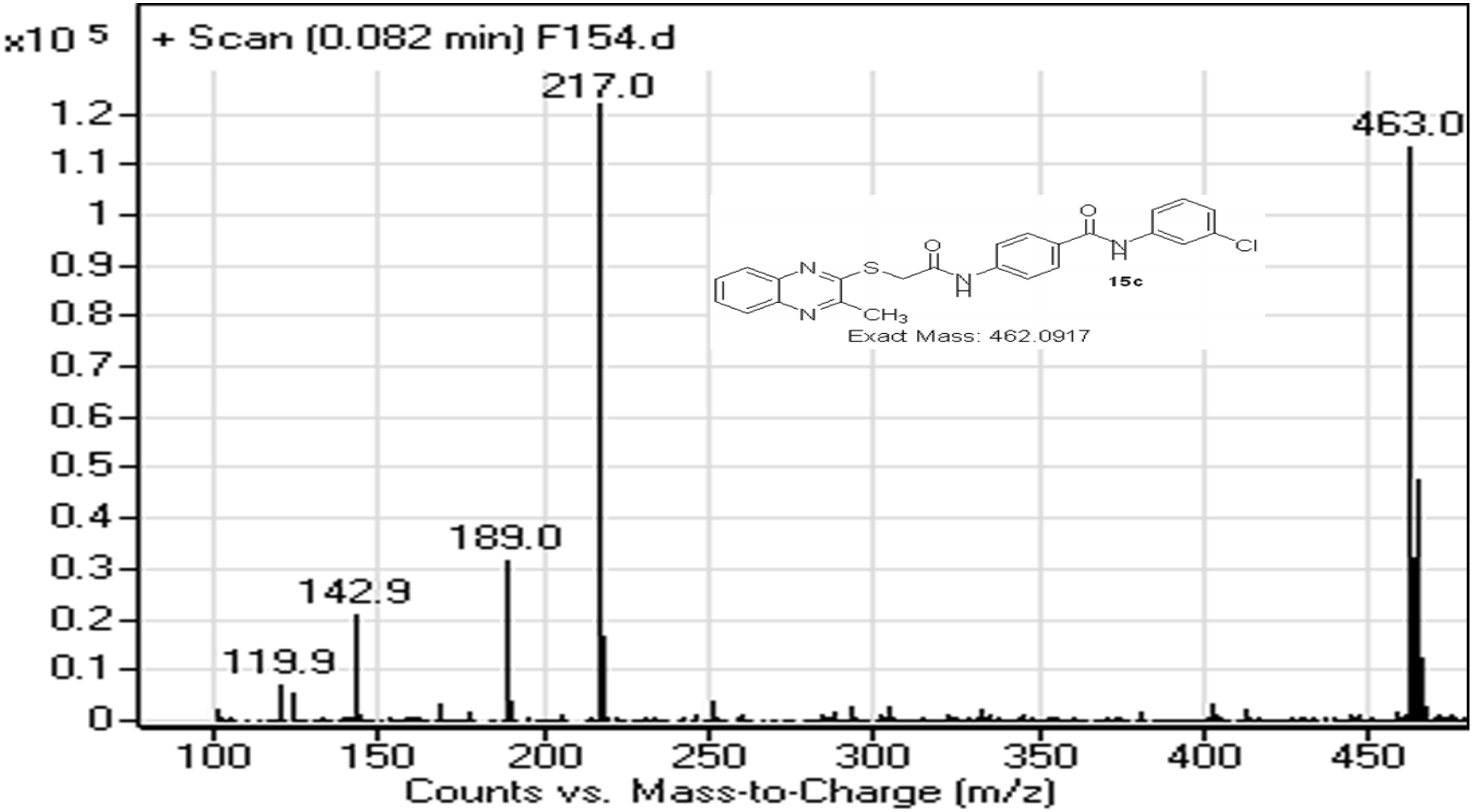
15b


drNawaf-F-154.11.fid
C13CPD DMSO {C:\Bruker\TOPSPIN} abari 45
167.12
165.58
155.43
151.96
142.71
141.25
140.81
139.36
133.39
130.75
130.02
129.45
129.30
128.90
128.68
127.37
123.63
120.09
119.02
118.82
35.43
22.18
N
N
S
O
NH
O
NH
Cl
CH3
40000
35000
30000
25000
20000
15000
10000
5000
0
-5000
00
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

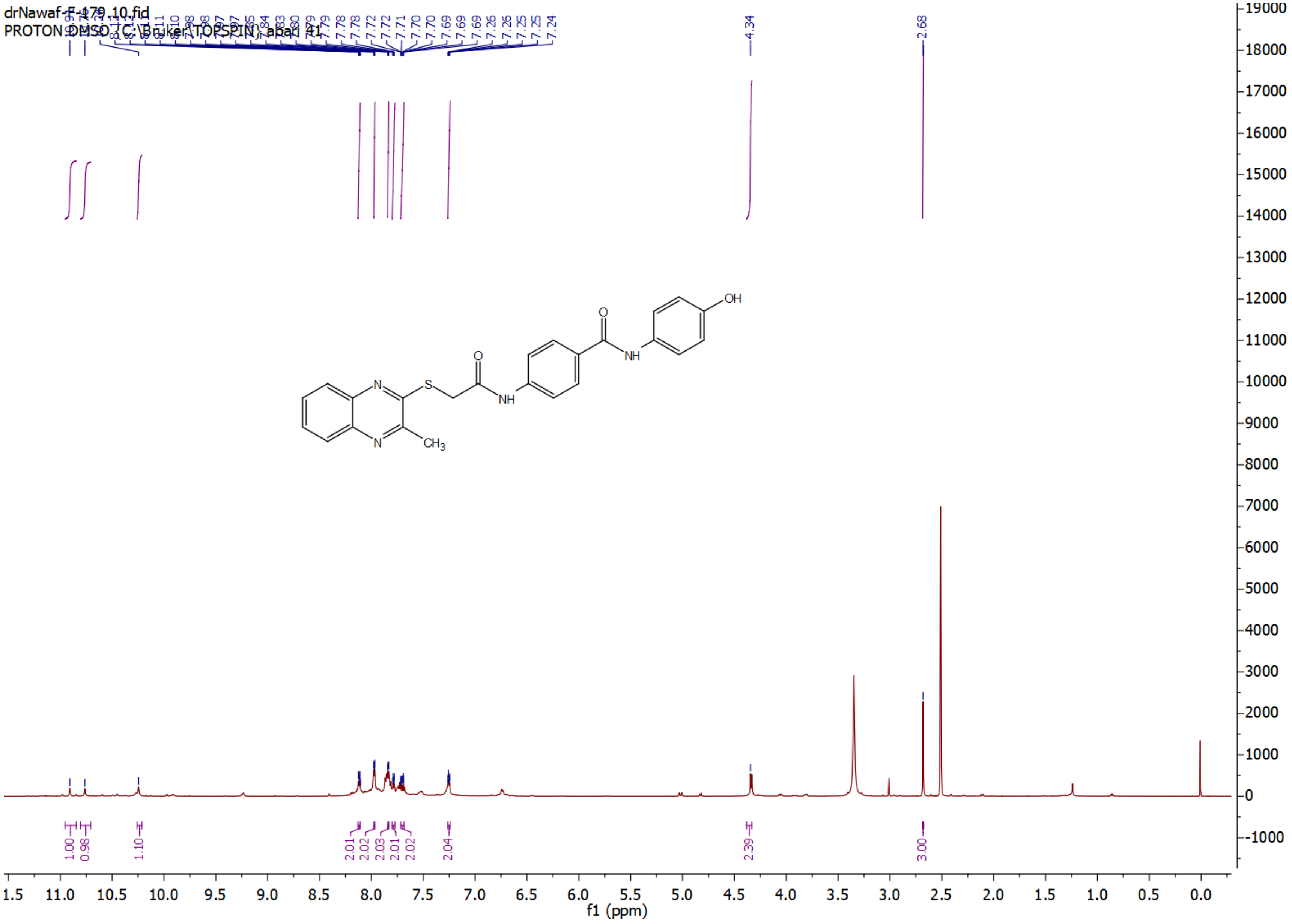
15b


Peak Find - Memory-147
100
90
80
70
60
%T
4000
3000
2000
1000
400
Wavenumber [cm-1]
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

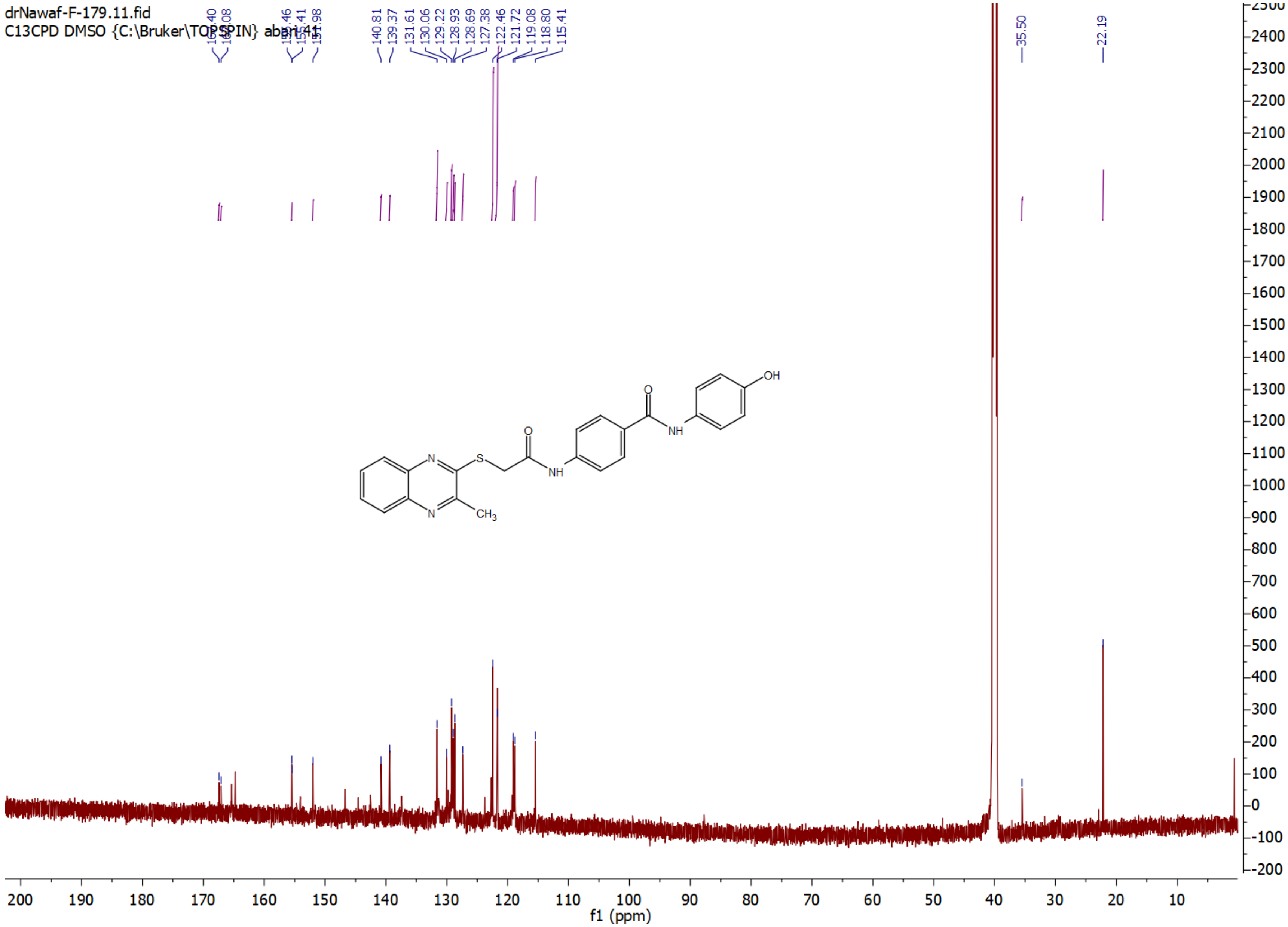
15b


x10 5
+ Scan (0.082 min) F154.d
217.0
463.0
189.0
142.9
119.9
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
100
150
200
250
300
350
400
450
Counts vs. Mass-to-Charge (m/z)
N
N
S
O
N
H
O
N
H
Cl
CH3
15c
Exact Mass: 462.0917

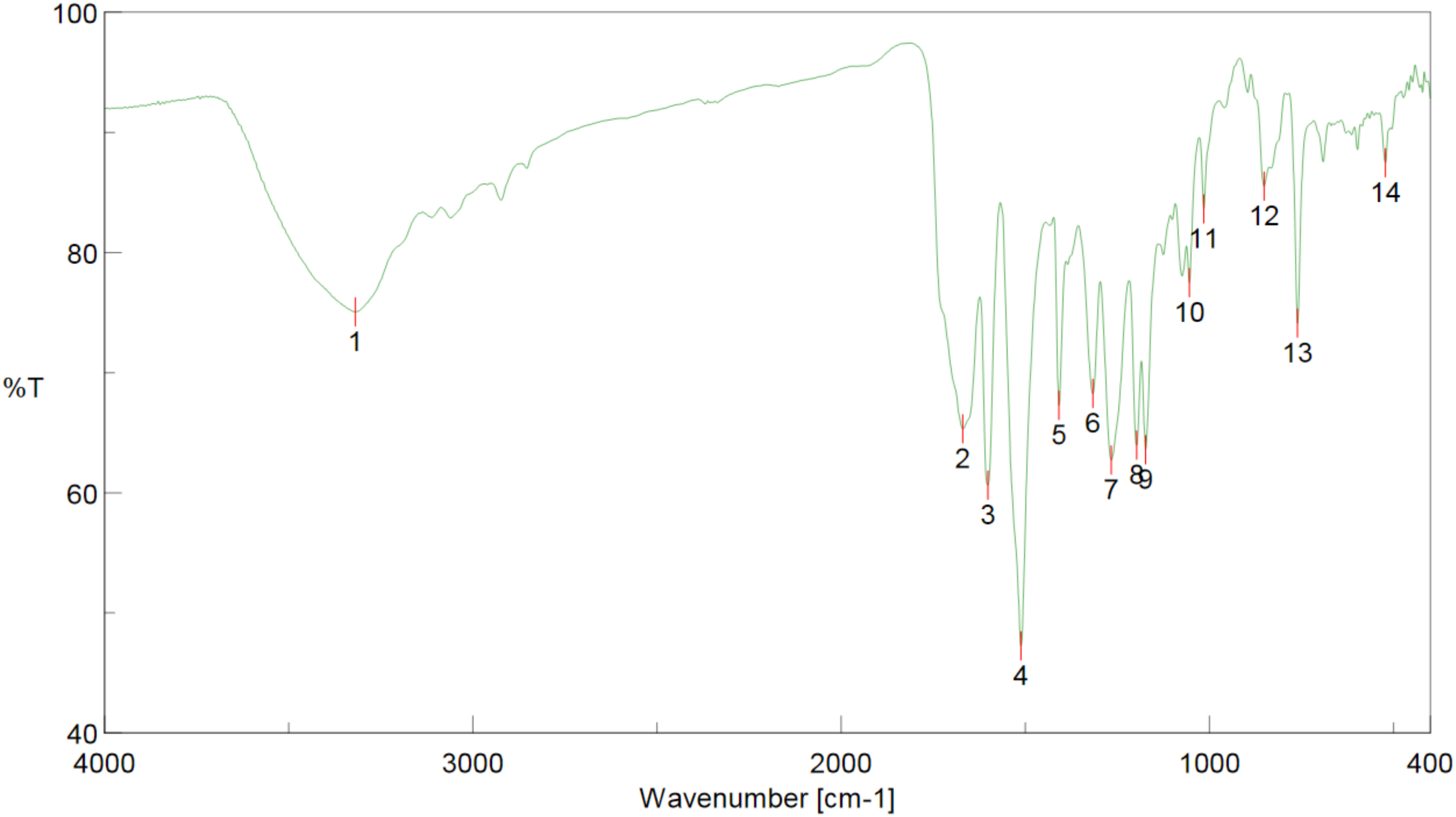
15c


drNawaf-F-179.10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 41
8.11
8.10
7.98
7.98
7.97
7.97
7.85
7.84
7.83
7.80
7.79
7.79
7.78
7.78
7.72
7.72
7.71
7.70
7.70
7.69
7.69
7.69
7.26
7.26
7.25
7.25
7.24
4.34
2.68
N
N
S
O
NH
O
NH
OH
CH3
1.00
0.98
1.10
2.01
2.02
2.03
2.01
2.02
2.04
2.39
3.00
19000
18000
17000
16000
15000
14000
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000
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
0.0
f1 (ppm)

15c

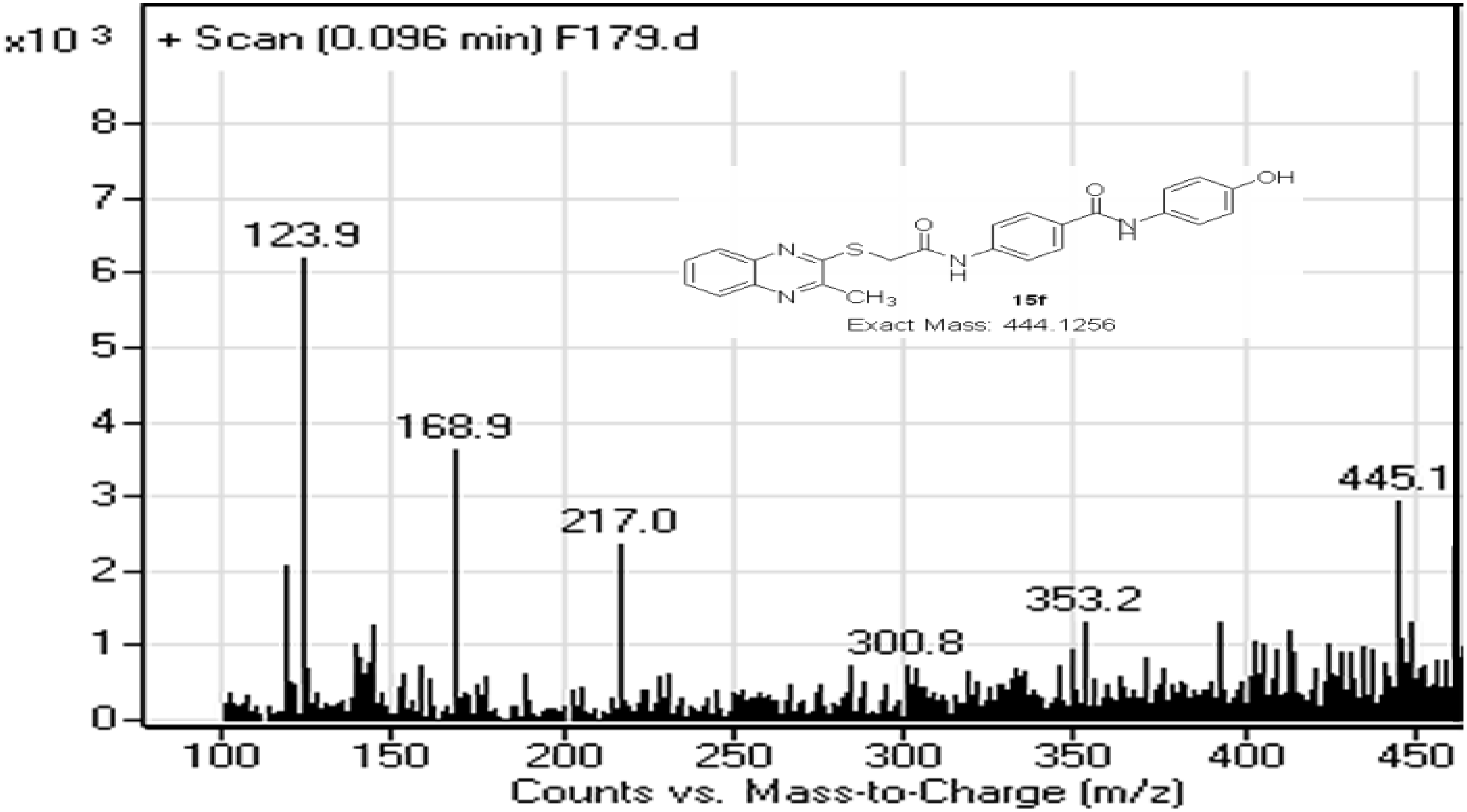

drNawaf-F-179.11.fid
C13CPD DMSO {C:\Bruker\TOPSPIN} abari 41
167.40
167.08
155.46
155.41
151.98
140.81
139.37
131.61
130.06
129.22
128.93
128.69
127.38
122.46
121.72
119.08
118.80
115.41
35.50
22.19
N
N
S
O
NH
O
NH
OH
CH3
2500
2400
2300
2200
2100
2000
1900
1800
1700
1600
1500
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
-100
-200
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
f1 (ppm)

15c



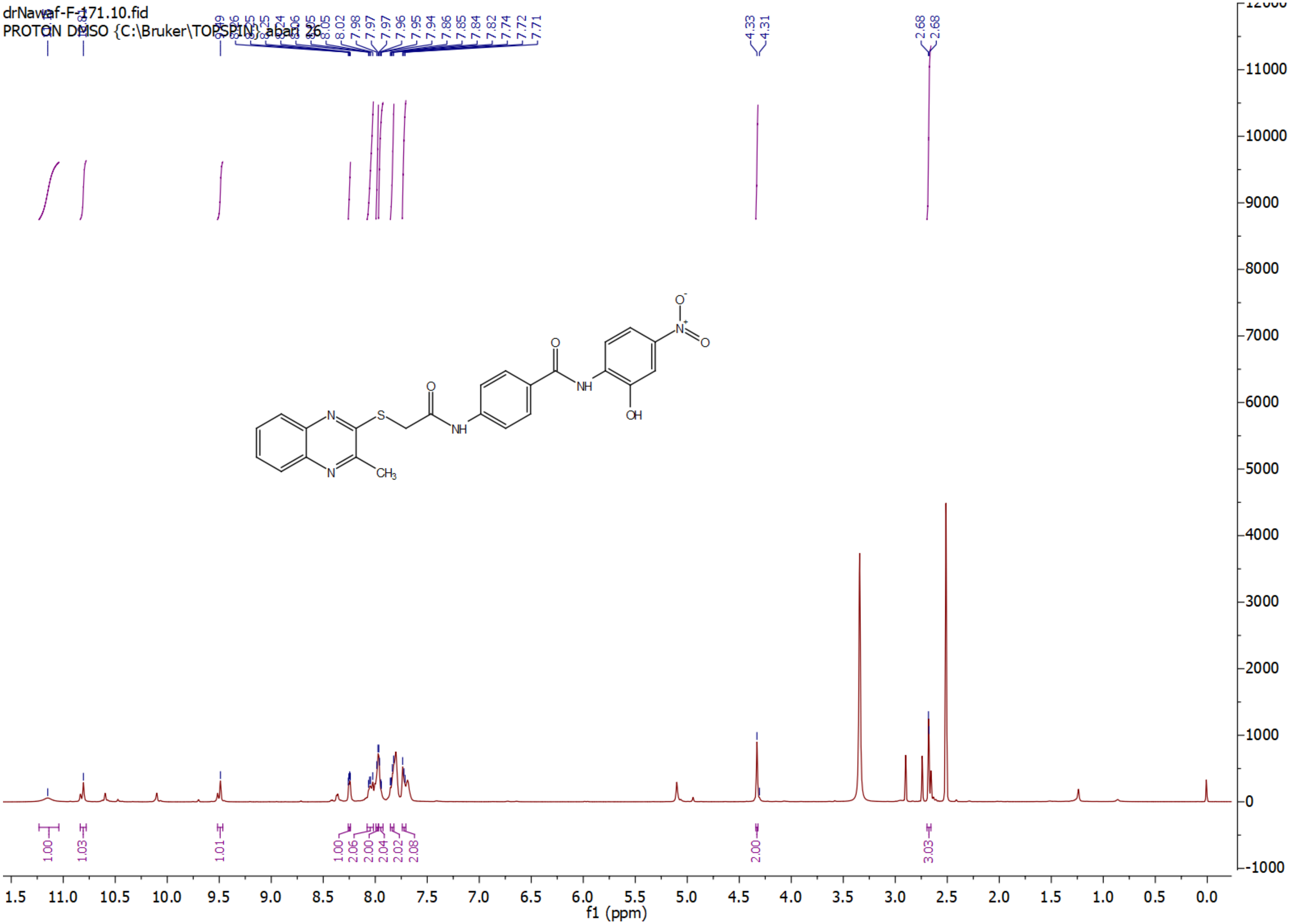
100
80
60
40
%T
4000
3000
2000
1000
400
Wavenumber [cm-1]
1
2
3
4
5
6
7
8
9
10
11
12
13
14

15c

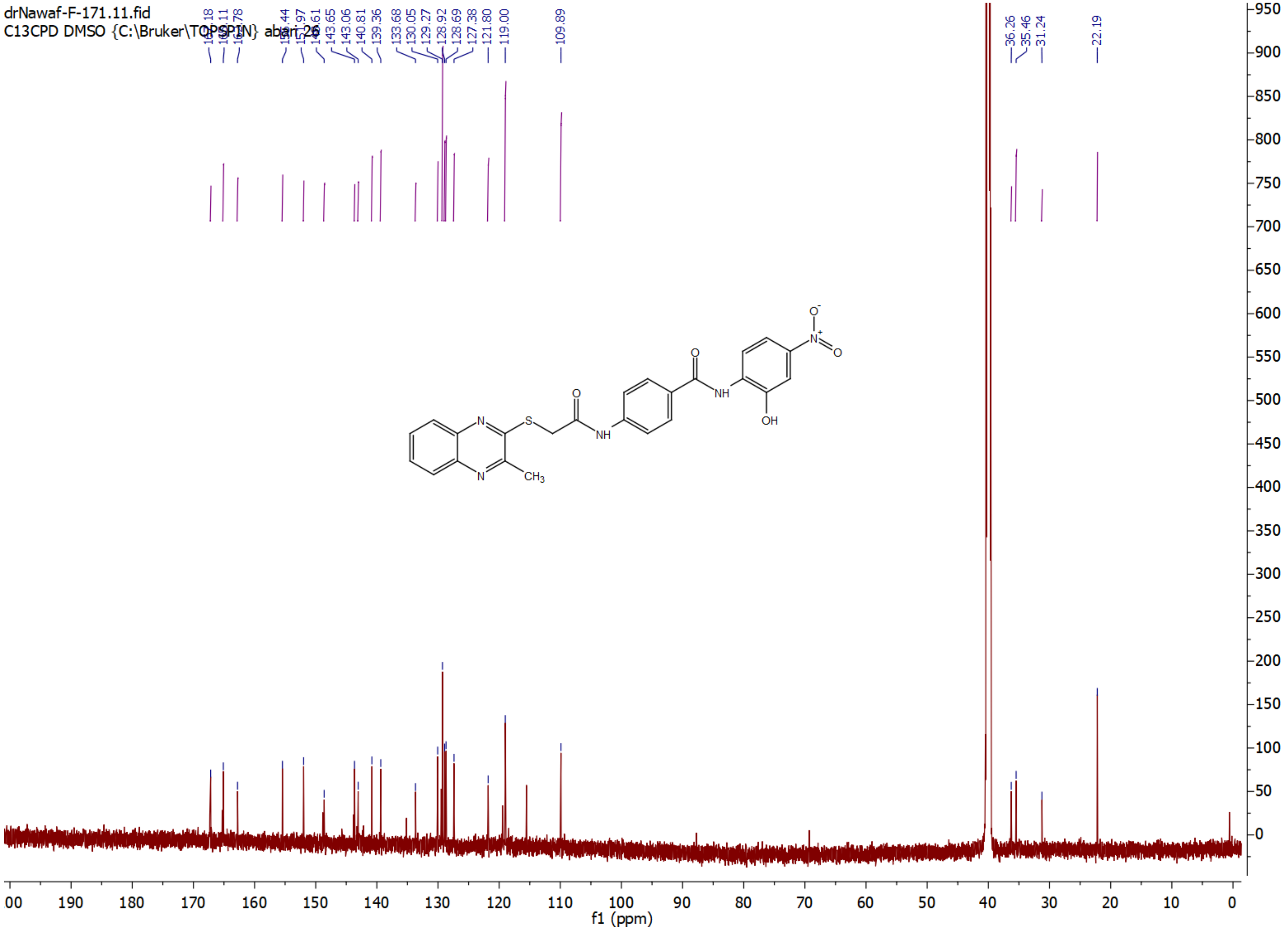


x10 3
+ Scan (0.096 min) F179.d
8
7
6
5
4
3
2
1
0
123.9
168.9
217.0
300.8
353.2
445.1
N
N
S
O
N
H
O
N
H
OH
CH3
15f
Exact Mass: 444.1256
100
150
200
250
300
350
400
450
Counts vs. Mass-to-Charge (m/z)

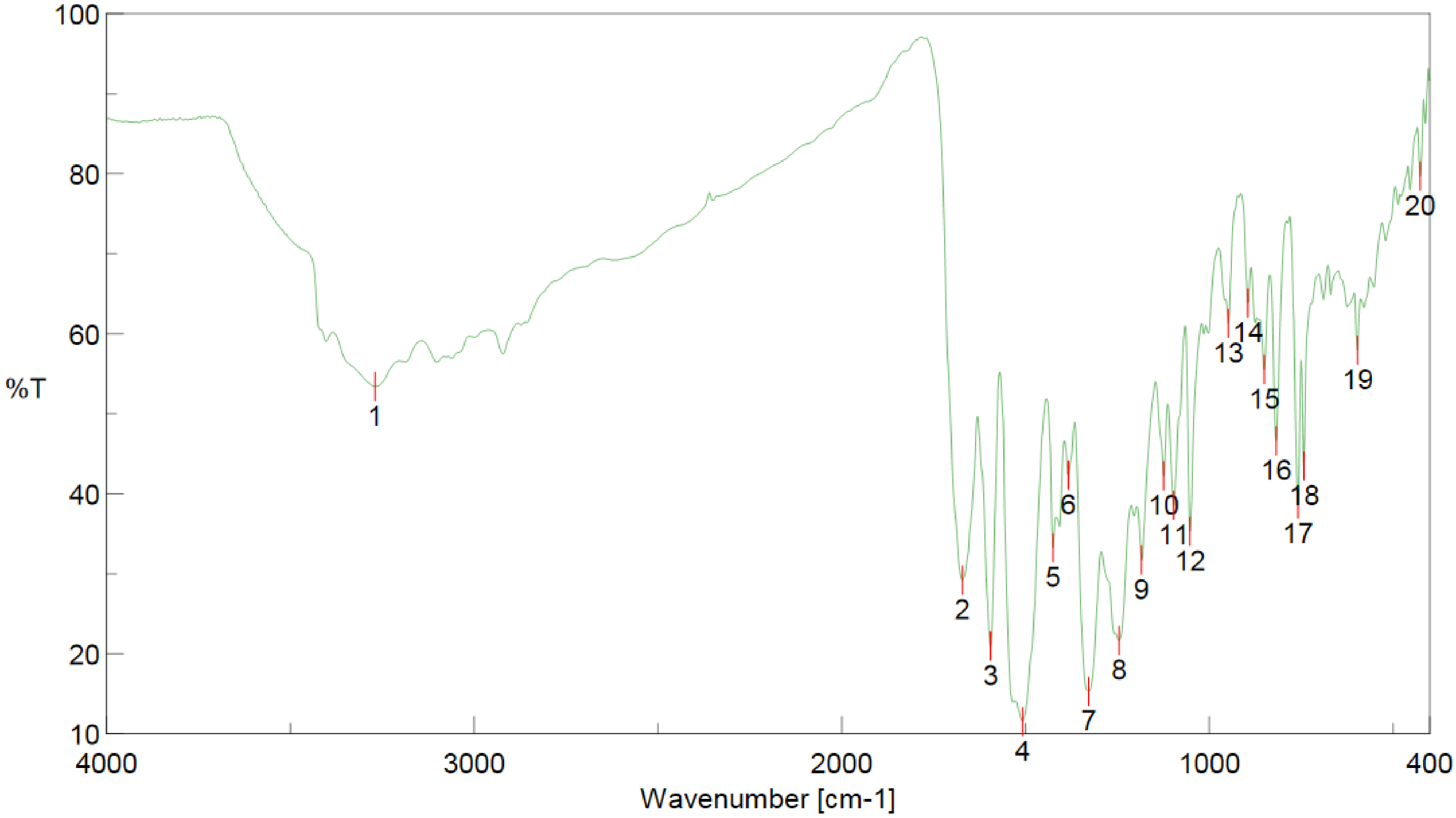
15d


drNawaf-F-171.10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 26
8.02
7.98
7.97
7.97
7.96
7.95
7.94
7.86
7.85
7.84
7.82
7.74
7.72
7.71
4.33
4.31
2.68
2.68
O
O-
N+
O
NH
O
OH
S
N
NH
N
CH3
1.00
1.03
1.01
1.00
2.06
2.00
2.04
2.02
2.08
2.00
3.03
f1 (ppm)
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
0.0
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000

15d

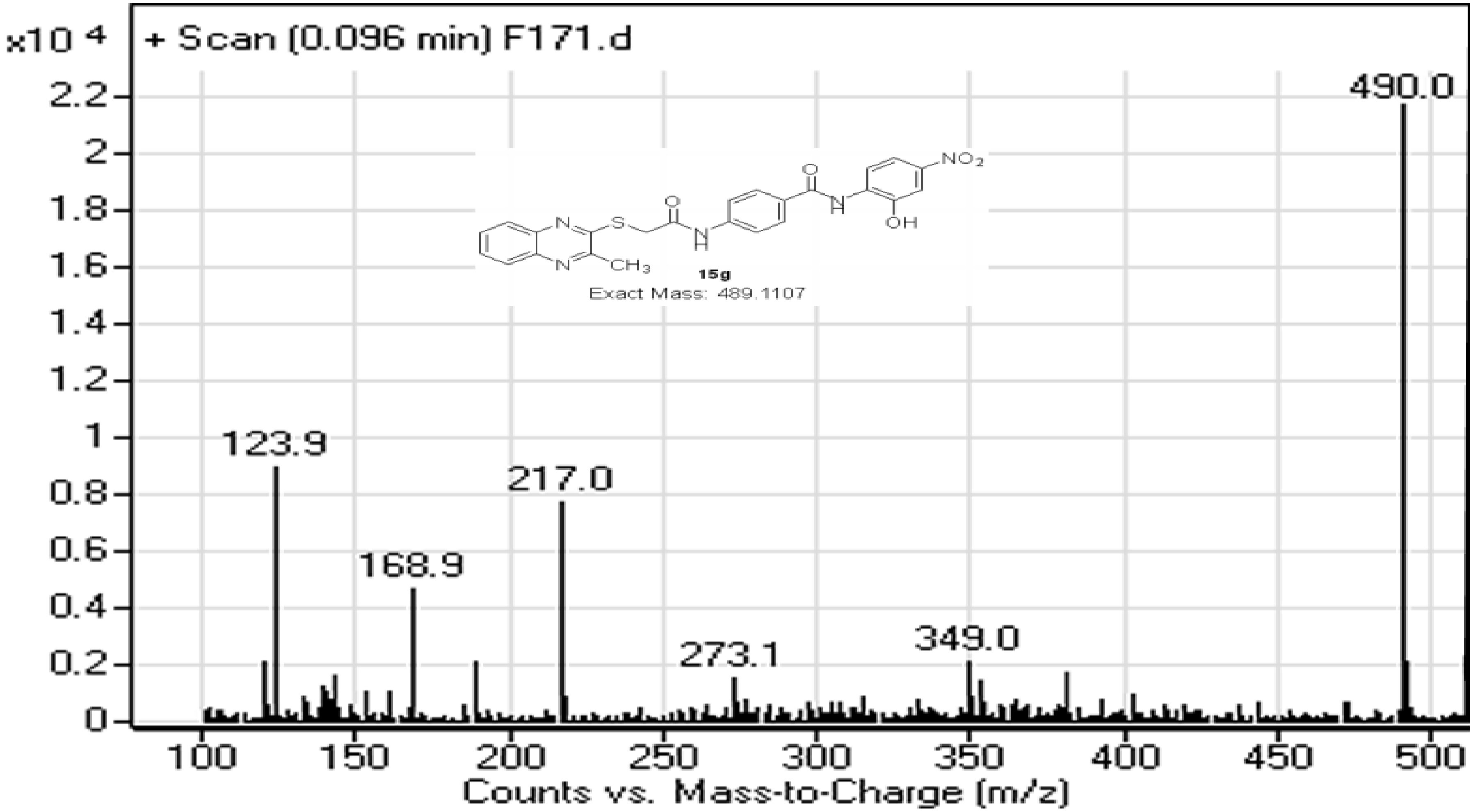

drNawaf-F-171.11.fid
C13CPD DMSO {C:\Bruker\TOPSPIN} abari 26
167.18
165.11
162.78
155.44
151.97
148.61
143.65
143.06
140.81
139.36
133.68
130.05
129.27
128.92
128.69
127.38
121.80
119.00
109.89
36.26
35.46
31.24
22.19
N
N
S
CH3
O
NH
O
NH
OH
O-
N+
O
950
900
850
800
750
700
650
600
550
500
450
400
350
300
250
200
150
100
50
0
00
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

15d

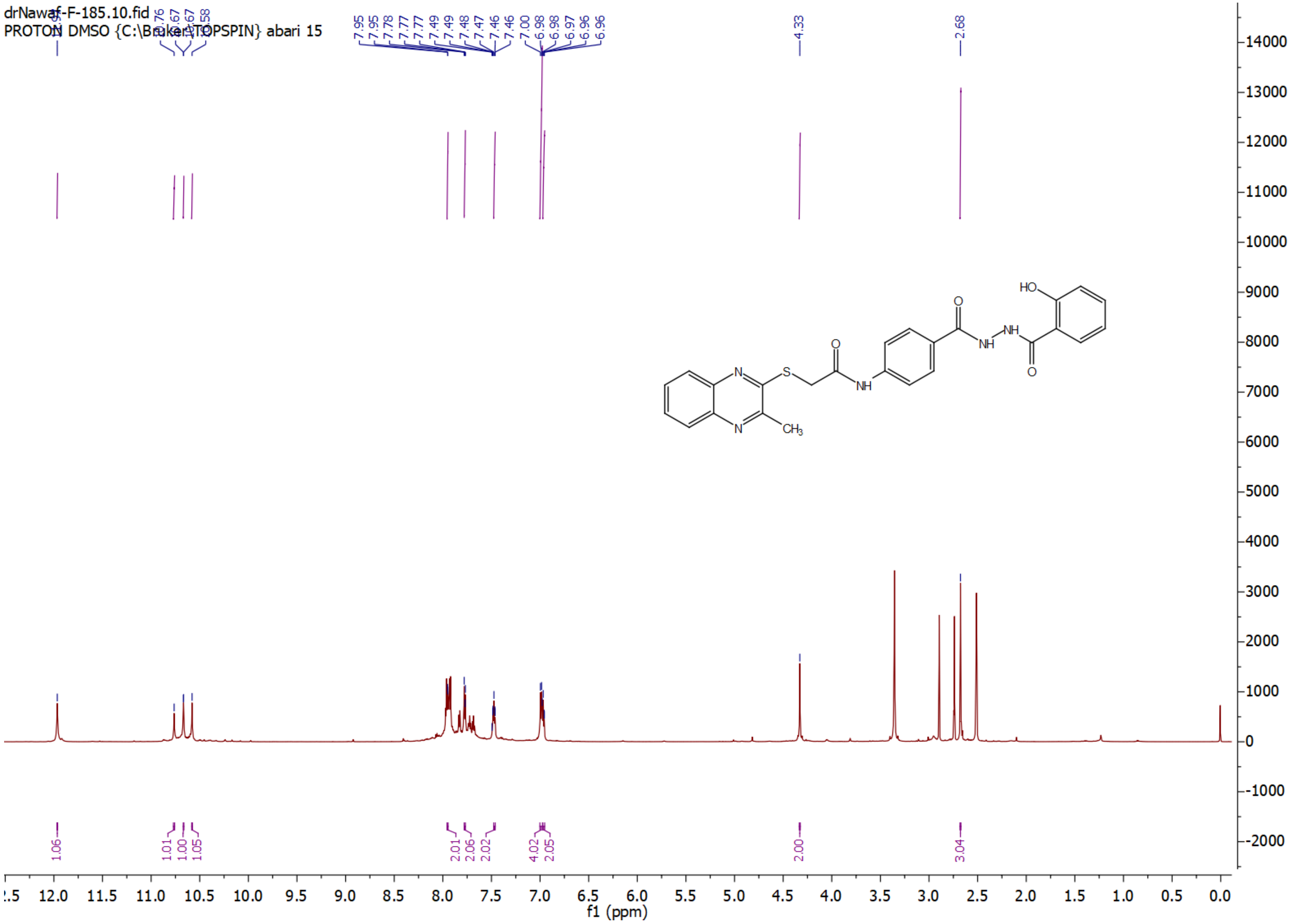


100
80
60
40
20
10
%T
4000
3000
2000
1000
400
Wavenumber [cm-1]
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

15d

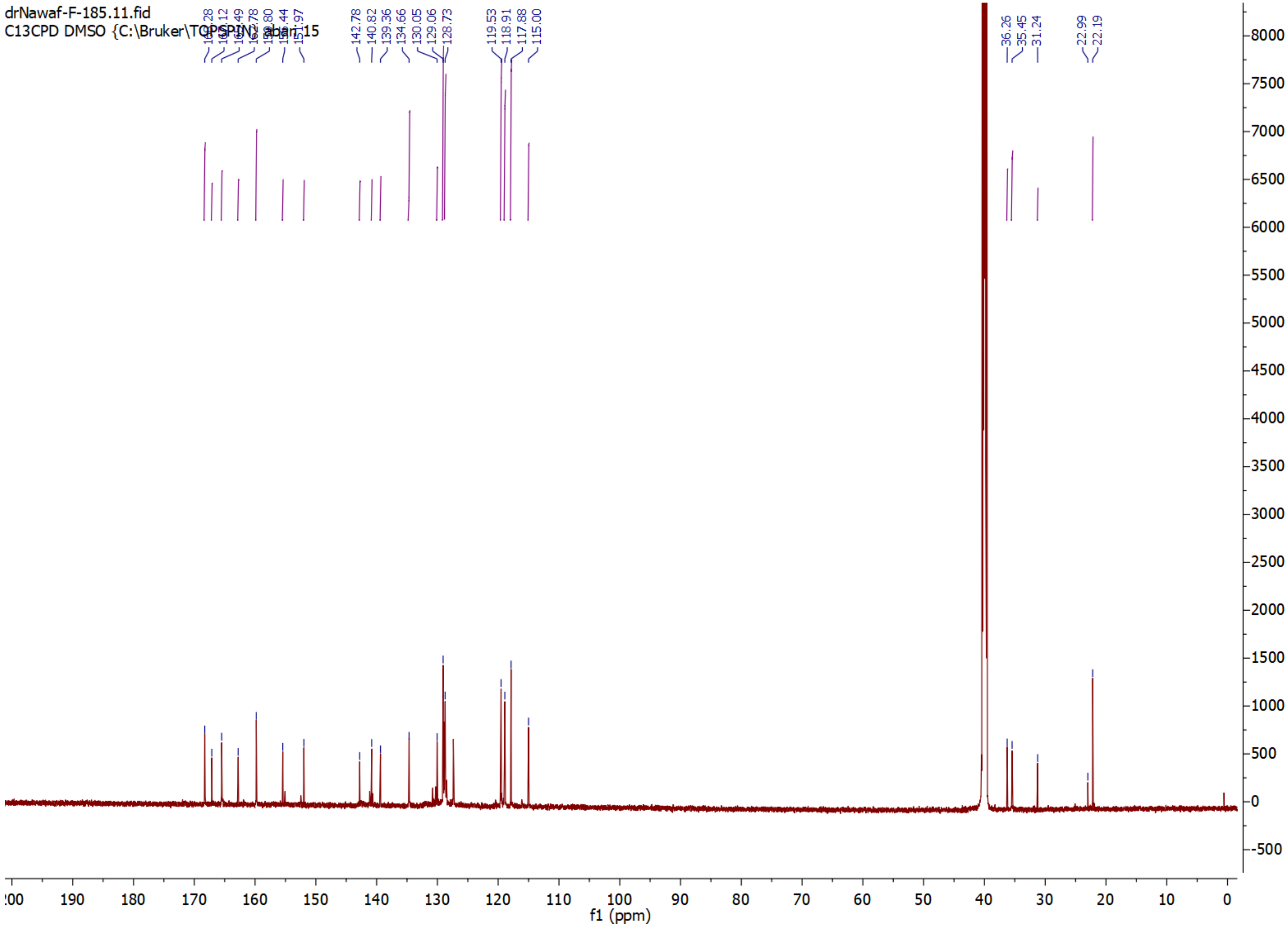

x10 4
+ Scan (0.096 min) F171.d
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
490.0
123.9
217.0
168.9
273.1
349.0
NO2
O
O
N
S
N
H
N
H
N
OH
CH3
15g
Exact Mass: 489.1107
100
150
200
250
300
350
400
450
500
Counts vs. Mass-to-Charge (m/z)

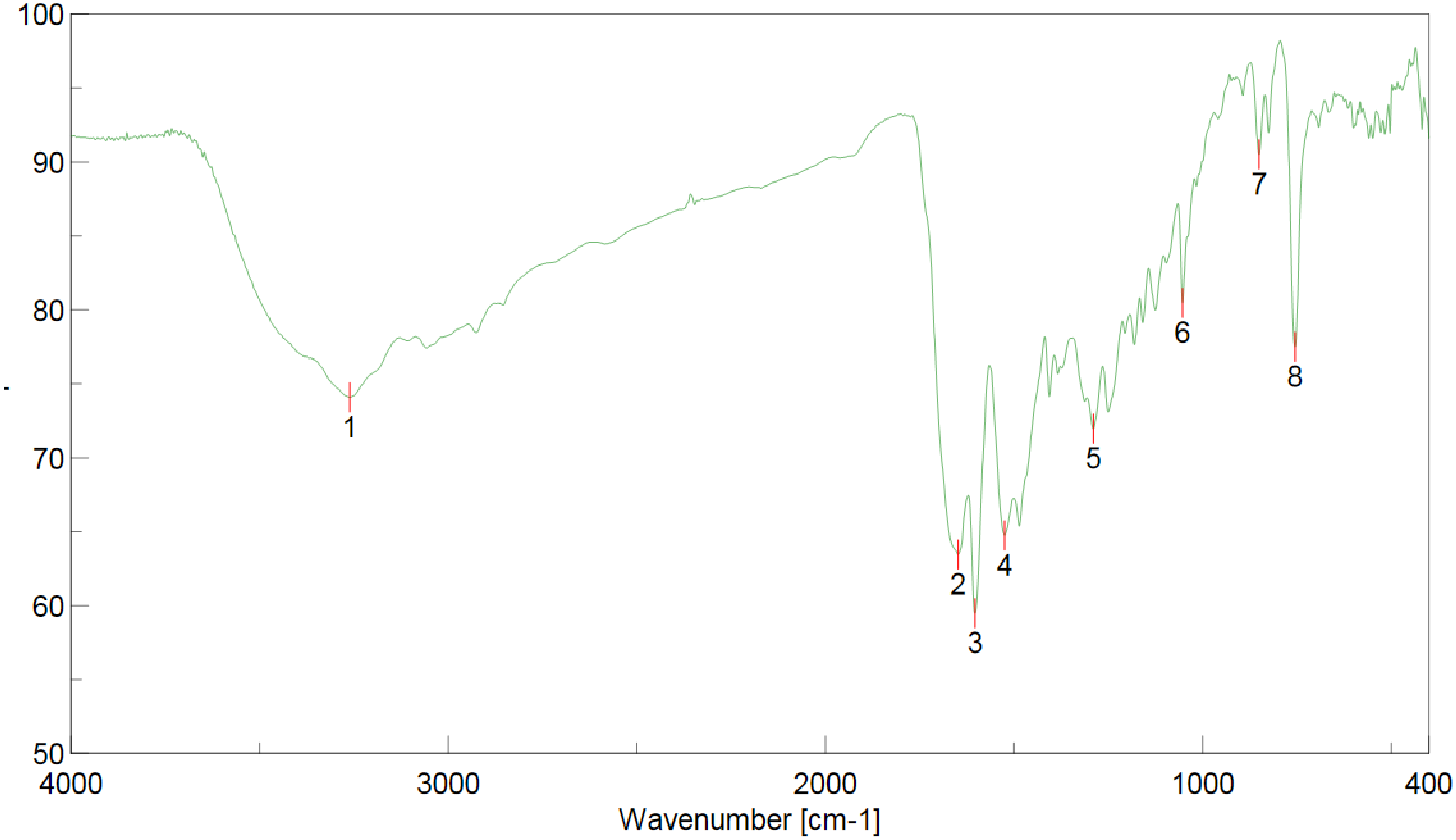
16



drNawaf-F-185.10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 15
10.76
10.67
10.67
10.58
7.95
7.95
7.78
7.77
7.77
7.49
7.49
7.48
7.47
7.46
7.46
7.00
6.98
6.98
6.97
6.96
6.96
4.33
2.68
1.06
1.01
1.00
1.05
2.01
2.06
2.02
4.02
2.05
2.00
3.04
HO
O
NH
NH
O
O
NH
S
N
N
CH3
14000
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000
-2000
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
0.0
f1 (ppm)

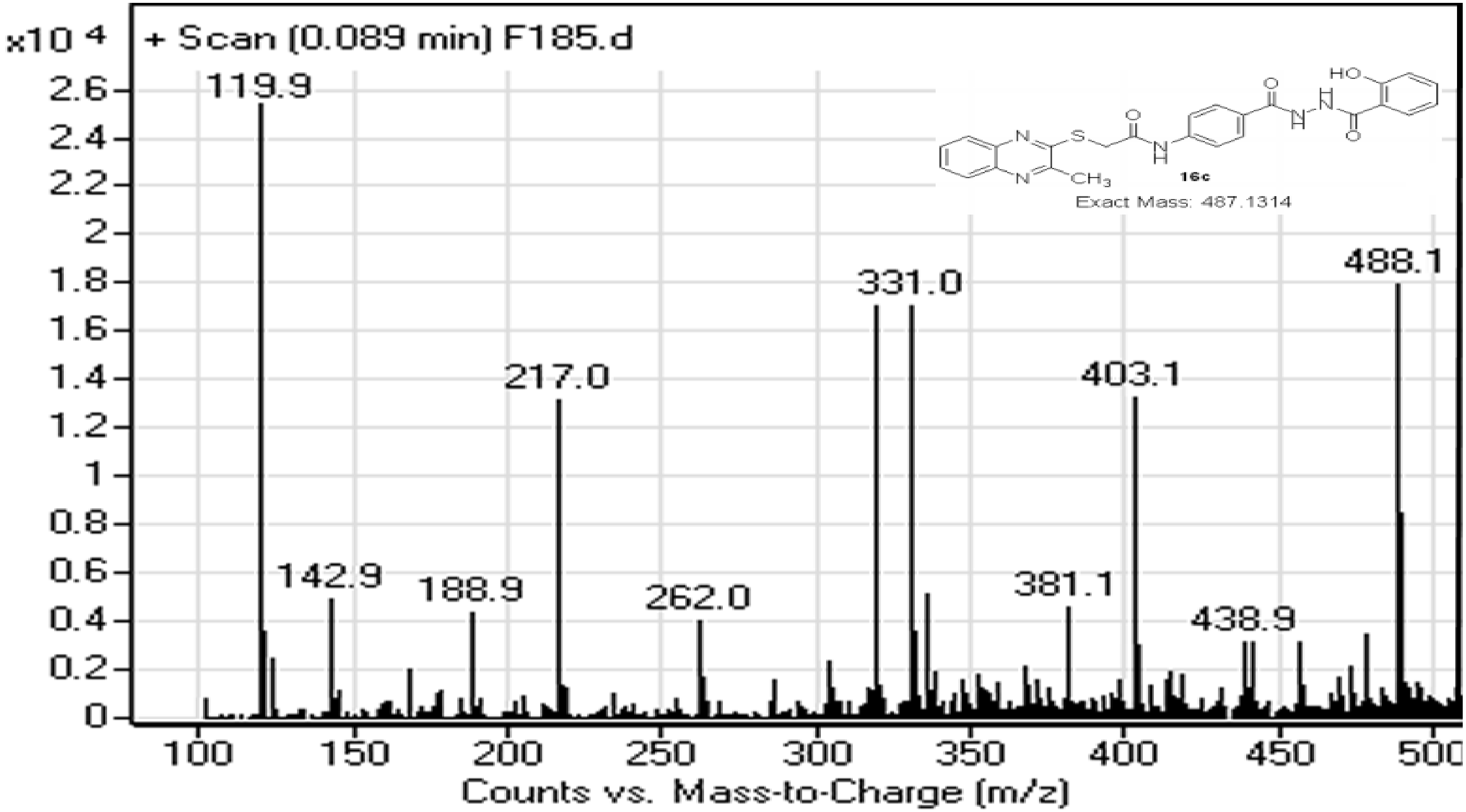
16


drNawaf-F-185.11.fid
C13CPD DMSO {C:\Bruker\TOPSPIN} abari 15
168.28
167.12
165.49
162.78
159.80
155.44
151.97
142.78
140.82
139.36
134.66
130.05
129.06
128.73
119.53
118.91
117.88
115.00
36.26
35.45
31.24
22.99
22.19
8000
7500
7000
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

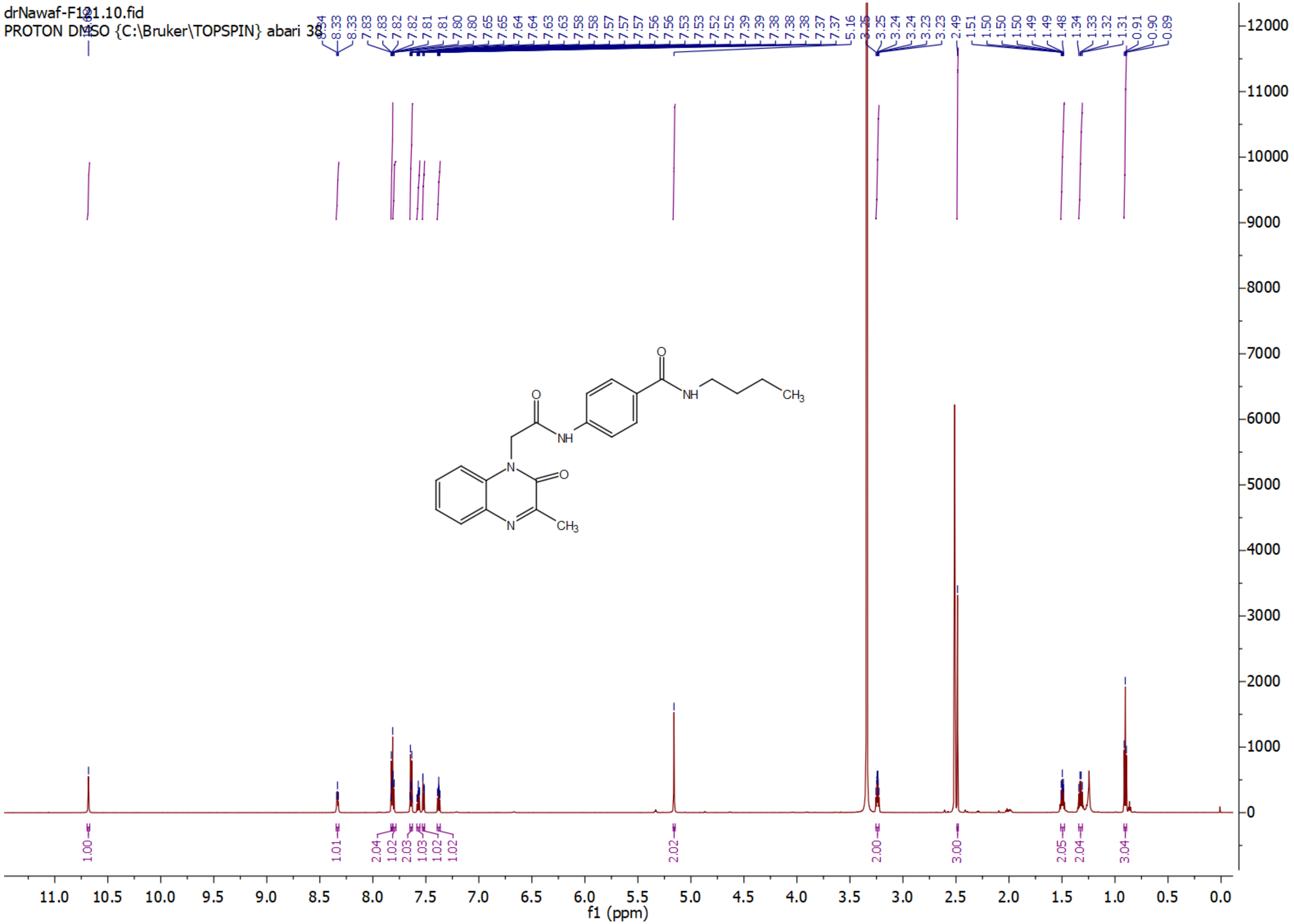


100
90
80
70
60
50
4000
3000
2000
1000
400
Wavenumber [cm-1]
1
2
3
4
5
6
7
8

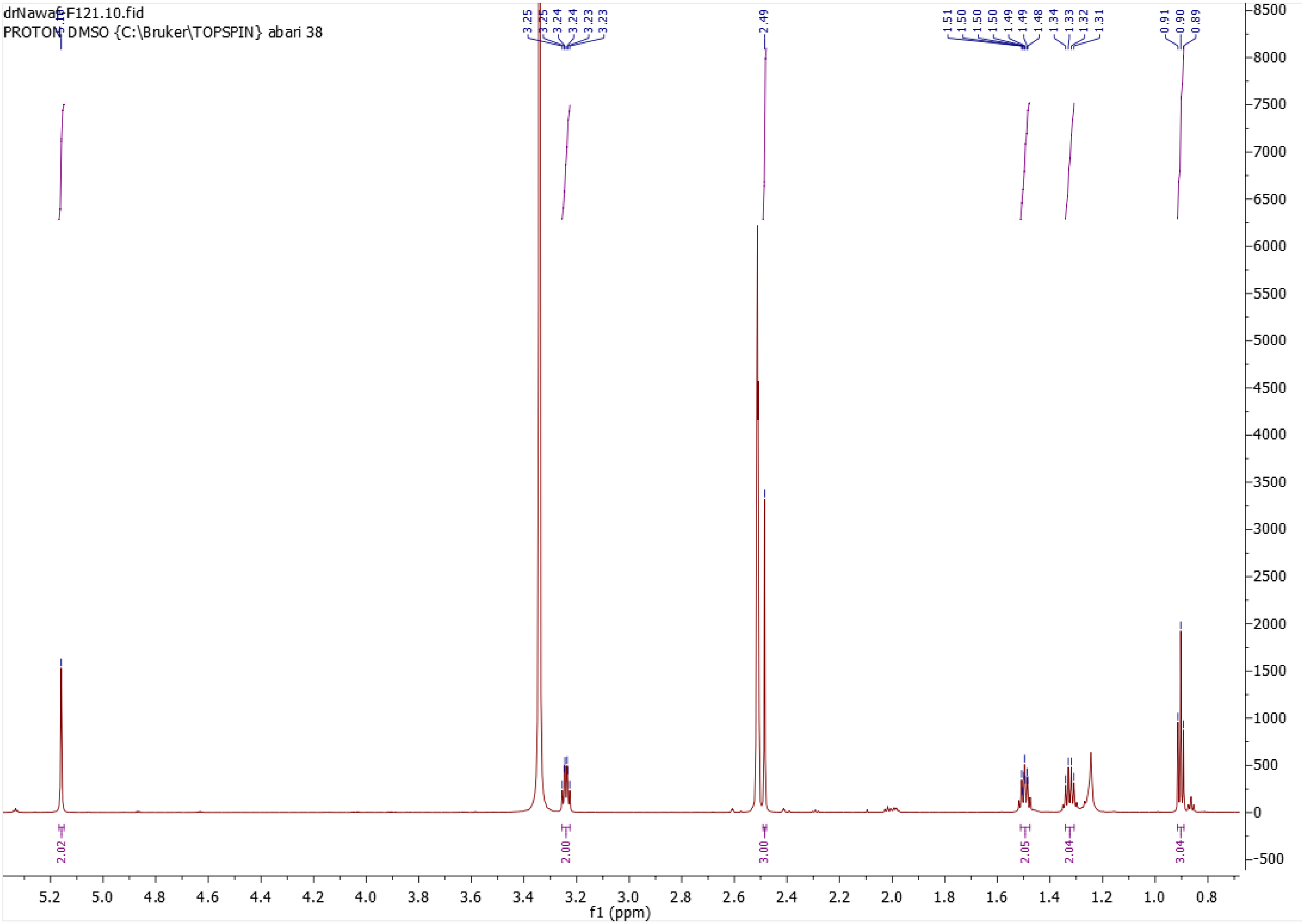
16


x10 4
+ Scan (0.089 min) F185.d
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
119.9
142.9
188.9
217.0
262.0
331.0
381.1
403.1
438.9
488.1
100
150
200
250
300
350
400
450
500
Counts vs. Mass-to-Charge (m/z)
N
N
S
CH3
O
N
H
O
H
N
N
H
O
HO
16c
Exact Mass: 487.1314

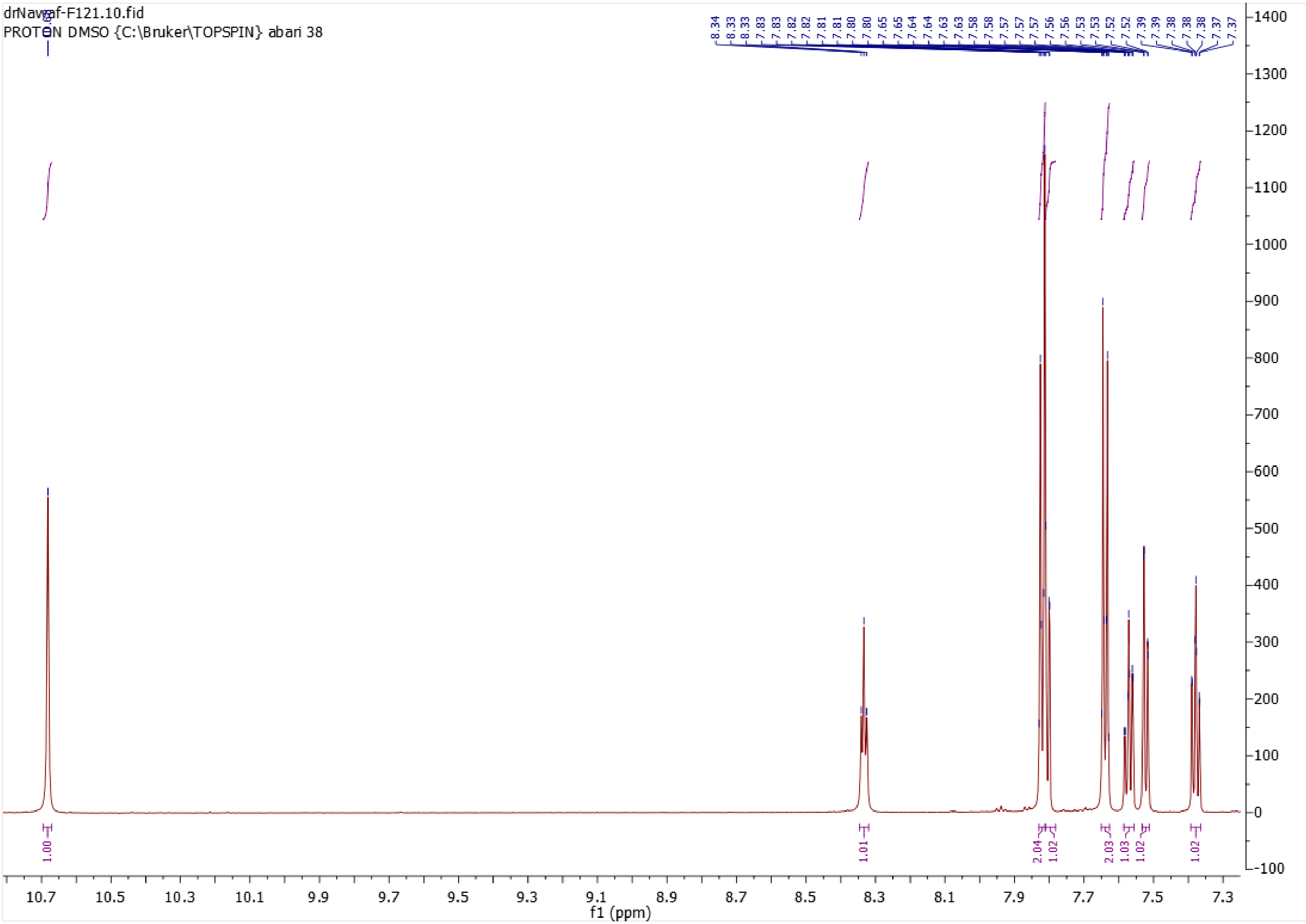
17a


drNawaf-F121.10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 38
10.68
8.34
8.33
8.33
7.83
7.82
7.82
7.81
7.80
7.80
7.65
7.65
7.64
7.64
7.63
7.63
7.58
7.58
7.57
7.57
7.57
7.56
7.56
7.53
7.53
7.52
7.52
7.39
7.39
7.38
7.38
7.38
7.37
7.37
5.16
3.35
3.25
3.24
3.23
3.23
2.49
1.51
1.50
1.50
1.50
1.49
1.49
1.48
1.34
1.33
1.32
1.31
0.91
0.90
0.89
O
NH
CH3
O
NH
N
O
N
CH3
1.00
1.01
2.04
1.02
2.03
1.03
1.02
1.02
2.02
2.00
3.00
2.05
2.04
3.04
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
0.0
f1 (ppm)
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000

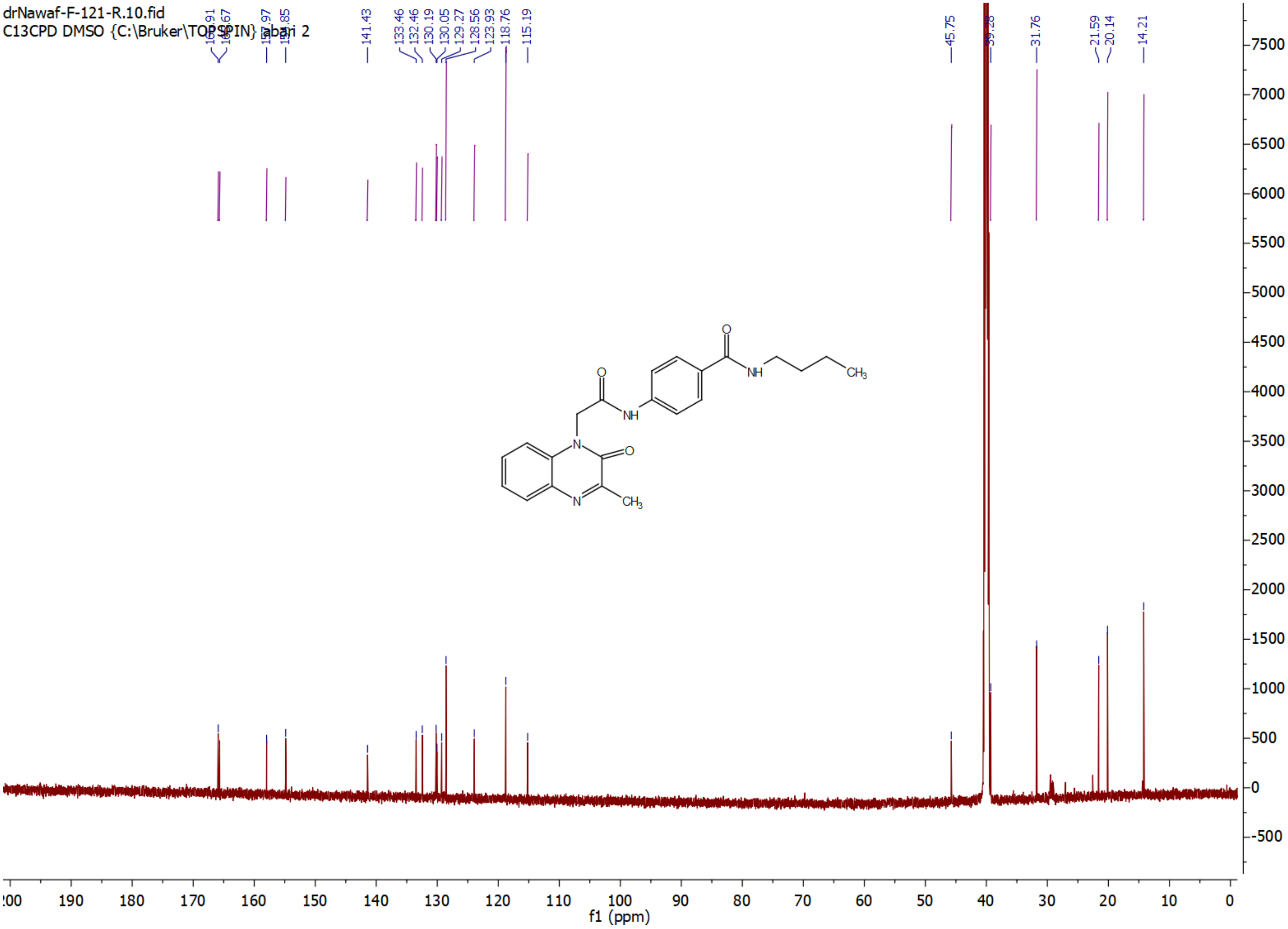
17a


drNawaf-F121.10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 38
5.17
3.25
3.25
3.24
3.24
3.23
3.23
2.49
1.51
1.50
1.50
1.50
1.49
1.49
1.48
1.34
1.33
1.32
1.31
0.91
0.90
0.89
2.02
2.00
3.00
2.05
2.04
3.04
8500
8000
7500
7000
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500
5.2
5.0
4.8
4.6
4.4
4.2
4.0
3.8
3.6
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
f1 (ppm)

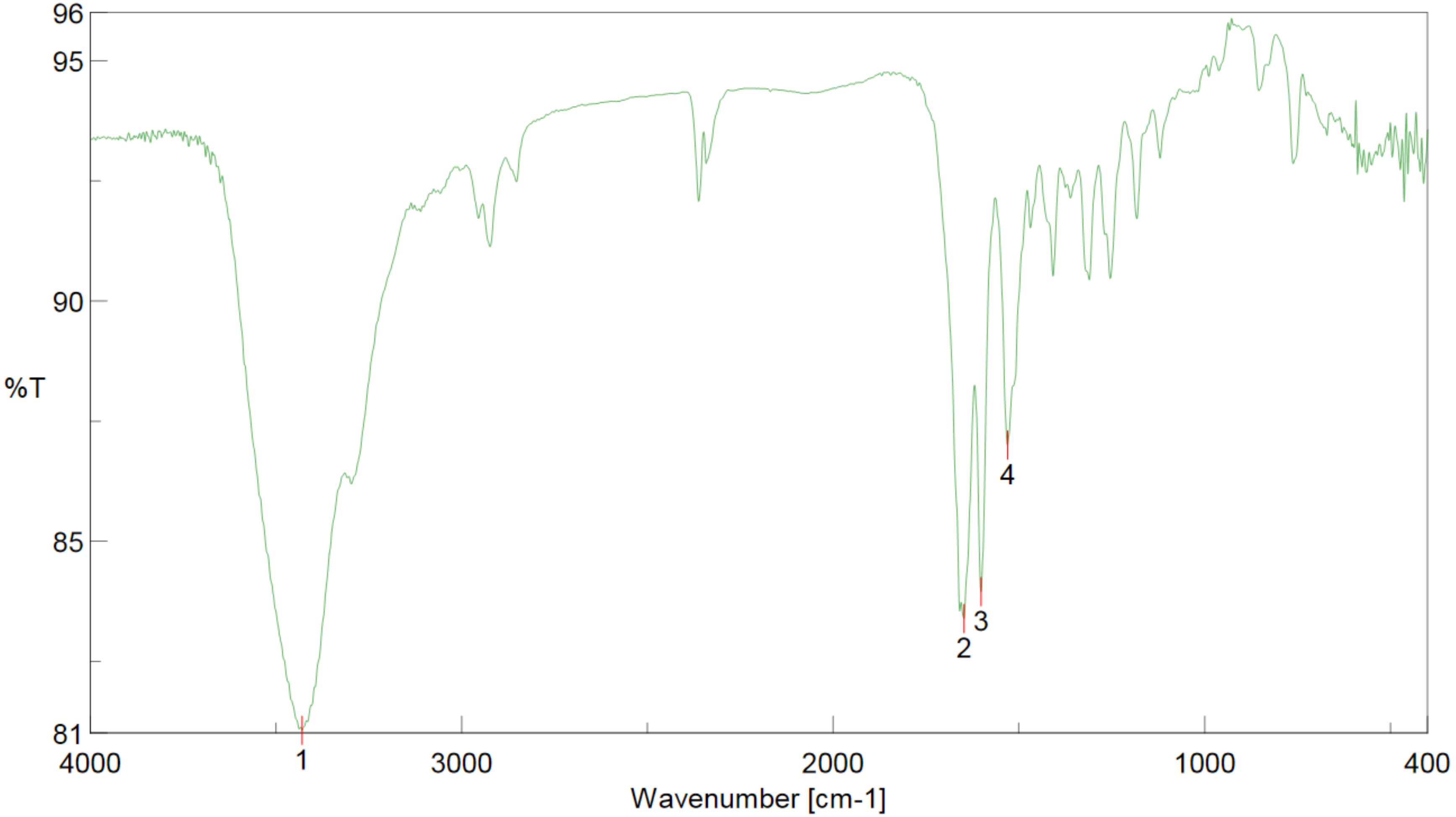
17a


drNawaf-F121.10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 38
8.34
8.33
8.33
7.83
7.83
7.82
7.82
7.81
7.81
7.80
7.80
7.65
7.65
7.64
7.64
7.63
7.63
7.58
7.58
7.57
7.57
7.57
7.56
7.56
7.53
7.53
7.52
7.52
7.39
7.39
7.38
7.38
7.38
7.37
7.37
1.00
1.01
2.04
1.02
2.03
1.03
1.02
1.02
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
-100
10.7
10.5
10.3
10.1
9.9
9.7
9.5
9.3
9.1
8.9
8.7
8.5
8.3
8.1
7.9
7.7
7.5
7.3
f1 (ppm)

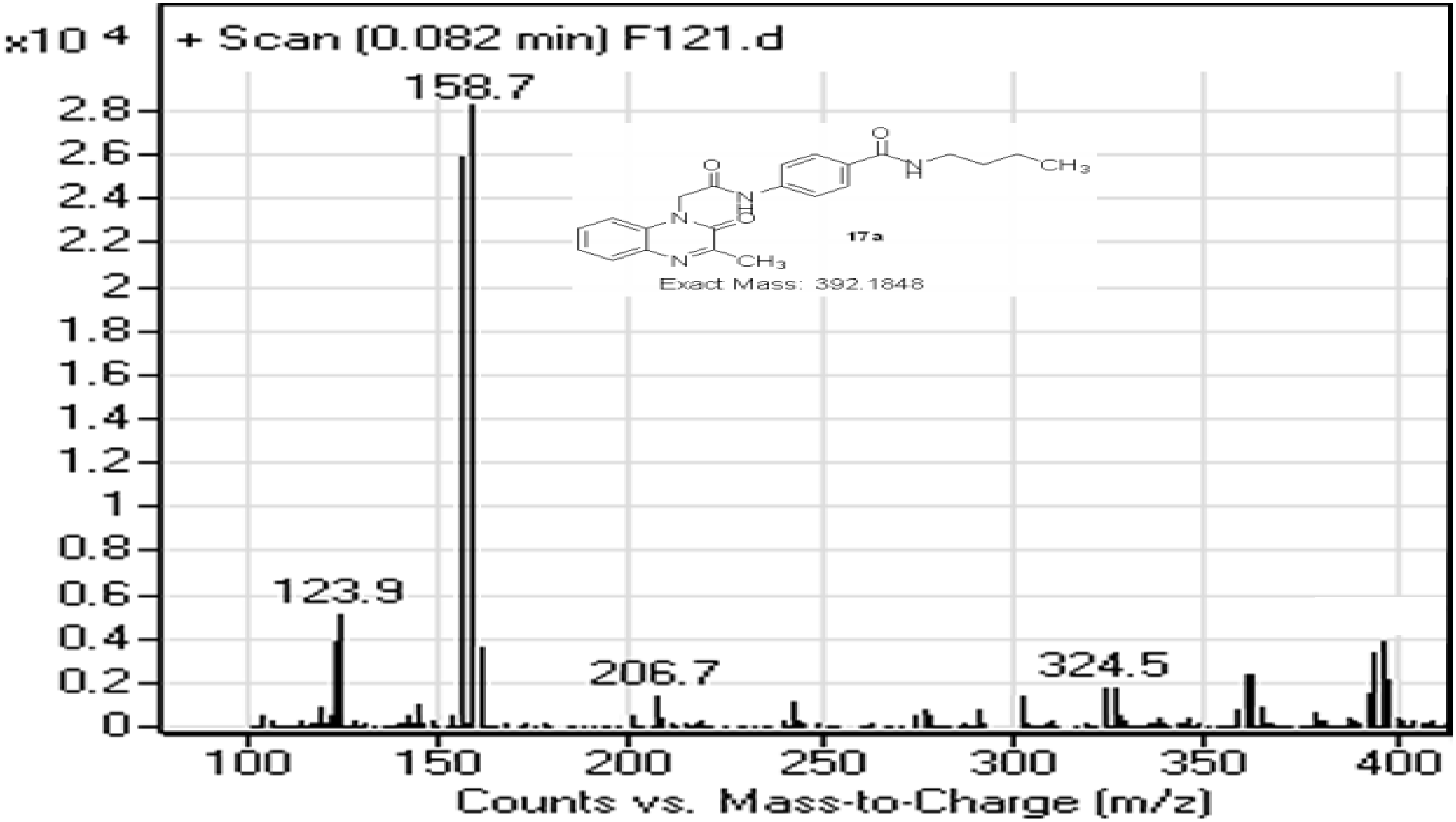
17a


drNawaf-F-121-R.10.fid
C13CPD DMSO {C:\Bruker\TOPSPIN} abari 2
165.91
165.67
157.97
154.85
141.43
133.46
132.46
130.19
130.05
129.27
128.56
123.93
118.76
115.19
45.75
39.28
31.76
21.59
20.14
14.21
O
NH
CH3
O
NH
N
O
N
CH3
7500
7000
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

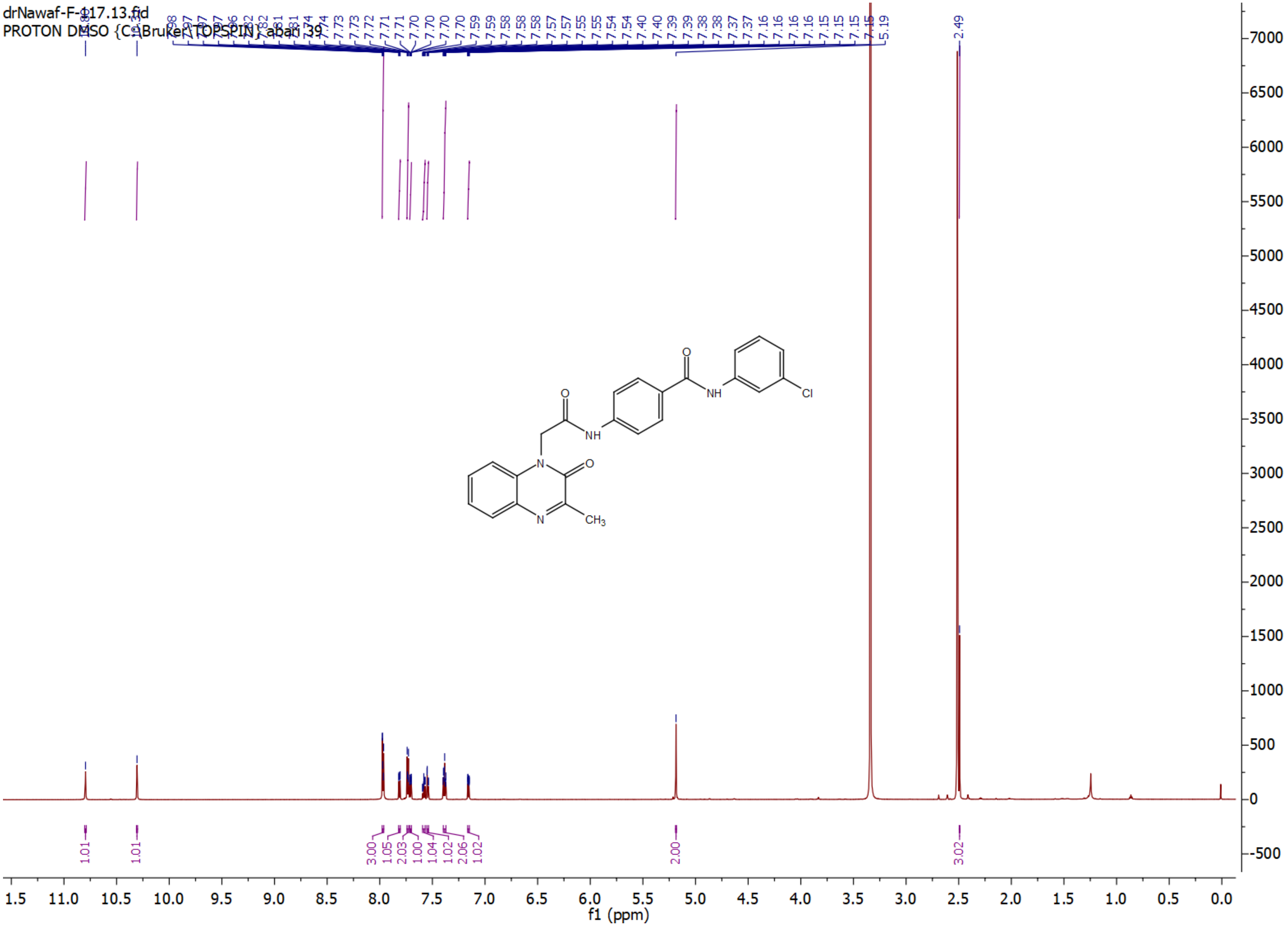
17a



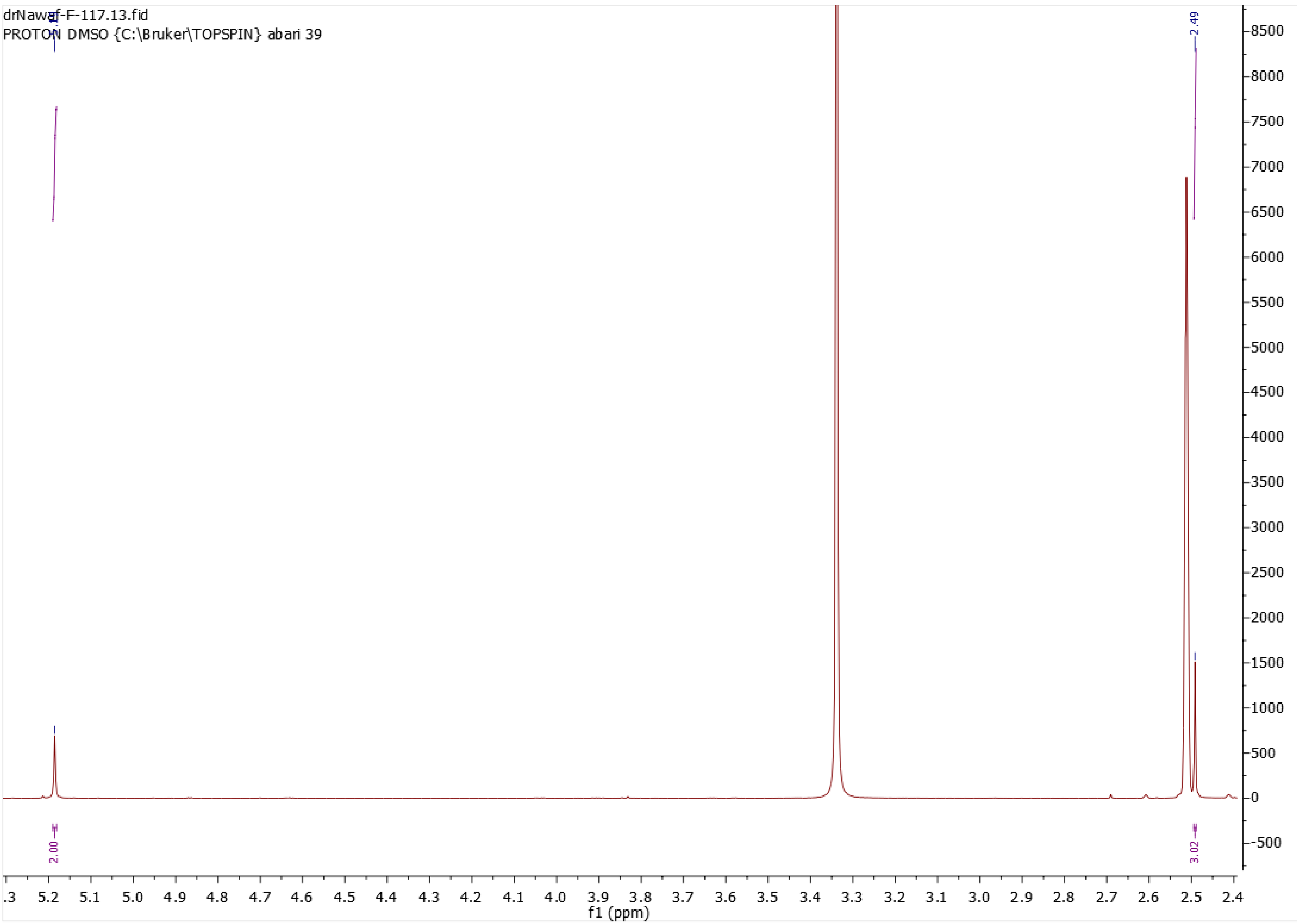
17a


x10 4
+ Scan (0.082 min) F121.d
158.7
123.9
206.7
324.5
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
100
150
200
250
300
350
400
Counts vs. Mass-to-Charge (m/z)
O
N
H
O
N
H
CH3
N
O
N
CH3
17a
Exact Mass: 392.1848

17b

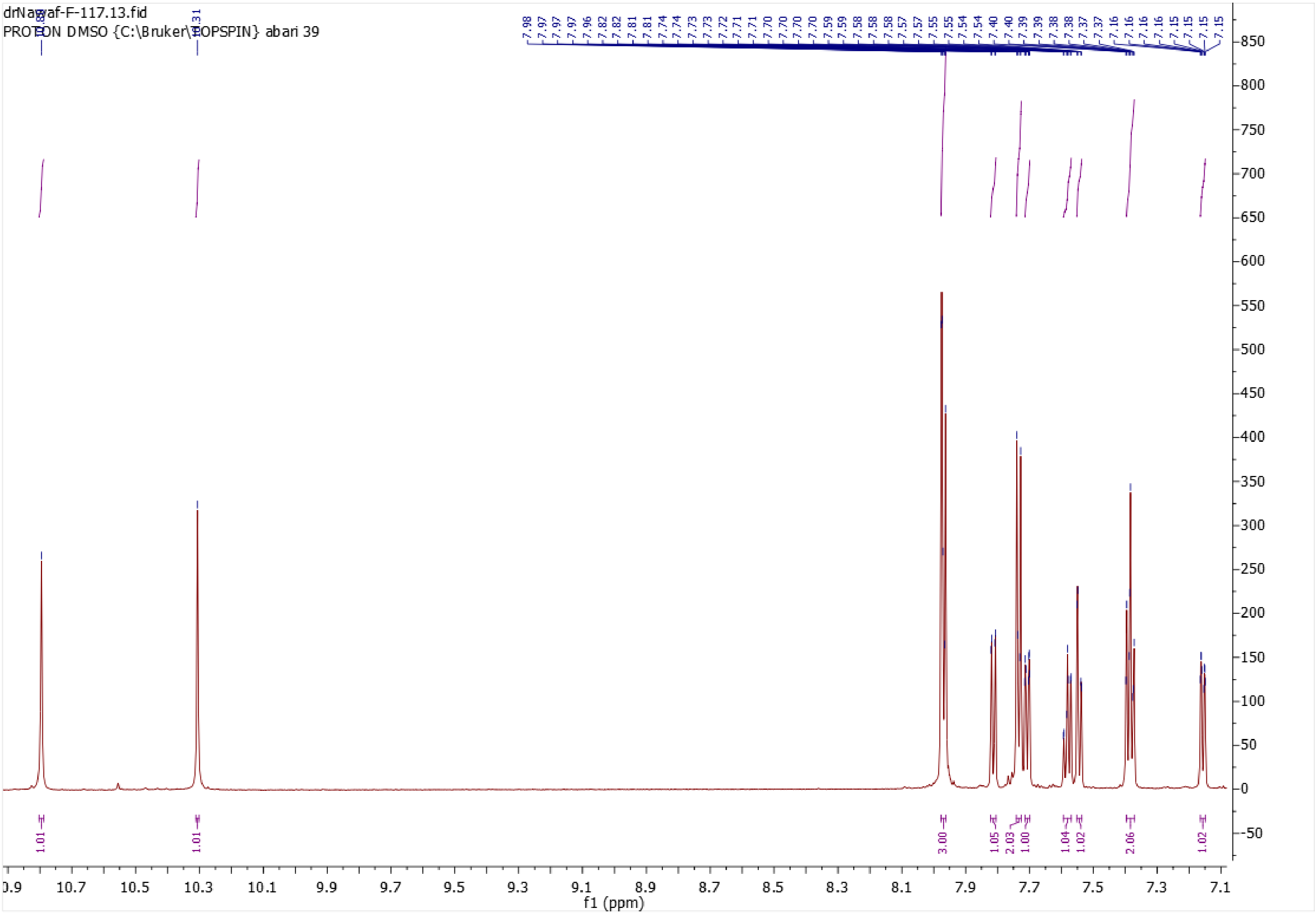

drNawaf-F-117.13.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 39
7.98
7.97
7.97
7.97
7.96
7.82
7.82
7.81
7.81
7.74
7.74
7.73
7.73
7.72
7.71
7.71
7.70
7.70
7.70
7.70
7.59
7.59
7.58
7.58
7.58
7.57
7.57
7.55
7.55
7.54
7.54
7.40
7.40
7.39
7.39
7.38
7.38
7.37
7.37
7.16
7.16
7.16
7.16
7.15
7.15
7.15
5.19
2.49
O
NH
O
NH
Cl
N
O
N
CH3
1.01
1.01
3.00
1.05
2.03
1.00
1.04
1.02
2.06
1.02
2.00
3.02
7000
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500
1.5
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
0.0
f1 (ppm)

17b

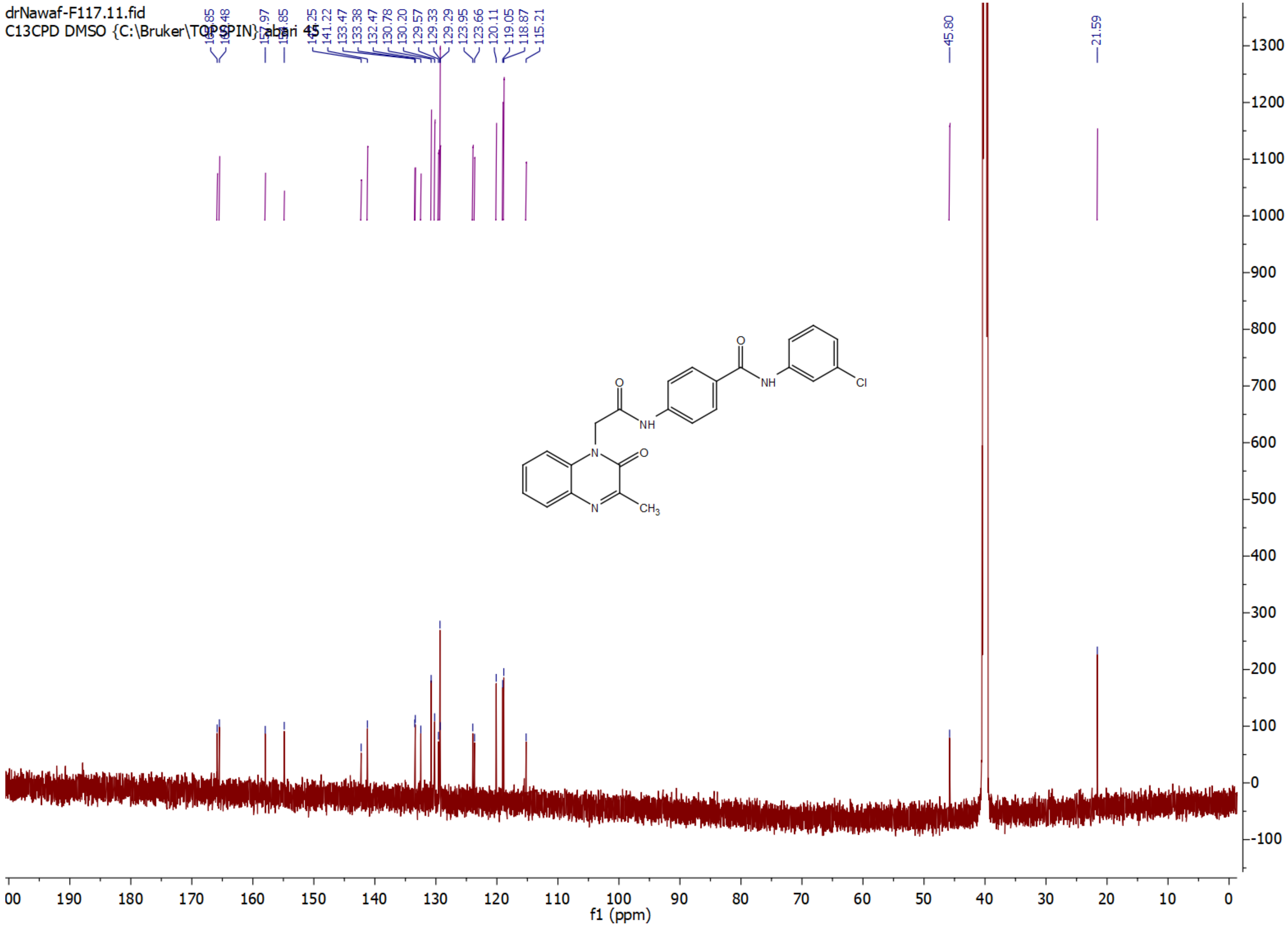


drNawaf-F-117.13.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 39
5.18
2.49
2.00
3.02
8500
8000
7500
7000
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500
.3
5.2
5.1
5.0
4.9
4.8
4.7
4.6
4.5
4.4
4.3
4.2
4.1
4.0
3.9
3.8
3.7
3.6
3.5
3.4
3.3
3.2
3.1
3.0
2.9
2.8
2.7
2.6
2.5
2.4
f1 (ppm)

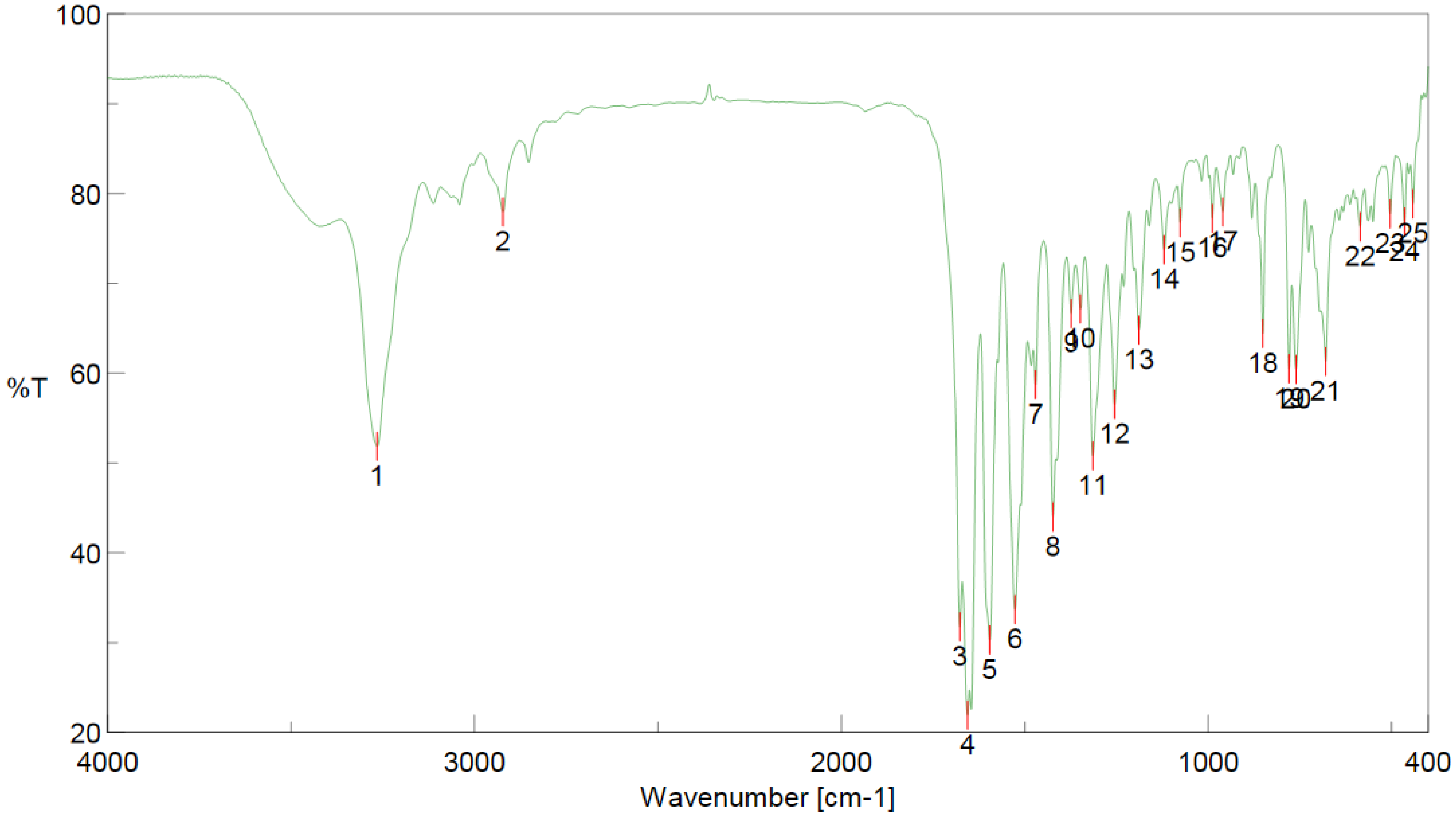
17b


drNawaf-F-117.13.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 39
10.88
10.31
7.98
7.97
7.97
7.97
7.96
7.82
7.82
7.81
7.81
7.74
7.74
7.73
7.73
7.72
7.71
7.71
7.70
7.70
7.70
7.70
7.59
7.59
7.58
7.58
7.58
7.57
7.57
7.55
7.55
7.54
7.54
7.40
7.40
7.39
7.39
7.38
7.38
7.37
7.37
7.16
7.16
7.16
7.15
7.15
7.15
7.15
1.01
1.01
3.00
1.05
2.03
1.00
1.04
1.02
2.06
1.02
850
800
750
700
650
600
550
500
450
400
350
300
250
200
150
100
50
0
-50
0.9
10.7
10.5
10.3
10.1
9.9
9.7
9.5
9.3
9.1
8.9
8.7
8.5
8.3
8.1
7.9
7.7
7.5
7.3
7.1
f1 (ppm)

17b

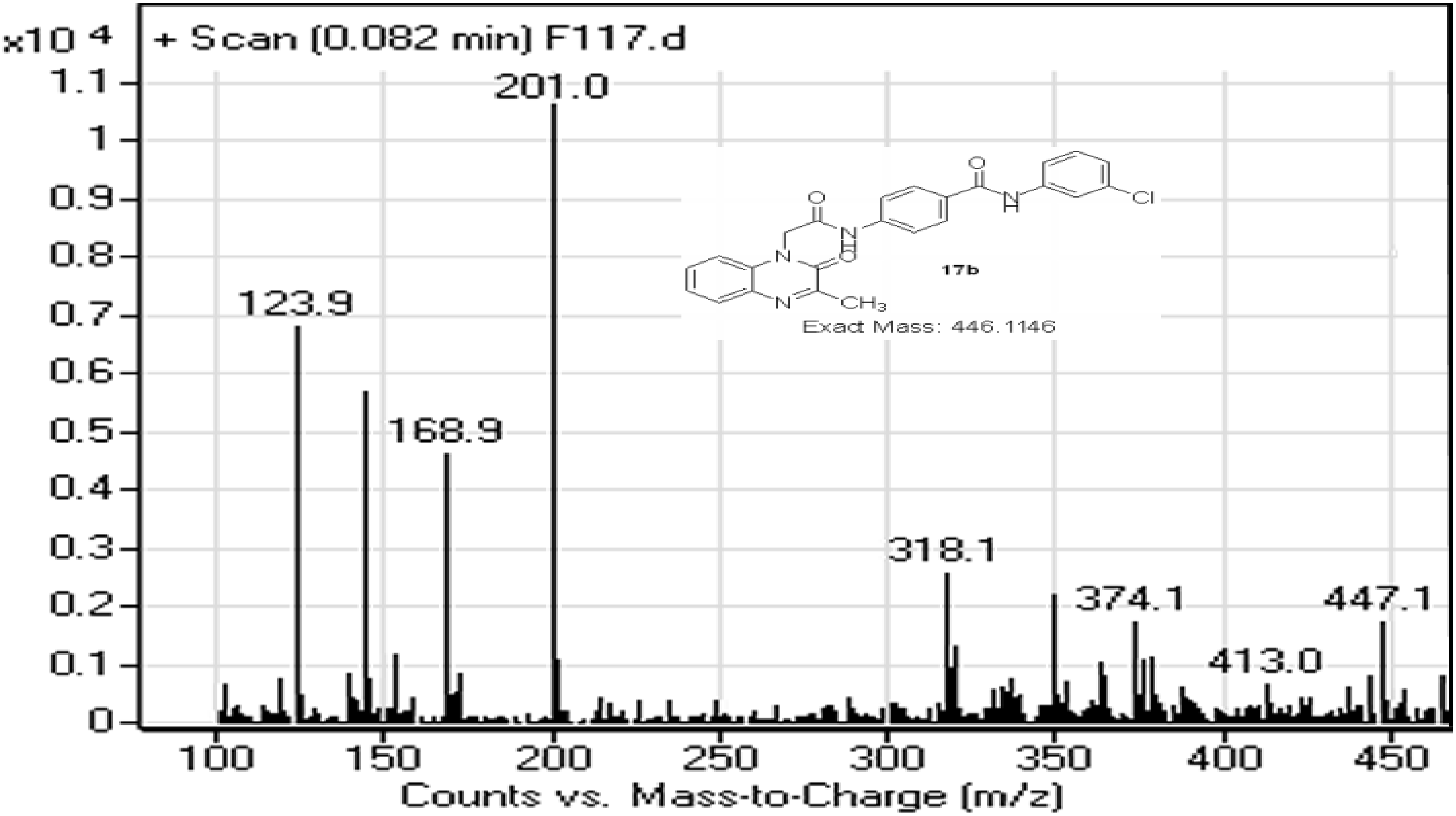

drNawaf-F117.11.fid
C13CPD DMSO {C:\Bruker\TOPSPIN} abari 45
165.85
165.48
157.97
154.85
142.25
141.22
133.47
133.38
132.47
130.78
130.20
129.57
129.33
129.29
123.95
123.66
120.11
119.05
118.87
115.21
45.80
21.59
O
NH
O
NH
Cl
N
O
N
CH3
1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
-100
00
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

17b

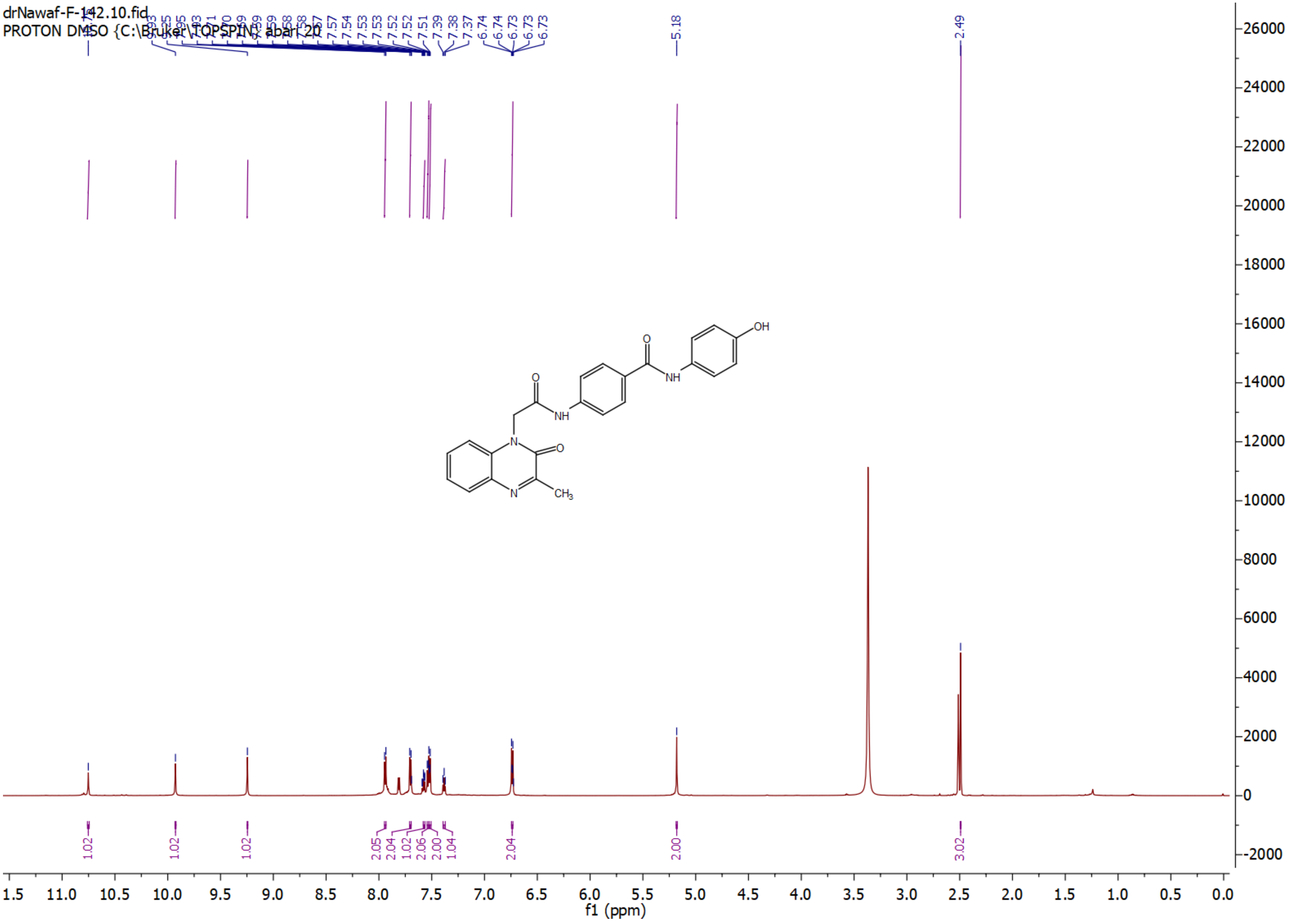


100
80
60
40
20
%T
4000
3000
2000
1000
400
Wavenumber [cm-1]
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

17b

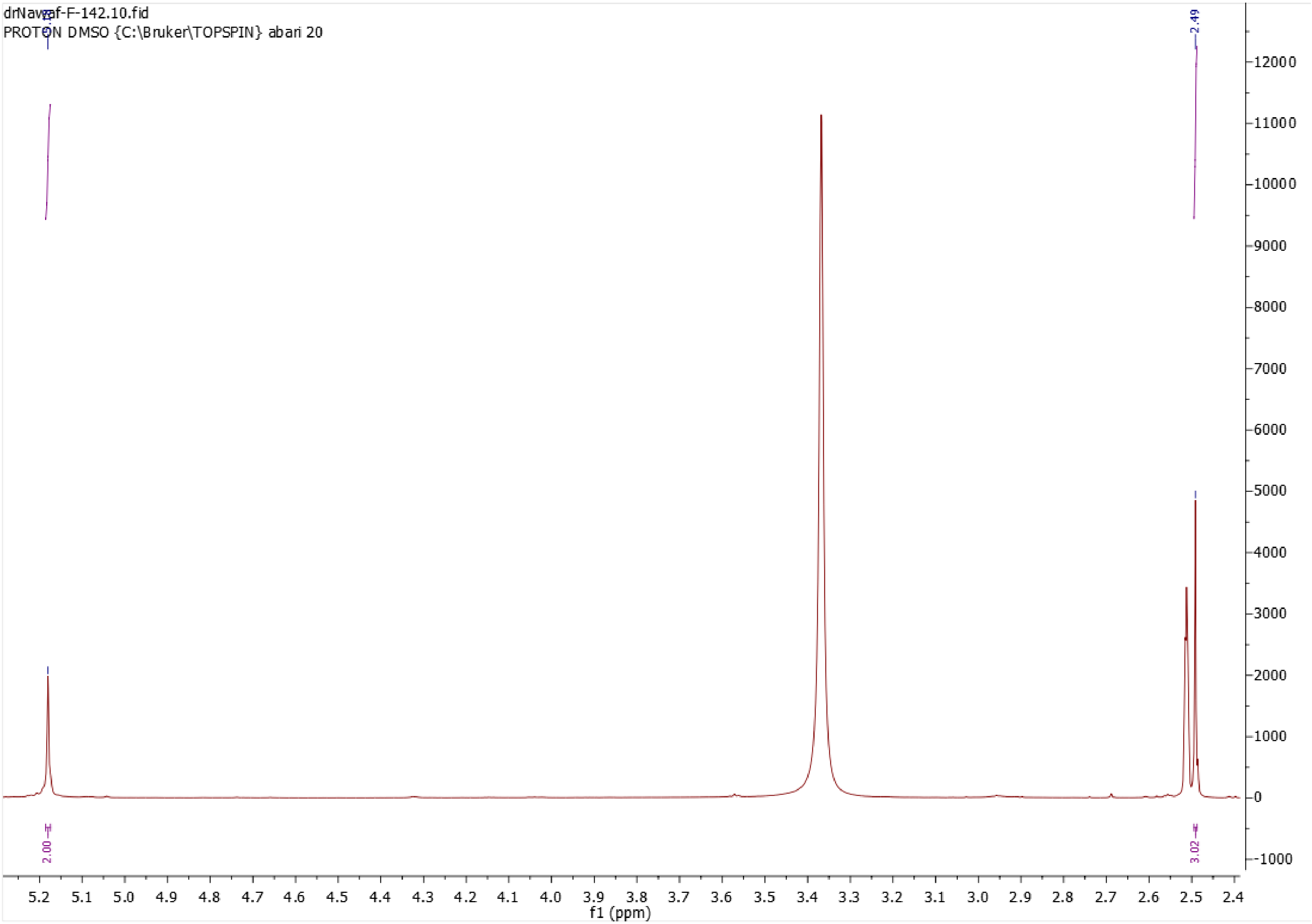

x10 4
+ Scan (0.082 min) F117.d
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
201.0
123.9
168.9
318.1
374.1
447.1
413.0
100
150
200
250
300
350
400
450
Counts vs. Mass-to-Charge (m/z)
O
N
H
Cl
CH3
17b
Exact Mass: 446.1146

17c



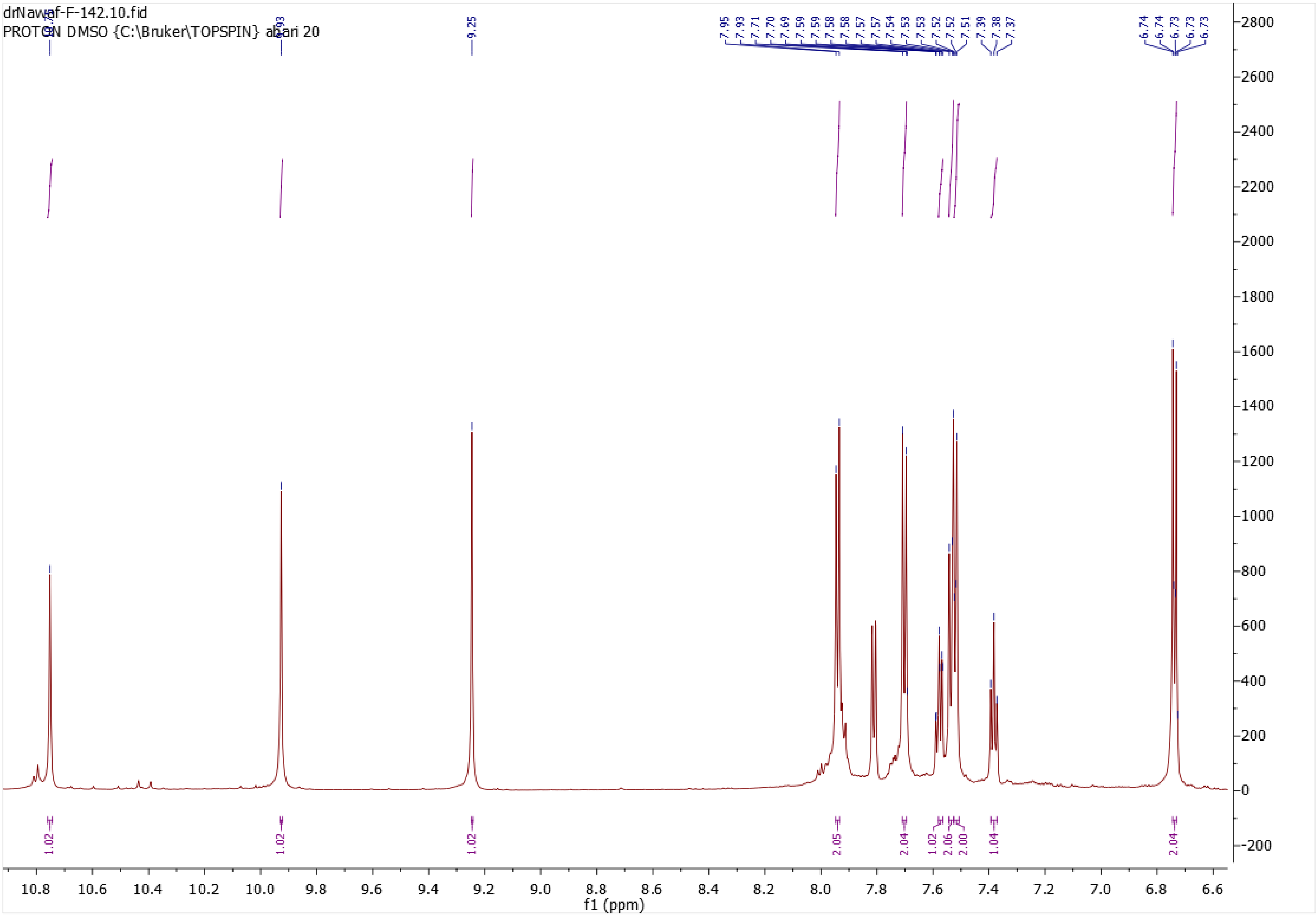
drNawaf-F-142.10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 20
10.73
9.93
9.25
7.95
7.93
7.71
7.70
7.59
7.59
7.59
7.58
7.58
7.57
7.57
7.54
7.53
7.53
7.52
7.52
7.51
7.39
7.38
7.37
6.74
6.74
6.73
6.73
6.73
5.18
2.49
O
NH
O
NH
OH
N
O
N
CH3
1.02
1.02
1.02
2.05
2.04
1.02
2.06
2.00
1.04
2.04
2.00
3.02
26000
24000
22000
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
-2000
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
0.0
f1 (ppm)

17c

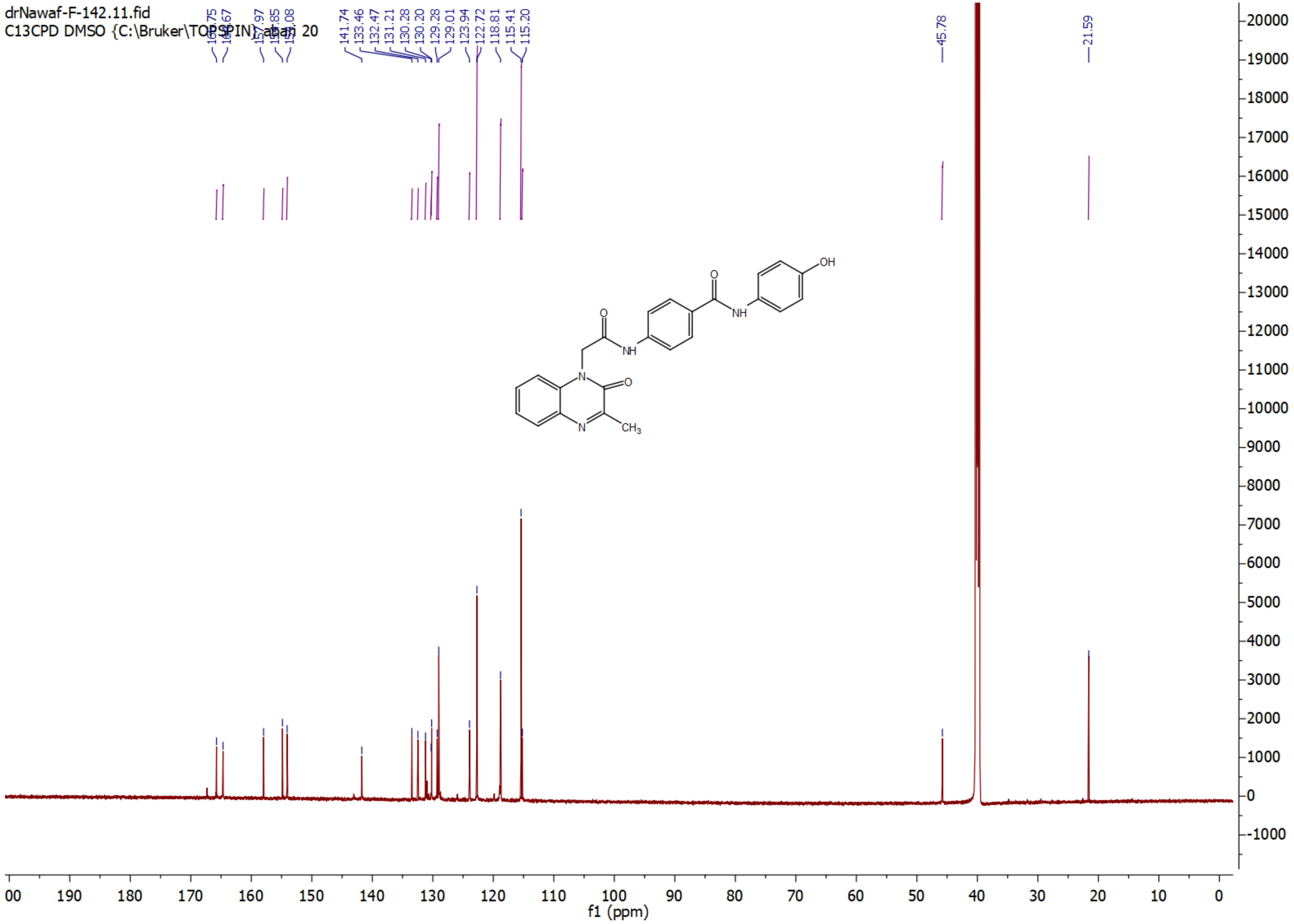


drNawaf-F-142.10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 20
5.18
2.49
2.00
3.02
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000
5.2 5.1 5.0 4.9 4.8 4.7 4.6 4.5 4.4 4.3 4.2 4.1 4.0 3.9 3.8 3.7 3.6 3.5 3.4 3.3 3.2 3.1 3.0 2.9 2.8 2.7 2.6 2.5 2.4
f1 (ppm)

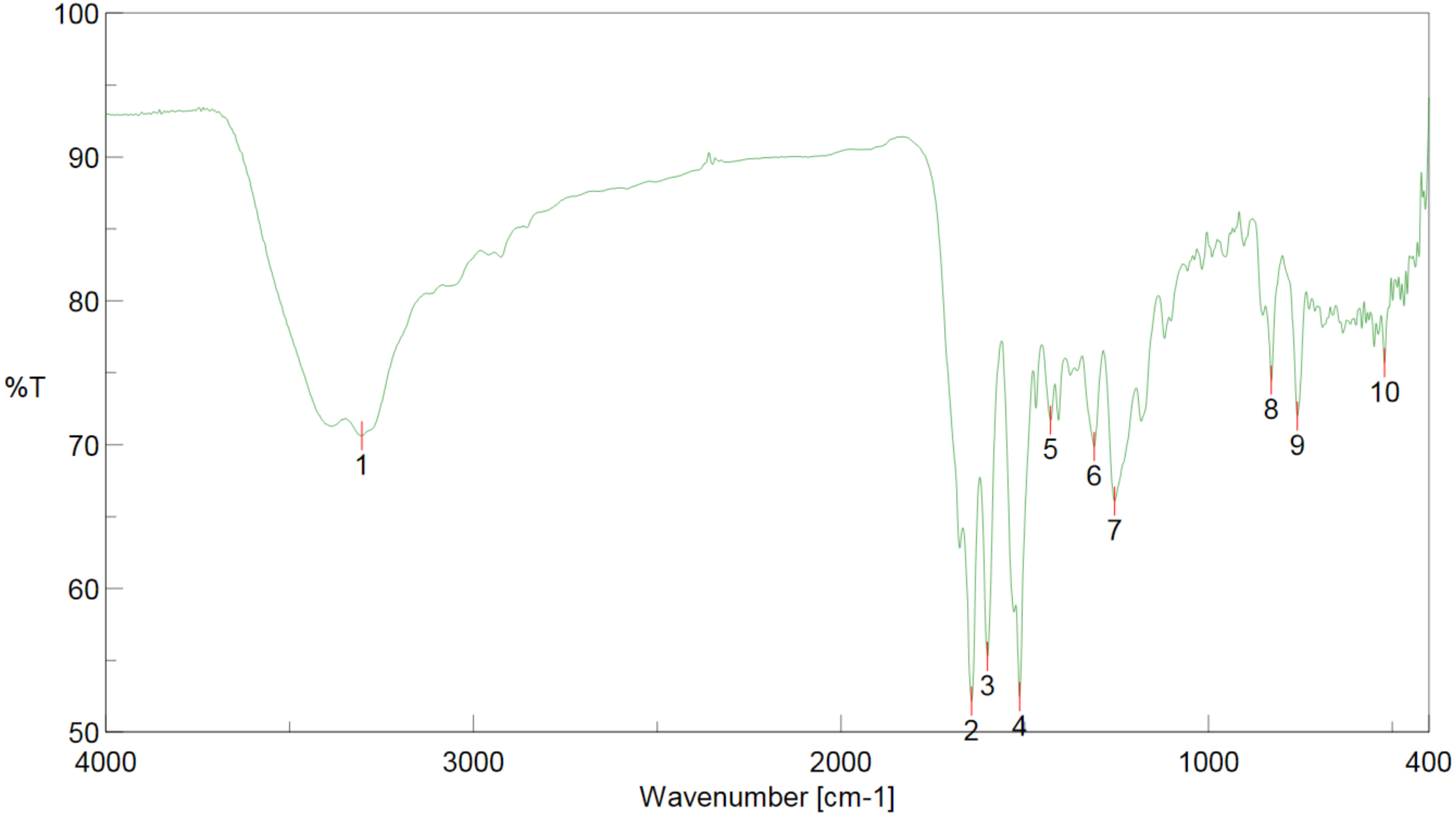
17c


drNawaf-F-142.10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 20
10.76
9.93
9.25
7.95
7.93
7.71
7.70
7.69
7.59
7.59
7.58
7.58
7.57
7.57
7.54
7.53
7.53
7.52
7.52
7.51
7.39
7.38
7.37
6.74
6.74
6.73
6.73
6.73
1.02
1.02
1.02
2.05
2.04
1.02
2.06
2.00
1.04
2.04
2800
2600
2400
2200
2000
1800
1600
1400
1200
1000
800
600
400
200
0
-200
10.8
10.6
10.4
10.2
10.0
9.8
9.6
9.4
9.2
9.0
8.8
8.6
8.4
8.2
8.0
7.8
7.6
7.4
7.2
7.0
6.8
6.6
f1 (ppm)

17c

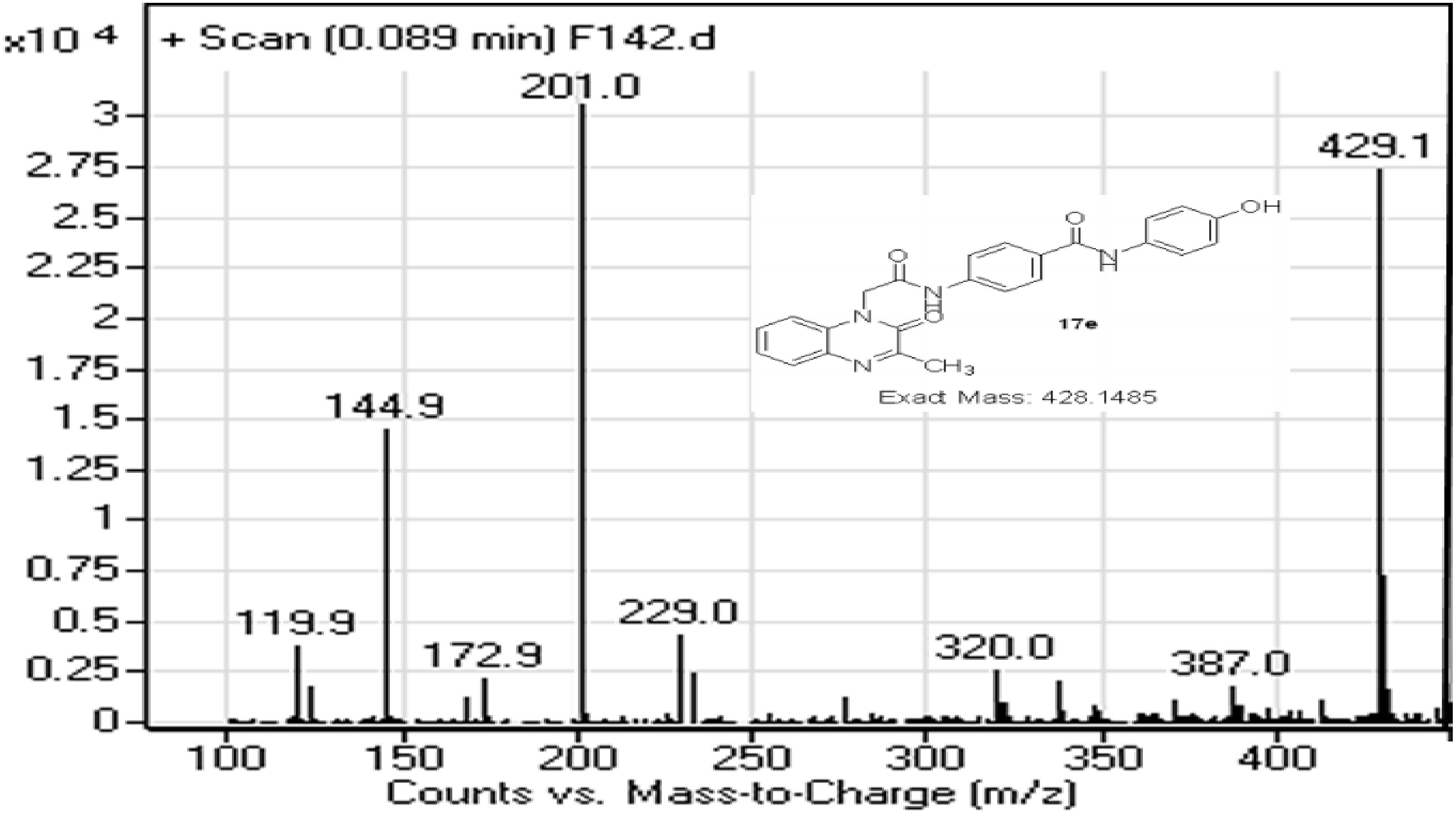

drNawaf-F-142.11.fid
C13CPD DMSO {C:\Bruker\TOPSPIN} abari 20
165.75
164.67
157.97
154.85
154.08
141.74
133.46
132.47
131.21
130.28
130.20
129.28
129.01
123.94
122.72
118.81
115.41
115.20
45.78
21.59
OH
O
NH
O
NH
N
O
N
CH3
20000
19000
18000
17000
16000
15000
14000
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000
00
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

17c



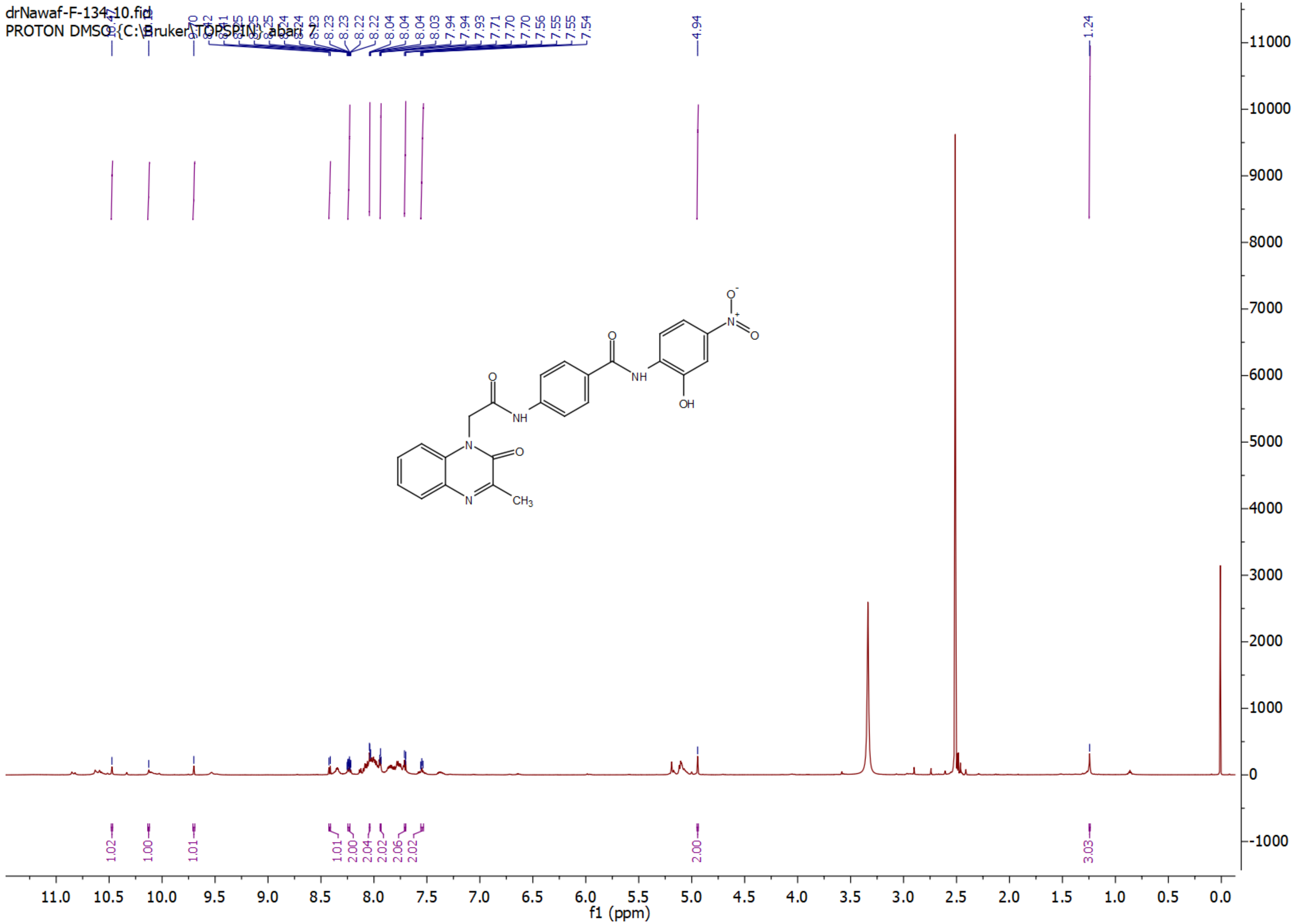
100
90
80
70
60
50
%T
4000
3000
2000
1000
400
Wavenumber [cm-1]
1
2
3
4
5
6
7
8
9
10

17c

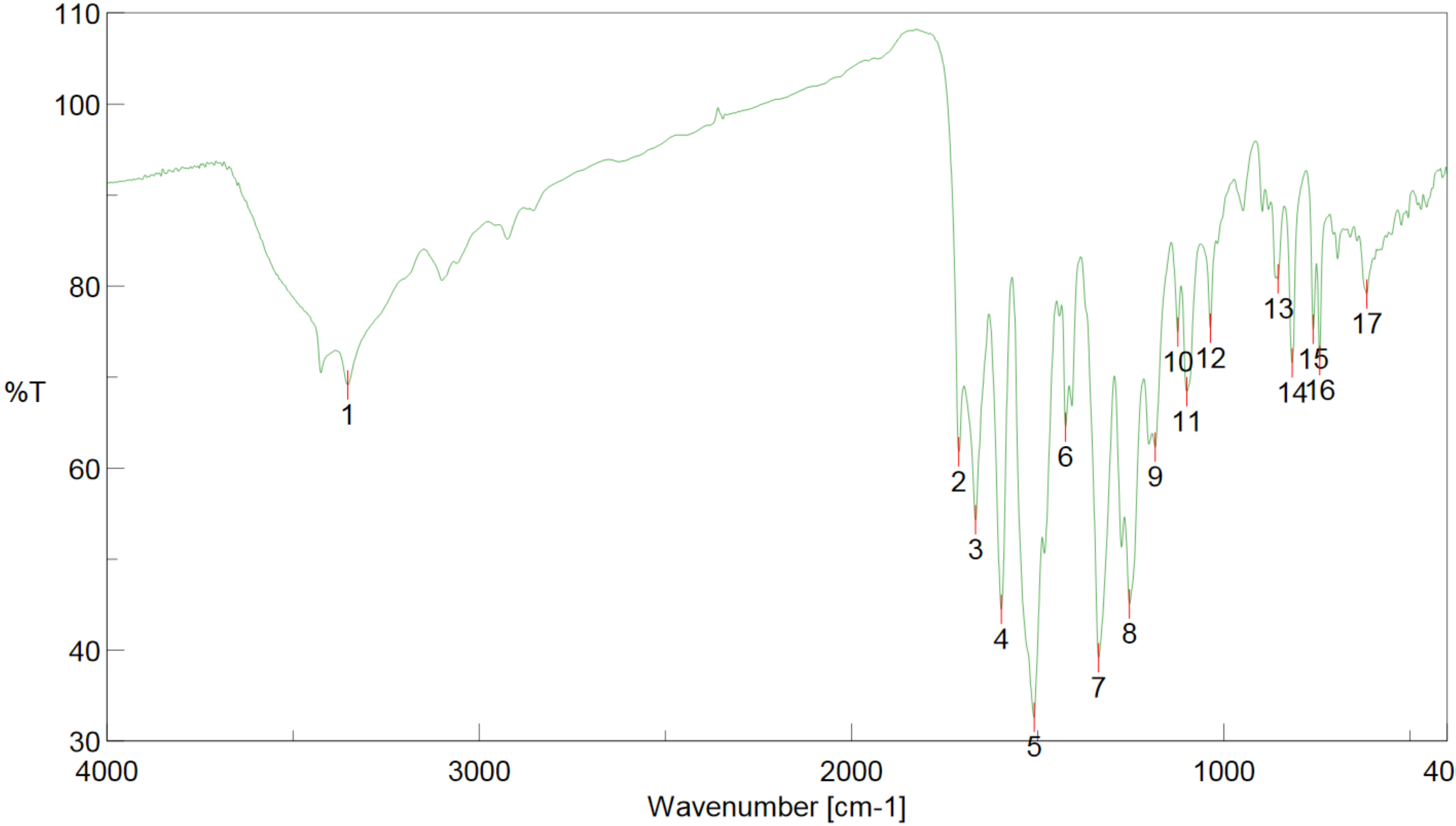


x10 4
+ Scan (0.089 min) F142.d
201.0
429.1
144.9
119.9
172.9
229.0
320.0
387.0
3
2.75
2.5
2.25
2
1.75
1.5
1.25
1
0.75
0.5
0.25
0
100
150
200
250
300
350
400
Counts vs. Mass-to-Charge (m/z)
O
N
H
O
H
N
OH
N
O
N
CH3
17e
Exact Mass: 428.1485

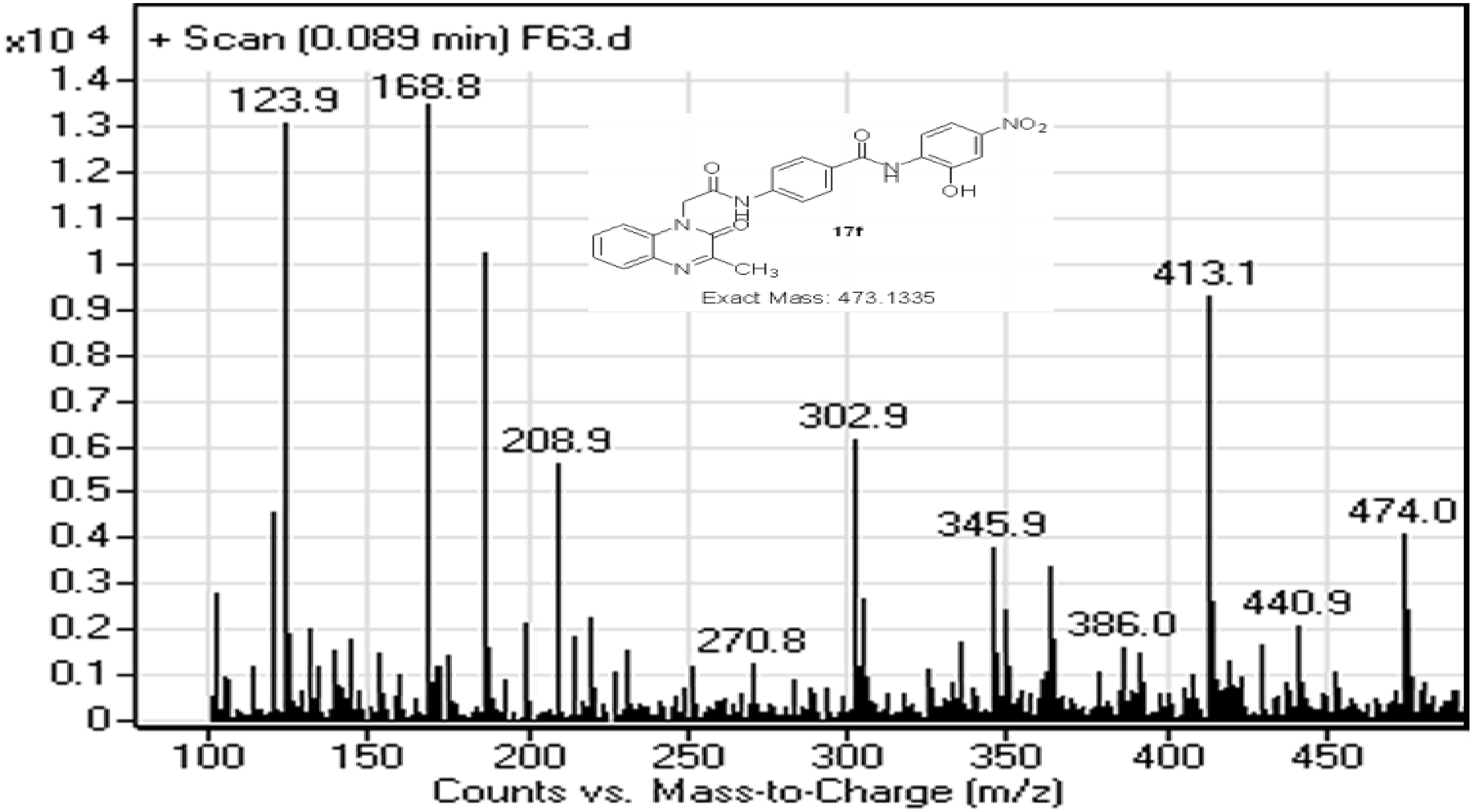
17d


drNawaf-F-134-10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 7
10.47
9.70
8.42
8.41
8.25
8.25
8.25
8.24
8.24
8.23
8.23
8.23
8.22
8.22
8.04
8.04
8.04
8.03
7.94
7.94
7.93
7.71
7.70
7.70
7.56
7.55
7.55
7.54
4.94
1.24
1.02
1.00
1.01
1.01
2.00
2.04
2.02
2.06
2.02
2.00
3.03
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000
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
0.0
f1 (ppm)
O
N
O
NH
O
NH
OH
O
N
O
CH3
N

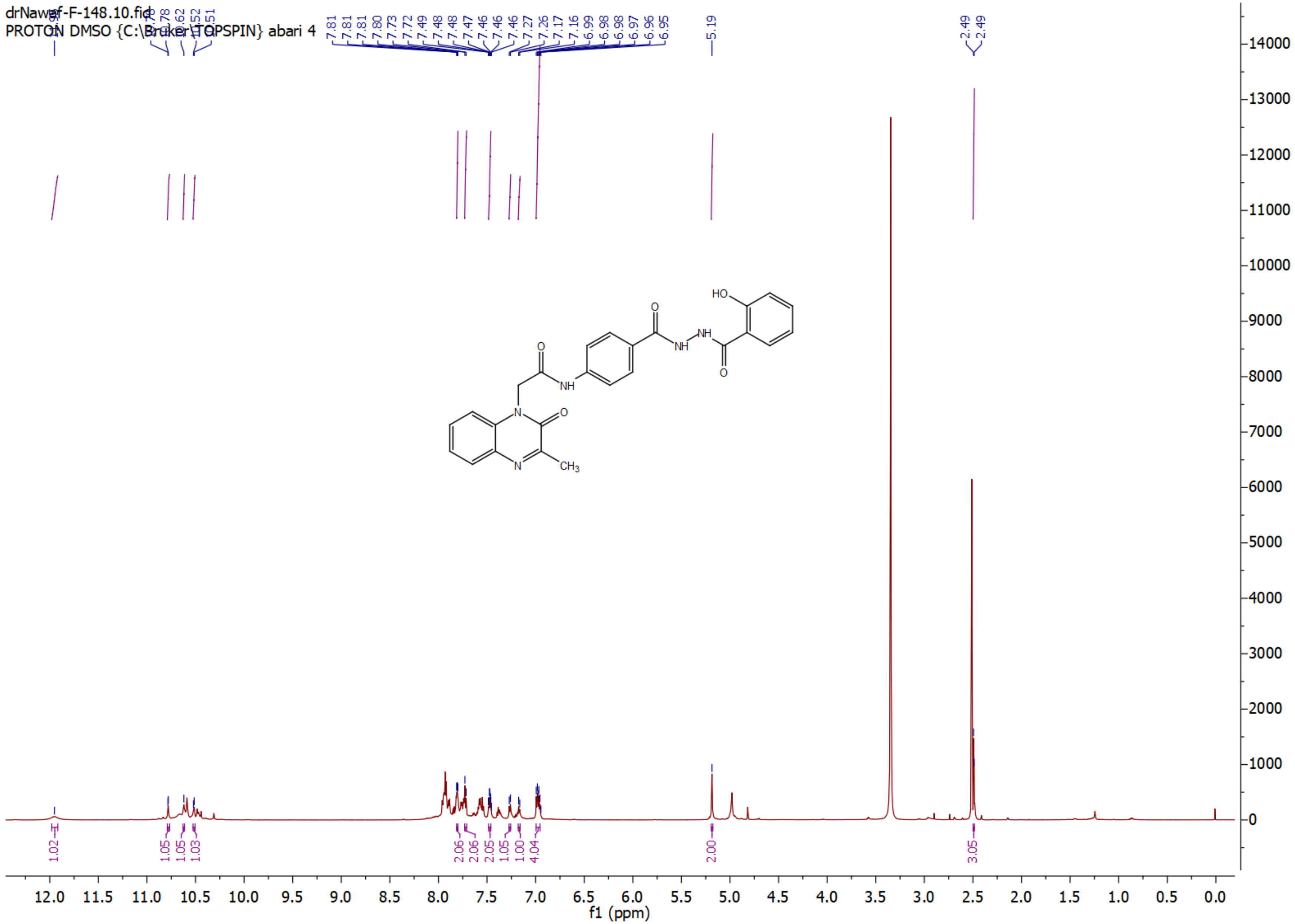
17d


110
100
80
60
40
30
%T
4000
3000
2000
1000
400
Wavenumber [cm-1]
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17

17d

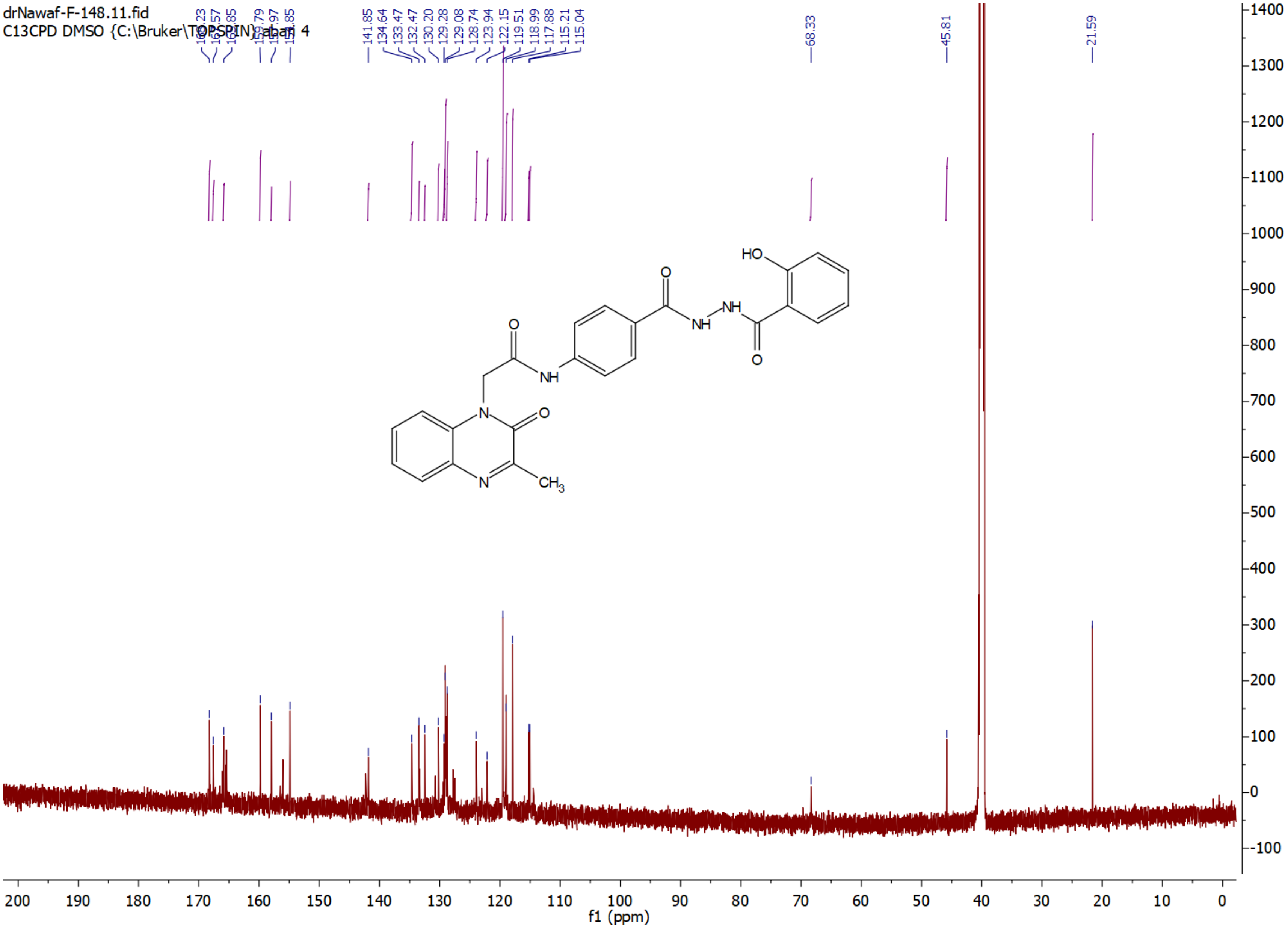

x10 4
+ Scan (0.089 min) F63.d
1.4
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
123.9
168.8
208.9
270.8
302.9
345.9
386.0
413.1
440.9
474.0
100
150
200
250
300
350
400
450
Counts vs. Mass-to-Charge (m/z)
NO2
O
N
H
OH
CH3
17f
Exact Mass: 473.1335

18

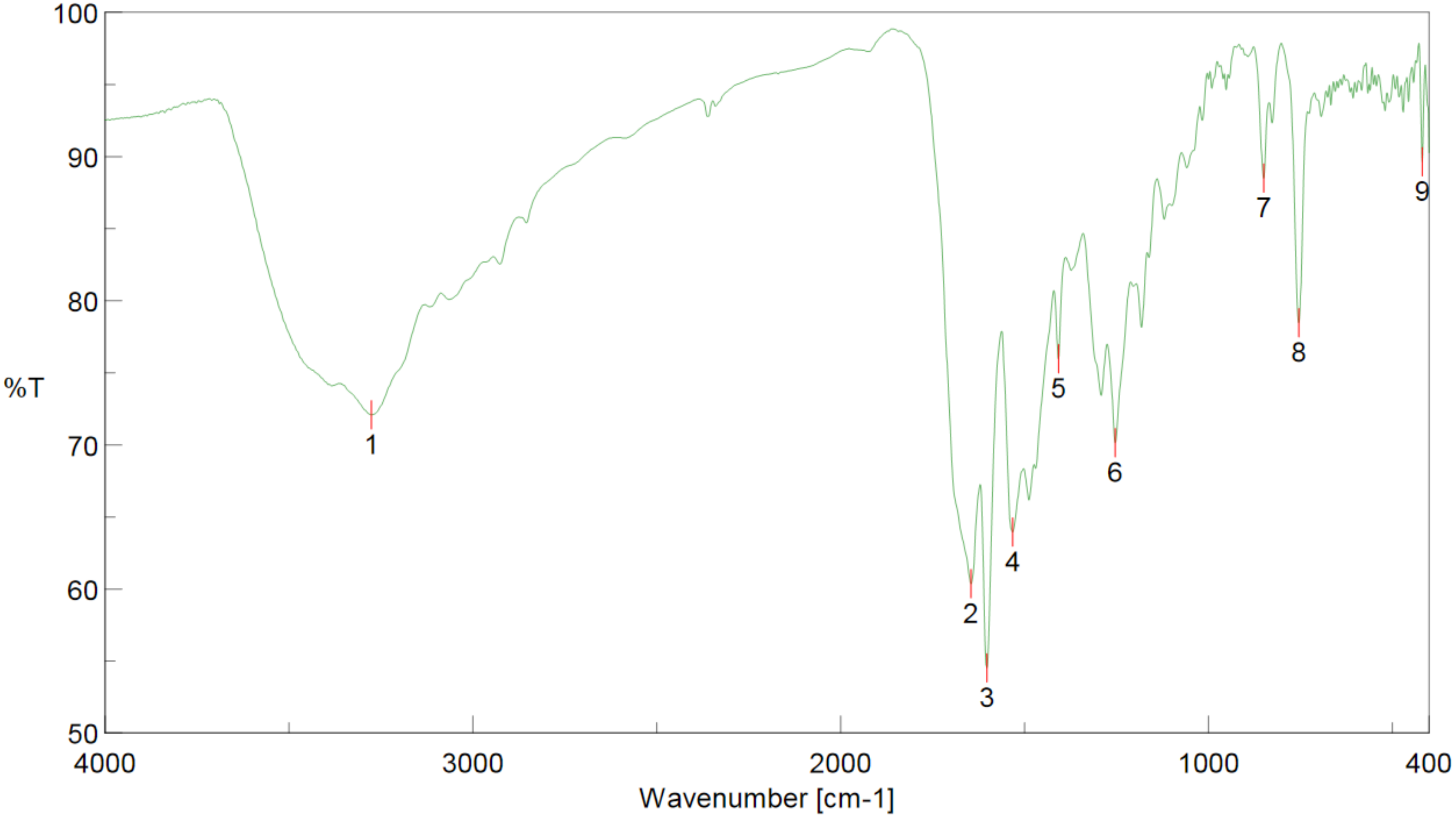


drNawaf-F-148.10.fid
PROTON DMSO {C:\Bruker\TOPSPIN} abari 4
10.78
10.62
10.52
10.51
7.81
7.81
7.81
7.80
7.73
7.72
7.49
7.48
7.48
7.47
7.46
7.46
7.46
7.27
7.26
7.17
7.16
6.99
6.98
6.98
6.97
6.96
6.95
5.19
2.49
2.49
HO
O
NH
NH
O
O
NH
N
O
N
CH3
1.02
1.05
1.05
1.03
2.06
2.06
2.05
1.05
1.00
4.04
2.00
3.05
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
0.0
f1 (ppm)
14000
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0

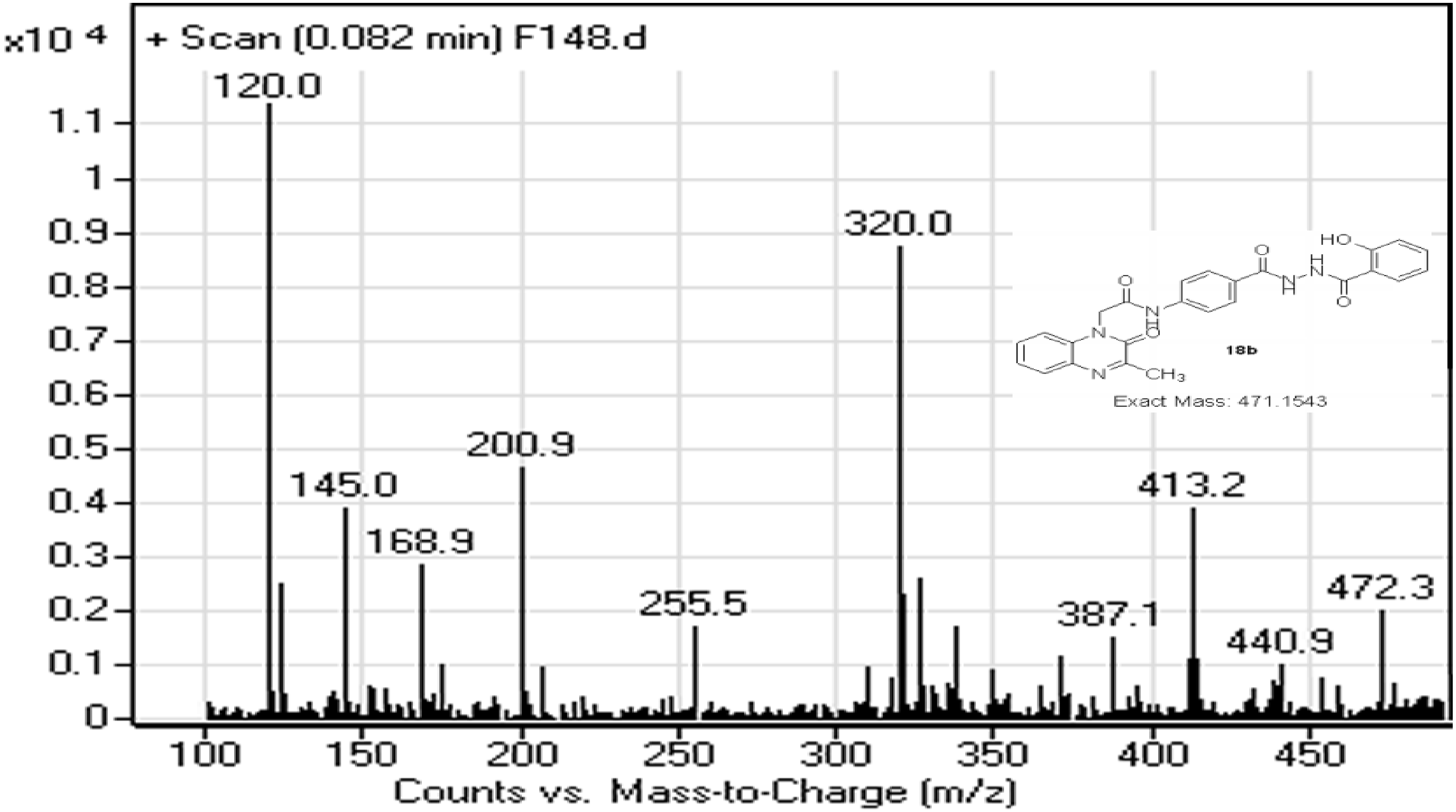
18


drNawaf-F-148.11.fid
C13CPD DMSO {C:\Bruker\TOPSPIN} abari 4
168.23
167.57
165.85
159.79
157.97
154.85
141.85
134.64
133.47
132.47
130.20
129.28
129.08
128.74
123.94
122.15
119.51
118.99
117.88
115.21
115.04
68.33
45.81
21.59
HO
O
NH
NH
O
O
NH
N
O
N
CH3
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
-100
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

18


100
90
80
70
60
50
%T
4000
3000
2000
1000
400
Wavenumber [cm-1]
1
2
3
4
5
6
7
8
9

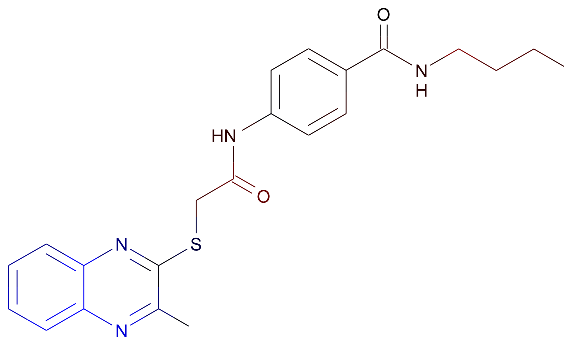
18


x10 4
+ Scan (0.082 min) F148.d
120.0
145.0
168.9
200.9
255.5
320.0
387.1
413.2
440.9
472.3
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
100
150
200
250
300
350
400
450
Counts vs. Mass-to-Charge (m/z)
HO
18b
CH3
Exact Mass: 471.1543

NMR spectra

15a

TOPKAT_Aerobic_Biodegradability



$C_{22}H_{24}N_4O_2S$

Molecular Weight: 408.51656

ALogP: 3.83

Rotatable Bonds: 8

Acceptors: 5

Donors: 2

Model Prediction

Prediction: Non-Degradable

Probability: 0.318

Enrichment: 0.729

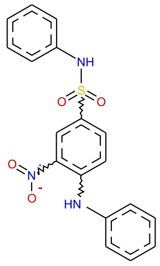
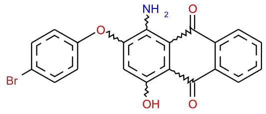
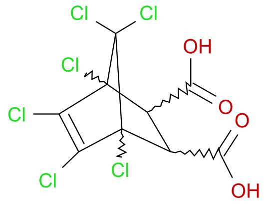
Bayesian Score: -3.803

Mahalanobis Distance: 15.266

Mahalanobis Distance p-value: 7.74e-013

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	Benzenesulfonamide,_3-nitro-N-phenyl-4-(phenylamino)-	9,10-Anthracenedione,_1-amino-2-(4-bromophenoxy)-4-hydroxy-	Bicyclo_2.2.1_hept-5-ene-2,3-dicarboxylic_acid,_1,4,5,6,7,7-hexachloro-
Structure			
Actual Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Predicted Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Distance	0.606	0.667	0.728
Reference	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.

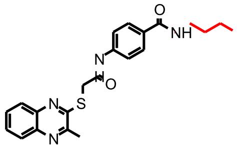
Model Applicability

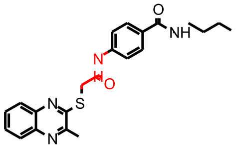
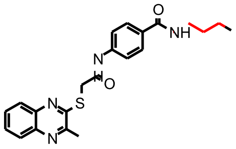
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. OPS PC18 out of range. Value: 4.61. Training min, max, SD, explained variance: -3.5239, 4.5287, 1.221, 0.0171.

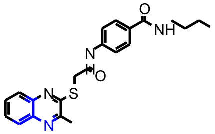
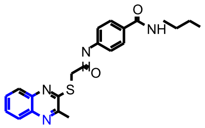
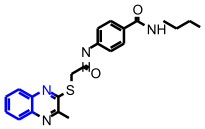
Feature Contribution

Top features for positive contribution

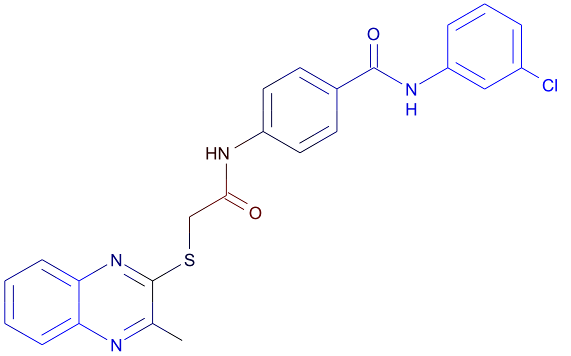
Fingerprint	Bit/Smiles	Feature Structure	Score	Degradable in training set
SCFP_12	-1577600103	 [*]CCCC	0.631	113 out of 147

SCFP_12	1256995004	 [*]CC(=O)N[*]	0.540	23 out of 32
SCFP_12	-1272798659	 [*]CCC[*]	0.518	160 out of 234

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Degradable in training set**
SCFP_12	112554633	 [*]:[cH]:[c](:n:[*]):[c](:[*]):[*]	-1.077	1 out of 12
SCFP_12	-1381862798	 [*]:n:[c]1:[cH]:[cH]:[cH]:[*]:[c]:1:[*]	-1.077	1 out of 12
SCFP_12	1655199790	 [*]:n:[c]1:[cH]:[cH]:[cH]:[cH]:[c]:1:[*]	-1.005	1 out of 11

15b



$C_{24}H_{19}ClN_4O_2S$

Molecular Weight: 462.95126

ALogP: 4.743

Rotatable Bonds: 6

Acceptors: 5

Donors: 2

Model Prediction

Prediction: Non-Degradable

Probability: 0.029

Enrichment: 0.067

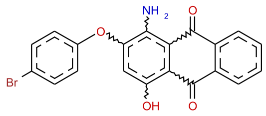
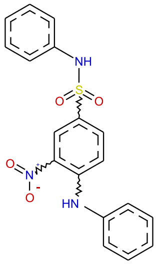
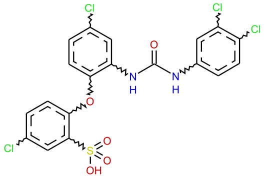
Bayesian Score: -18.792

Mahalanobis Distance: 15.602

Mahalanobis Distance p-value: 5.79e-014

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	**9,10-Anthracenedione,_1-amino-2-(4-bromophenoxy)-4-hydroxy-**	**Benzenesulfonamide,_3-nitro-N-phenyl-4-(phenylamino)-**	**Mitin_FF**
Structure			
Actual Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Predicted Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Distance	0.641	0.681	0.733
Reference	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.

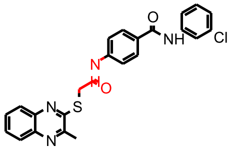
Model Applicability

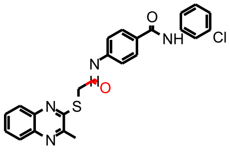
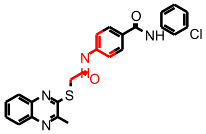
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

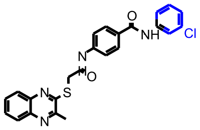
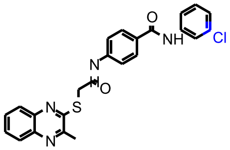
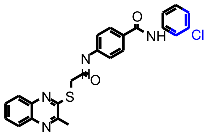
Feature Contribution

Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Degradable in training set**
SCFP_12	1256995004	 [*]CC(=O)N[*]	0.540	23 out of 32

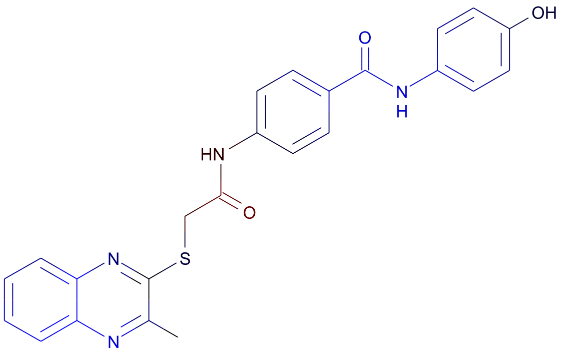
SCFP_12	1311071855	 [*]C(=O)[*]	0.461	173 out of 268
SCFP_12	2097618059	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	0.446	5 out of 7

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Degradable in training set
SCFP_12	-1378360678	 Cl[c]1:[cH]:[*]:[cH]:[cH]:[cH]:1	-1.878	3 out of 62
SCFP_12	-52074512	 [*]:[c](:[*])Cl	-1.865	5 out of 93
SCFP_12	-601571304	 [*]:[cH]:[c](Cl):[cH]:[*]	-1.844	5 out of 91

15c

TOPKAT_Aerobic_Biodegradability



$C_{24}H_{20}N_4O_3S$

Molecular Weight: 444.5056

ALogP: 3.836

Rotatable Bonds: 6

Acceptors: 6

Donors: 3

Model Prediction

Prediction: Non-Degradable

Probability: 0.093

Enrichment: 0.213

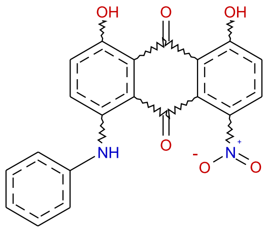
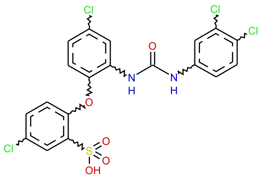
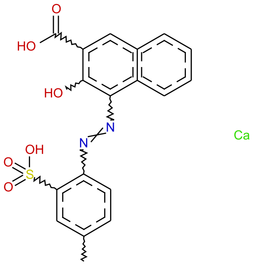
Bayesian Score: -12.054

Mahalanobis Distance: 15.808

Mahalanobis Distance p-value: 1.13e-014

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	9,10-Anthracenedione,_1,8-dihydroxy-4-nitro-5-(phenylamino)-	Mitin_FF	Brilliant_Blue_6B
Structure			
Actual Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Predicted Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Distance	0.689	0.707	0.723
Reference	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.

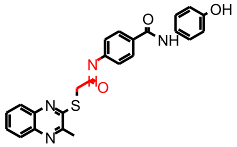
Model Applicability

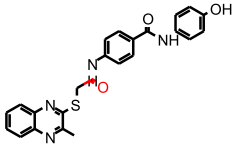
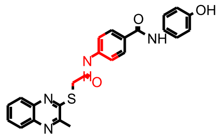
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. OPS PC23 out of range. Value: 3.602. Training min, max, SD, explained variance: -3.0926, 3.437, 1.042, 0.0125.

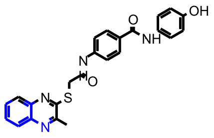
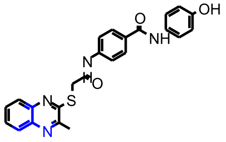
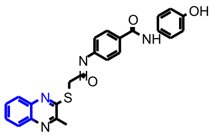
Feature Contribution

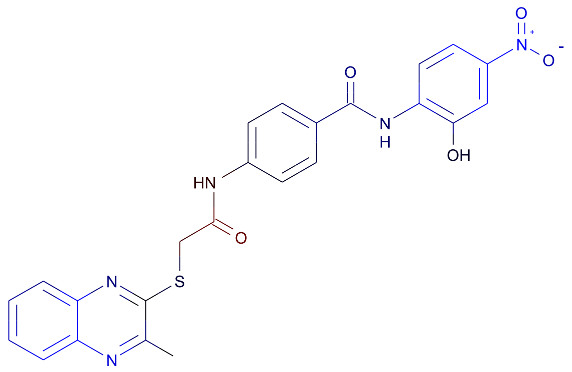
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Degradable in training set
SCFP_12	1256995004	 [*]CC(=O)N[*]	0.540	23 out of 32

SCFP_12	1311071855	 [*]C(=O)[*]	0.461	173 out of 268
SCFP_12	2097618059	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	0.446	5 out of 7

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Degradable in training set**
SCFP_12	-1381862798	 [*]:n:[c]1:[cH]:[cH]: [cH]:[*]:[c]:1:[*]	-1.077	1 out of 12
SCFP_12	112554633	 [*]:[cH]:[c](:n:[*]): [c](:[*]):[*]	-1.077	1 out of 12
SCFP_12	1655199790	 [*]:n:[c]1:[cH]:[cH]: [cH]:[cH]:[c]:1:[*]	-1.005	1 out of 11



$C_{24}H_{19}N_5O_5S$

Molecular Weight: 489.50316

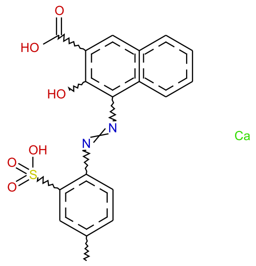
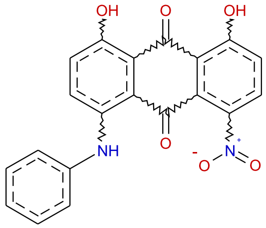
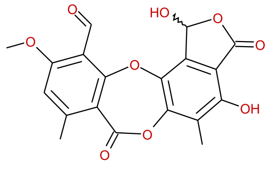
ALogP: 3.731

Rotatable Bonds: 7

Acceptors: 8

Donors: 3

Structural Similar Compounds

Name	**Brilliant_Blue_6B**	**9,10-Anthracenedione,_1,8-dihydroxy-4-nitro-5-(phenylamino)-**	**Methylpyridine**
Structure			
Actual Endpoint	Non-Degradable	Non-Degradable	Degradable
Predicted Endpoint	Non-Degradable	Non-Degradable	Degradable
Distance	0.621	0.722	0.838
Reference	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.

Model Prediction

Prediction: Non-Degradable

Probability: 0.022

Enrichment: 0.051

Bayesian Score: -20.302

Mahalanobis Distance: 21.366

Mahalanobis Distance p-value: 3.58e-038

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.

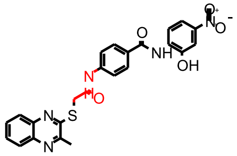
Model Applicability

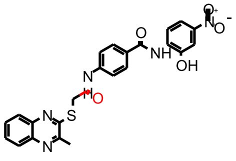
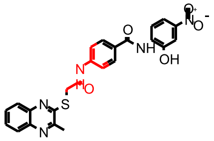
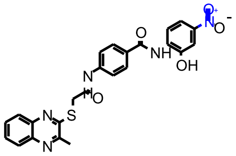
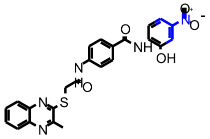
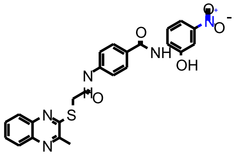
Unknown features are fingerprint features in the query molecule, but not found in the training set.

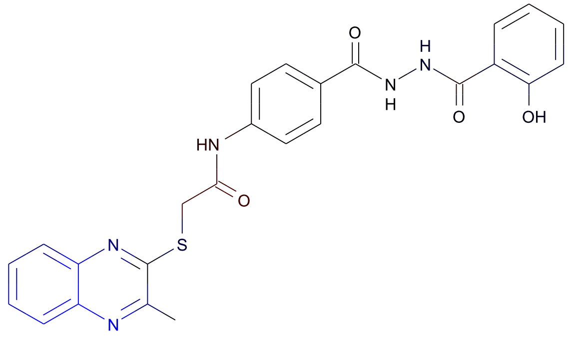
1. OPS PC18 out of range. Value: 4.9115. Training min, max, SD, explained variance: -3.5239, 4.5287, 1.221, 0.0171.

Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Degradable in training set
SCFP_12	1256995004	 [*]CC(=O)N[*]	0.540	23 out of 32

<table>
<tr><td>SCFP_12</td><td>1311071855</td><td>
[*]C(=O)[*]</td><td>0.461</td><td>173 out of 268</td></tr>
<tr><td>SCFP_12</td><td>2097618059</td><td>
[*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]</td><td>0.446</td><td>5 out of 7</td></tr>
<tr><th colspan="5">Top Features for negative contribution</th></tr>
<tr><th>Fingerprint</th><th>Bit/Smiles</th><th>Feature Structure</th><th>Score</th><th>Degradable in training set</th></tr>
<tr><td>SCFP_12</td><td>1311339974</td><td>
[*][N+](=O)[*]</td><td>-1.780</td><td>3 out of 56</td></tr>
<tr><td>SCFP_12</td><td>1334669481</td><td>
[*][N+](=[*])[c](:[cH]:[*]):[cH]:[*]</td><td>-1.631</td><td>10 out of 136</td></tr>
<tr><td>SCFP_12</td><td>10</td><td>
[*][N+](=[*])[*]</td><td>-1.607</td><td>11 out of 145</td></tr>
</table>



$C_{25}H_{21}N_5O_4S$

Molecular Weight: 487.53034

ALogP: 3.093

Rotatable Bonds: 7

Acceptors: 7

Donors: 4

Model Prediction

Prediction: Non-Degradable

Probability: 0.172

Enrichment: 0.394

Bayesian Score: -8.173

Mahalanobis Distance: 15.835

Mahalanobis Distance p-value: 9.05e-015

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.

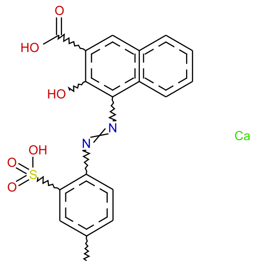
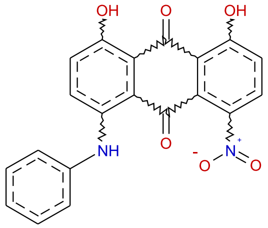
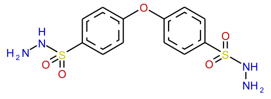
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.

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	**Brilliant_Blue_6B**	**9,10-Anthracenedione,_1,8-dihydroxy-4-nitro-5-(phenylamino)-**	**Benzenesulfonic_acid,_4,4'-oxybis-,_dihydrazide**
Structure			
Actual Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Predicted Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Distance	0.782	0.807	0.880
Reference	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.

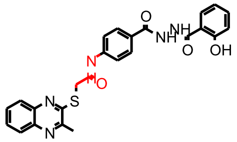
Model Applicability

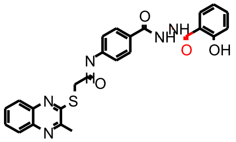
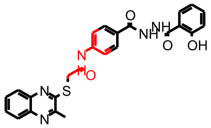
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

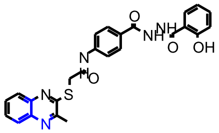
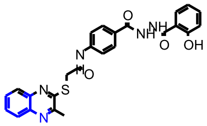
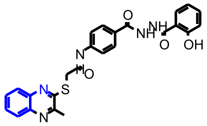
Feature Contribution

Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Degradable in training set**
SCFP_12	1256995004	 [*]CC(=O)N[*]	0.540	23 out of 32

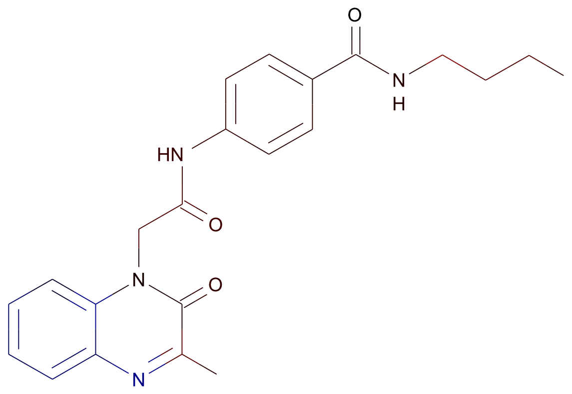
SCFP_12	1311071855	 [*]C(=O)[*]	0.461	173 out of 268
SCFP_12	2097618059	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	0.446	5 out of 7

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Degradable in training set**
SCFP_12	112554633	 [*]:[cH]:[c](:n:[*]):[c](:[*]):[*]	-1.077	1 out of 12
SCFP_12	-1381862798	 [*]:n:[c]1:[cH]:[cH]:[cH]:[*]:[c]:1:[*]	-1.077	1 out of 12
SCFP_12	1655199790	 [*]:n:[c]1:[cH]:[cH]:[cH]:[cH]:[c]:1:[*]	-1.005	1 out of 11

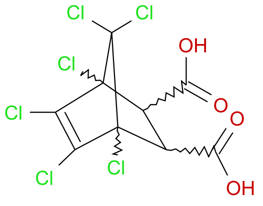
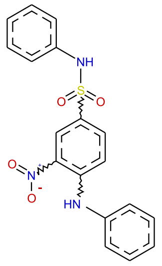
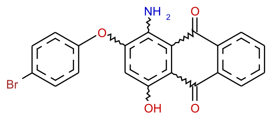
17a

TOPKAT_Aerobic_Biodegradability



$C_{22}H_{24}N_4O_3$
Molecular Weight: 392.45096
ALogP: 2.212
Rotatable Bonds: 7
Acceptors: 4
Donors: 2

Structural Similar Compounds

Name	Bicyclo_2.2.1_hept-5-ene-2,3-dicarboxylic_acid,_1,4,5,6,7,7-hexachloro-	Benzenesulfonamide,_3-nitro-N-phenyl-4-(phenylamino)-	9,10-Anthracenedione,_1-amino-2-(4-bromophenoxy)-4-hydroxy-
Structure			
Actual Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Predicted Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Distance	0.647	0.689	0.723
Reference	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.

Model Prediction

Prediction: Degradable
Probability: 0.613
Enrichment: 1.406
Bayesian Score: 2.598
Mahalanobis Distance: 13.176
Mahalanobis Distance p-value: 8.54e-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.

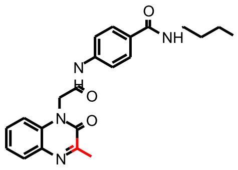
Model Applicability

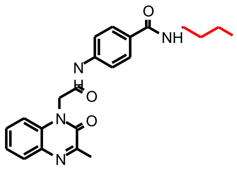
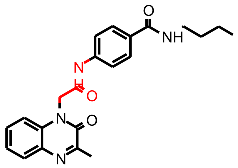
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

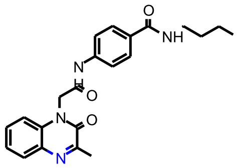
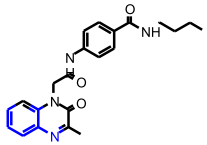
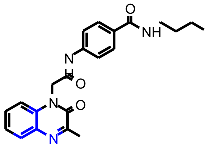
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Degradable in training set
SCFP_12	136627117	 [*]C(=[*])C	0.662	56 out of 70

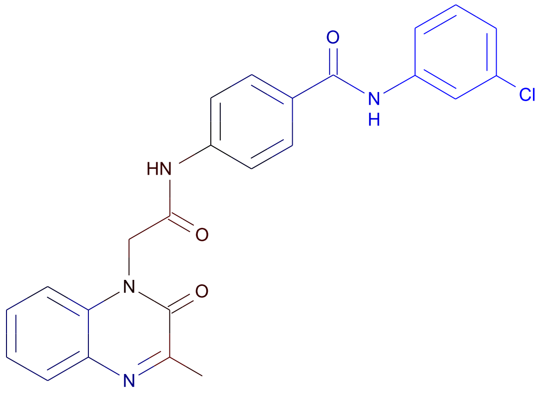
SCFP_12	-1577600103	 [*]CCCC	0.631	113 out of 147
SCFP_12	1256995004	 [*]CC(=O)N[*]	0.540	23 out of 32

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Degradable in training set
SCFP_12	6	 [*]N=[*]	-1.018	2 out of 18
SCFP_12	-1377141613	 [*][c]1:[*]:[cH]:[cH] :[cH]:[c]:1N=[*]	-0.964	0 out of 4
SCFP_12	1851000357	 [*][c](:[*]):[c](:[cH]:[*])N=[*]	-0.964	0 out of 4

17b

TOPKAT_Aerobic_Biodegradability



$C_{24}H_{19}ClN_4O_3$

Molecular Weight: 446.88566

ALogP: 3.125

Rotatable Bonds: 5

Acceptors: 4

Donors: 2

Model Prediction

Prediction: Non-Degradable

Probability: 0.082

Enrichment: 0.187

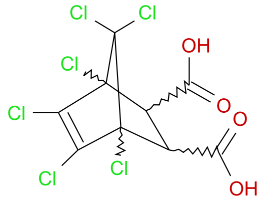
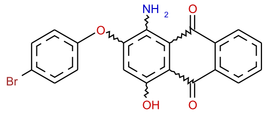
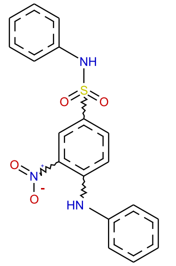
Bayesian Score: -12.860

Mahalanobis Distance: 12.676

Mahalanobis Distance p-value: 1.27e-005

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	**Bicyclo_2.2.1_hept-5-ene-2,3-dicarboxylic_acid,_1,4,5,6,7,7-hexachloro-**	**9,10-Anthracenedione,_1-amino-2-(4-bromophenoxy)-4-hydroxy-**	**Benzenesulfonamide,_3-nitro-N-phenyl-4-(phenylamino)-**
Structure			
Actual Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Predicted Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Distance	0.633	0.669	0.711
Reference	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.

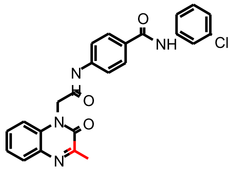
Model Applicability

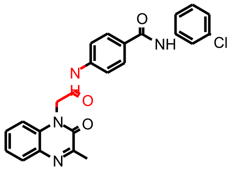
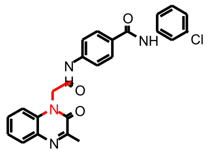
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

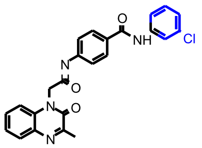
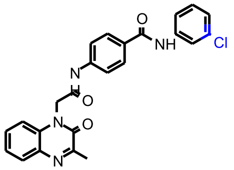
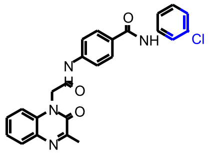
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Degradable in training set
SCFP_12	136627117	 [*]C(=[*])C	0.662	56 out of 70

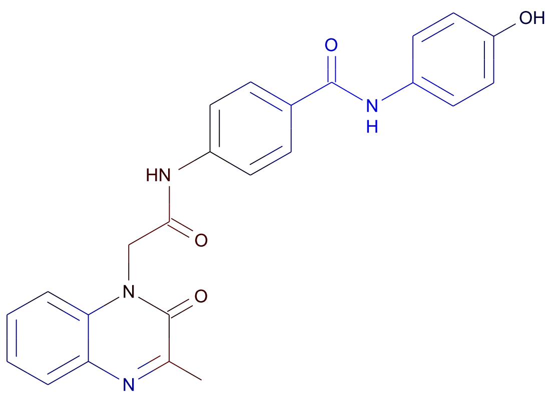
SCFP_12	1256995004	 [*]CC(=O)N[*]	0.540	23 out of 32
SCFP_12	-587539325	 [*]N([*])CC(=[*])[*]	0.504	2 out of 2

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Degradable in training set
SCFP_12	-1378360678	 Cl[c]1:[cH]:[*]:[cH]:[cH]:[cH]:1	-1.878	3 out of 62
SCFP_12	-52074512	 [*]:[c](:[*])Cl	-1.865	5 out of 93
SCFP_12	-601571304	 [*]:[cH]:[c](Cl):[cH]:[*]	-1.844	5 out of 91

17c

TOPKAT_Aerobic_Biodegradability



$C_{24}H_{20}N_4O_4$

Molecular Weight: 428.44

ALogP: 2.218

Rotatable Bonds: 5

Acceptors: 5

Donors: 3

Model Prediction

Prediction: Non-Degradable

Probability: 0.219

Enrichment: 0.503

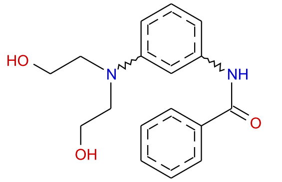
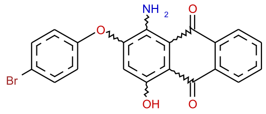
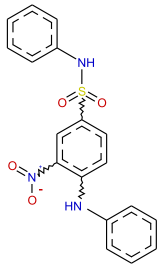
Bayesian Score: -6.527

Mahalanobis Distance: 13.089

Mahalanobis Distance p-value: 1.39e-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.

Structural Similar Compounds

Name	**Benzamide,_N-_3-_bis(2-hydroxyethyl)amino_phenyl_-**	**9,10-Anthracenedione,_1-amino-2-(4-bromophenoxy)-4-hydroxy-**	**Benzenesulfonamide,_3-nitro-N-phenyl-4-(phenylamino)-**
Structure			
Actual Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Predicted Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Distance	0.682	0.756	0.770
Reference	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.

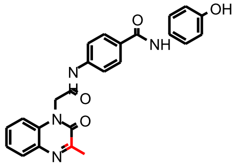
Model Applicability

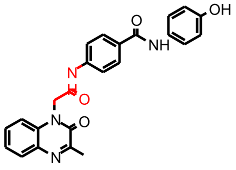
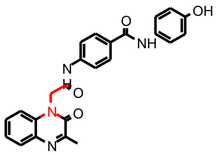
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. OPS PC19 out of range. Value: 3.3958. Training min, max, SD, explained variance: -3.6394, 3.3101, 1.187, 0.0162.

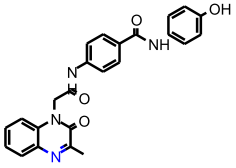
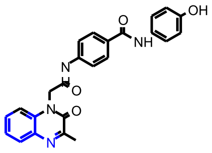
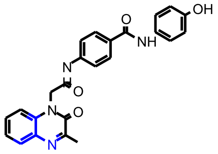
Feature Contribution

Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Degradable in training set**
SCFP_12	136627117	 [*]C(=[*])C	0.662	56 out of 70

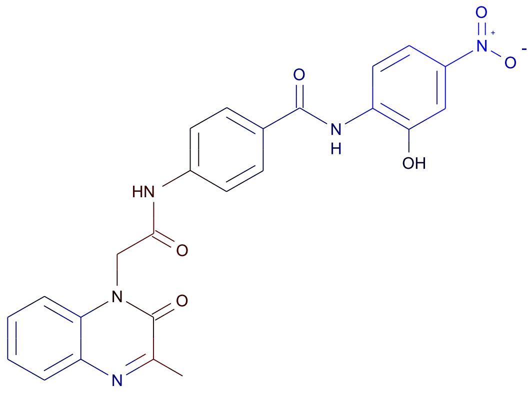
SCFP_12	1256995004	 [*]CC(=O)N[*]	0.540	23 out of 32
SCFP_12	-587539325	 [*]N([*])CC(=[*])[*]	0.504	2 out of 2

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Degradable in training set
SCFP_12	6	 [*]N=[*]	-1.018	2 out of 18
SCFP_12	-1377141613	 [*][c]1:[*]:[cH]:[cH] :[cH]:[c]:1N=[*]	-0.964	0 out of 4
SCFP_12	1851000357	 [*][c](:[*]):[c](:[cH]:[*])N=[*]	-0.964	0 out of 4

17d

TOPKAT_Aerobic_Biodegradability



$C_{24}H_{19}N_5O_6$

Molecular Weight: 473.43756

ALogP: 2.113

Rotatable Bonds: 6

Acceptors: 7

Donors: 3

Model Prediction

Prediction: Non-Degradable

Probability: 0.058

Enrichment: 0.132

Bayesian Score: -14.947

Mahalanobis Distance: 19.514

Mahalanobis Distance p-value: 1.48e-029

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.

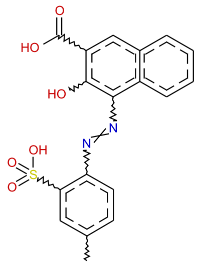
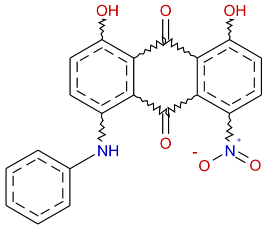
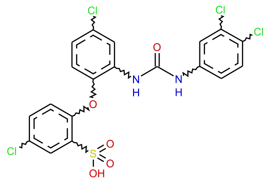
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.

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	Brilliant_Blue_6B	9,10-Anthracenedione,_1,8-dihydroxy-4-nitro-5-(phenylamino)-	Mitin_FF
Structure			
Actual Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Predicted Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Distance	0.679	0.689	0.853
Reference	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.

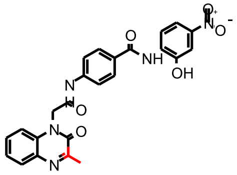
Model Applicability

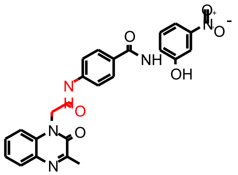
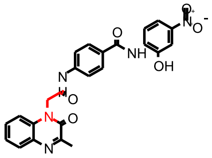
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

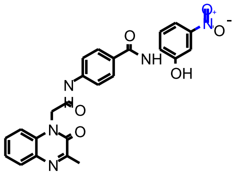
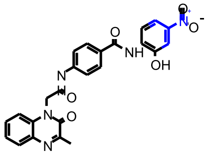
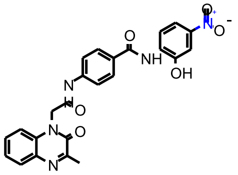
Feature Contribution

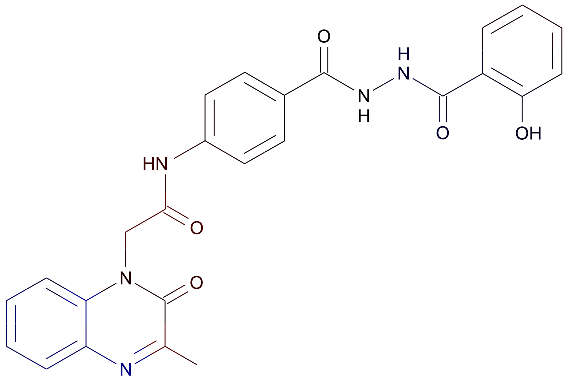
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Degradable in training set
SCFP_12	136627117	 [*]C(=[*])C	0.662	56 out of 70

SCFP_12	1256995004	 [*]CC(=O)N[*]	0.540	23 out of 32
SCFP_12	-587539325	 [*]N([*])CC(=[*])[*]	0.504	2 out of 2

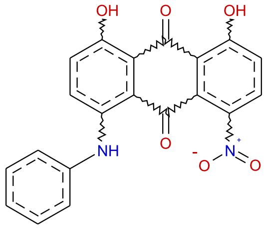
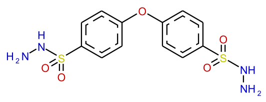
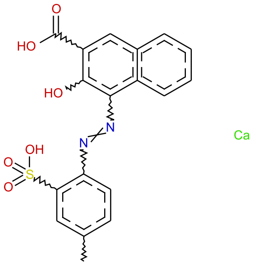
Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Degradable in training set**
SCFP_12	1311339974	 [*][N+](=O)[*]	-1.780	3 out of 56
SCFP_12	1334669481	 [*][N+](=[*])[c](:[cH]:[*]):[cH]:[*]	-1.631	10 out of 136
SCFP_12	10	 [*][N+](=[*])[*]	-1.607	11 out of 145



$C_{25}H_{21}N_5O_5$
Molecular Weight: 471.46474
ALogP: 1.475
Rotatable Bonds: 6
Acceptors: 6
Donors: 4

Structural Similar Compounds

Name	9,10-Anthracenedione,_1,8-dihydroxy-4-nitro-5-(phenylamino)-	Benzenesulfonic_acid,_4,4'-oxybis-,_dihydrazide	Brilliant_Blue_6B
Structure			
Actual Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Predicted Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Distance	0.842	0.858	0.866
Reference	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.

Model Prediction

Prediction: Non-Degradable
Probability: 0.382
Enrichment: 0.875
Bayesian Score: -2.323
Mahalanobis Distance: 13.528
Mahalanobis Distance p-value: 1.09e-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.

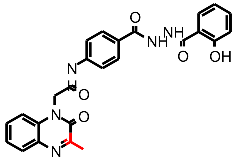
Model Applicability

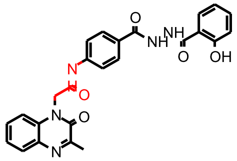
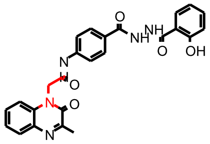
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

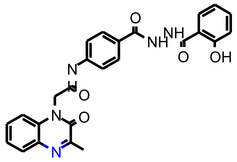
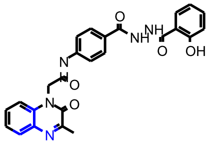
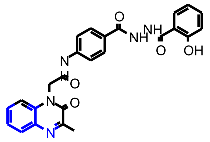
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Degradable in training set
SCFP_12	136627117	 [*]C(=[*])C	0.662	56 out of 70

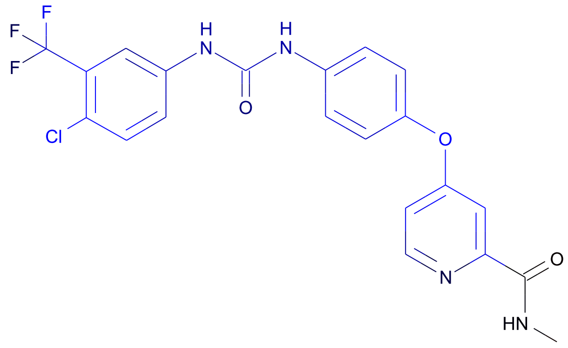
SCFP_12	1256995004	 [*]CC(=O)N[*]	0.540	23 out of 32
SCFP_12	-587539325	 [*]N([*])CC(=[*])[*]	0.504	2 out of 2

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Degradable in training set
SCFP_12	6	 [*]N=[*]	-1.018	2 out of 18
SCFP_12	1851000357	 [*][c](:[*]):[c](:[cH]:[*])N=[*]	-0.964	0 out of 4
SCFP_12	-1377141613	 [*][c]1:[*]:[cH]:[cH]:[cH]:[c]:1N=[*]	-0.964	0 out of 4

Sorafenib

TOPKAT_Aerobic_Biodegradability



$C_{21}H_{16}ClF_3N_4O_3$

Molecular Weight: 464.82494

ALogP: 4.175

Rotatable Bonds: 6

Acceptors: 4

Donors: 3

Model Prediction

Prediction: Non-Degradable

Probability: 0.013

Enrichment: 0.030

Bayesian Score: -23.288

Mahalanobis Distance: 15.358

Mahalanobis Distance p-value: 3.82e-013

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.

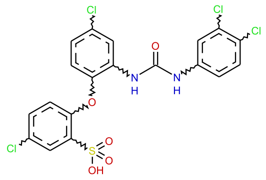
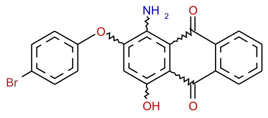
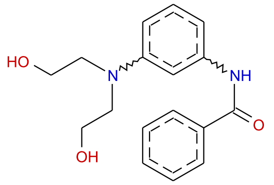
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.

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	Mitin_FF	9,10-Anthracenedione,_1-amino-2-(4-bromophenoxy)-4-hydroxy-	Benzamide,_N-_3-_bis(2-hydroxyethyl)amino_phenyl_-
Structure			
Actual Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Predicted Endpoint	Non-Degradable	Non-Degradable	Non-Degradable
Distance	0.634	0.783	0.789
Reference	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.	Environmental Toxicology & Chemistry 18(9), 1763-1768, 1999.

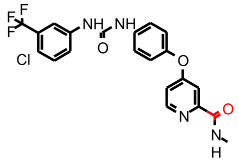
Model Applicability

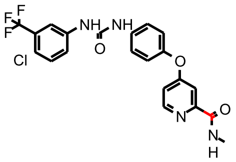
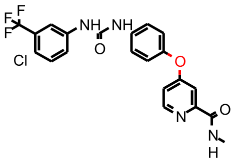
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

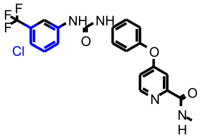
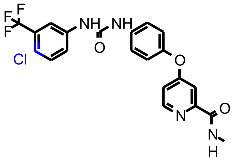
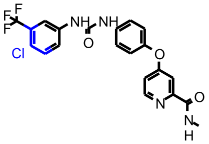
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Degradable in training set
SCFP_12	1311071855	 [*]C(=O)[*]	0.461	173 out of 268

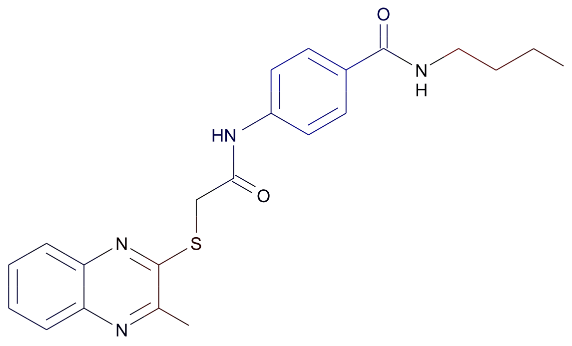
SCFP_12	1	 [*]C(=[*])[*]	0.359	194 out of 333
SCFP_12	12	 [*]O	0.264	245 out of 463

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Degradable in training set
SCFP_12	-1378360678	 Cl[c]1:[cH]:[*]:[cH]:[cH]:[cH]:1	-1.878	3 out of 62
SCFP_12	-52074512	 [*]:[c](:[*])Cl	-1.865	5 out of 93
SCFP_12	-601571304	 [*]:[cH]:[c](Cl):[cH]:[*]	-1.844	5 out of 91

15a

TOPKAT_Ocular_Irritancy_Mild_vs_Moderate_Severe



$C_{22}H_{24}N_4O_2S$

Molecular Weight: 408.51656

ALogP: 3.83

Rotatable Bonds: 8

Acceptors: 5

Donors: 2

Model Prediction

Prediction: Mild

Probability: 0.781

Enrichment: 1.133

Bayesian Score: -1.662

Mahalanobis Distance: 9.107

Mahalanobis Distance p-value: 0.454

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.

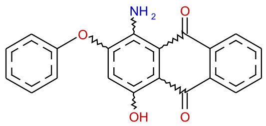
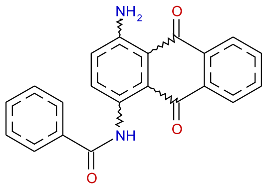
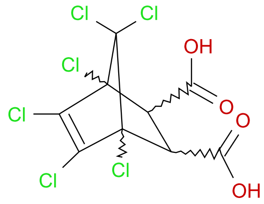
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.

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	**ANTHRAQUINONE; 1-AMINO-4-HYDROXY-2-PHENOXY-**	**1-AMINO-4-BENZOYLAMINO-ANTHRAQUINONE**	**5-NORBORNENE-2;3-DICARBOXYLIC ACID; 1;4;5;6;7;7-HEXACHLORO-**
Structure			
Actual Endpoint	Mild	Mild	Moderate_Severe
Predicted Endpoint	Mild	Mild	Moderate_Severe
Distance	0.693	0.697	0.728
Reference	28ZPAK 239;72	28ZPAK-;124;72	28ZPAK-;92;72

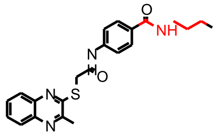
Model Applicability

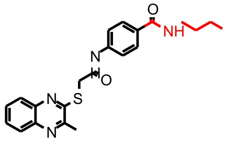
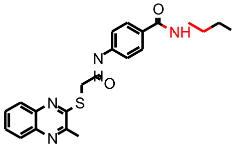
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

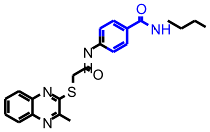
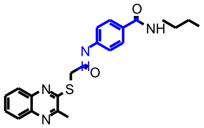
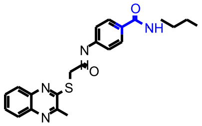
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Moderate_Severe in training set
FCFP_10	159404153	 [*]CCCNC(=[*])[*]	0.324	15 out of 16

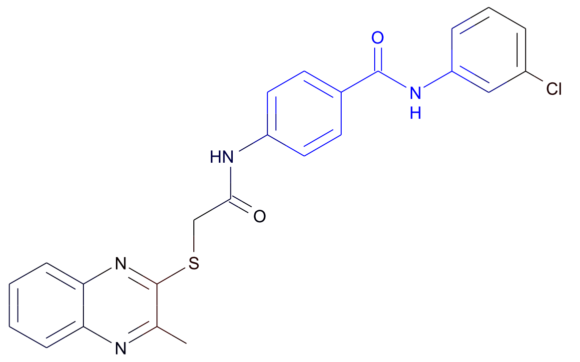
FCFP_10	-1106976657	 [*]C(=[*])NCCCC	0.317	4 out of 4
FCFP_10	-1272709286	 [*]CCN[*]	0.285	234 out of 266

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Moderate_Severe in training set**
FCFP_10	-581879738	 [*]NC(=O)[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	-1.293	0 out of 4
FCFP_10	-790336137	 [*]C(=[*])N[c]1:[cH]:[cH]:[c](:[cH]:[cH]:1)C(=[*])[*]	-0.507	0 out of 1
FCFP_10	-1549103449	 [*]NC(=O)[c](:[*]):[*]	-0.504	2 out of 6

15b

TOPKAT_Ocular_Irritancy_Mild_vs_Moderate_Severe



$C_{24}H_{19}ClN_4O_2S$
Molecular Weight: 462.95126
ALogP: 4.743
Rotatable Bonds: 6
Acceptors: 5
Donors: 2

Model Prediction

Prediction: Mild

Probability: 0.263

Enrichment: 0.382

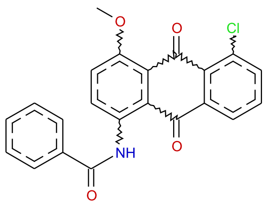
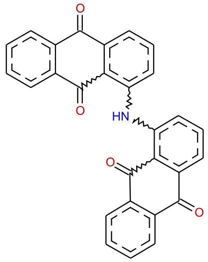
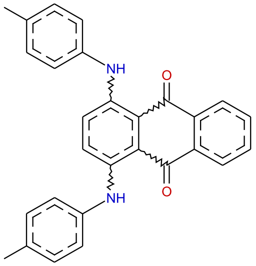
Bayesian Score: -8.589

Mahalanobis Distance: 7.465

Mahalanobis Distance p-value: 0.987

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	**1-BENZOYLAMINO-4-METHOXY-5-CHLORANTHRAQUINONE**	**ANTHRAQUINONE; 1;1'-IMINODI-**	**ANTHRAQUINONE; 1;4-BIS(p-TOLYLAMINO)-**
Structure			
Actual Endpoint	Mild	Mild	Moderate_Severe
Predicted Endpoint	Mild	Mild	Mild
Distance	0.721	0.730	0.745
Reference	28ZPAK-;90;72	28ZPAK-;125;72	28ZPAK -;124;72

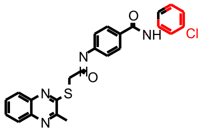
Model Applicability

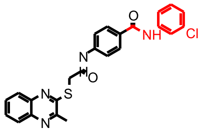
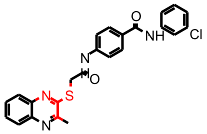
Unknown features are fingerprint features in the query molecule, but not found in the training set.

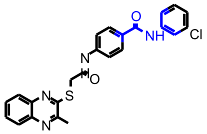
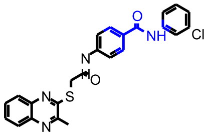
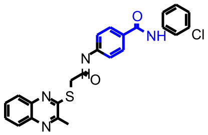
1. All properties and OPS components are within expected ranges.

Feature Contribution

Top features for positive contribution

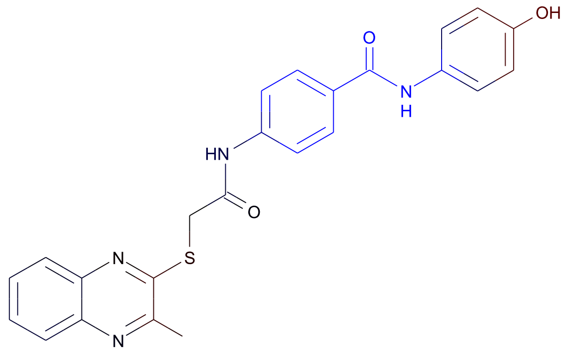
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Moderate_Severe in training set**
FCFP_10	-745491832	 [*][c]1:[*]:[cH]:[cH]:[c](Cl):[cH]:1	0.304	29 out of 32

FCFP_10	-1513589583	 [*]C(=[*])N[c]1:[cH]:[cH]:[cH]:[c](Cl):[cH]:1	0.256	2 out of 2
FCFP_10	-1410049896	 [*]S[c](:n:[*]):[c]([*]):[*]	0.256	2 out of 2

Top Features for negative contribution				
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Moderate_Severe in training set**
FCFP_10	1175232969	 [*]:[cH]:[c](NC(=O)[c](:[*]):[*]):[cH]:[*]	-1.293	0 out of 4
FCFP_10	-1925475824	 [*]:[cH]:[c](:[cH]:[*])C(=O)N[c](:[*]):[*]	-1.293	0 out of 4
FCFP_10	-581879738	 [*]NC(=O)[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	-1.293	0 out of 4

15c

TOPKAT_Ocular_Irritancy_Mild_vs_Moderate_Severe



$C_{24}H_{20}N_4O_3S$

Molecular Weight: 444.5056

ALogP: 3.836

Rotatable Bonds: 6

Acceptors: 6

Donors: 3

Model Prediction

Prediction: Mild

Probability: 0.415

Enrichment: 0.602

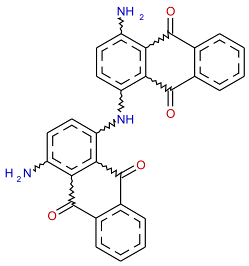
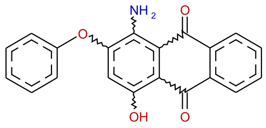
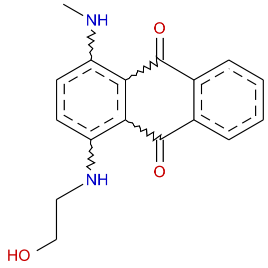
Bayesian Score: -7.039

Mahalanobis Distance: 7.499

Mahalanobis Distance p-value: 0.985

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	**4;4'-DIAMINO-1;1'-DIANTHRIMIDE**	**ANTHRAQUINONE; 1-AMINO-4-HYDROXY-2-PHENOXY-**	**ANTHRAQUINONE; 1-((2-HYDROXYETHYL)AMINO)-4-(METHYLAMINO)-**
Structure			
Actual Endpoint	Mild	Mild	Mild
Predicted Endpoint	Mild	Mild	Mild
Distance	0.609	0.793	0.798
Reference	28ZPAK-;125;72	28ZPAK 239;72	28ZPAK 245;72

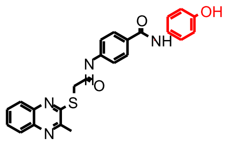
Model Applicability

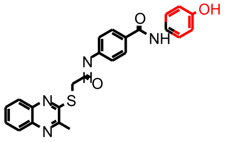
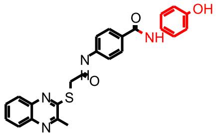
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

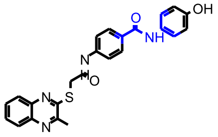
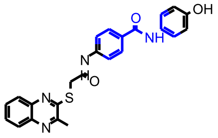
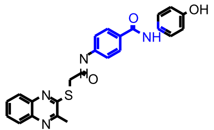
Feature Contribution

Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Moderate_Severe in training set**
FCFP_10	-1066794953	 [*][c]1:[cH]:[cH]:[c](O):[cH]:[cH]:1	0.378	13 out of 13

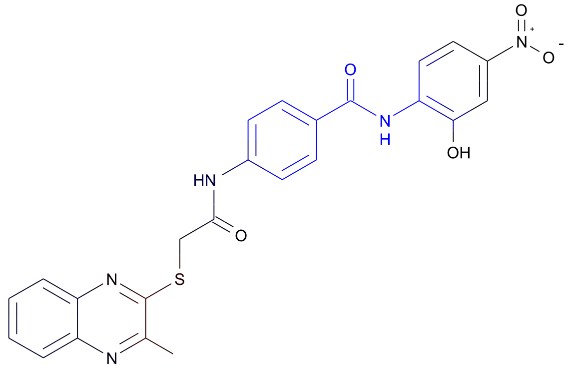
FCFP_10	-158888774	 O[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	0.356	24 out of 25
FCFP_10	-1849978862	 [*]N[c]1:[cH]:[cH]:[c](O):[cH]:[cH]:1	0.256	2 out of 2

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Moderate_Severe in training set
FCFP_10	1175232969	 [*]:[cH]:[c](NC(=O)[c](:[*]):[*]):[cH]:[*]	-1.293	0 out of 4
FCFP_10	241406177	 [*]:[cH]:[c](NC(=O)[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1):[cH]:[*]	-1.293	0 out of 4
FCFP_10	384221478	 [*][c]1:[cH]:[cH]:[c](:[cH]:[cH]:1)C(=O)N[c](:[*]):[*]	-1.293	0 out of 4

15d

TOPKAT_Ocular_Irritancy_Mild_vs_Moderate_Severe



$C_{24}H_{19}N_5O_5S$

Molecular Weight: 489.50316

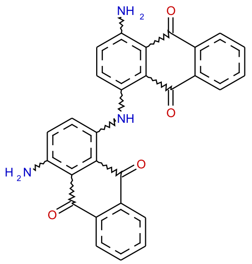
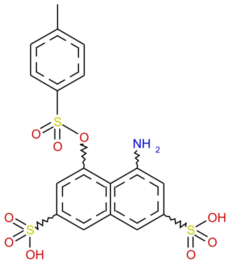
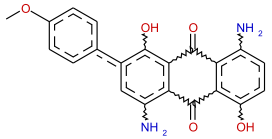
ALogP: 3.731

Rotatable Bonds: 7

Acceptors: 8

Donors: 3

Structural Similar Compounds

Name	4;4'-DIAMINO-1;1'-DIANTHRIMIDE	2;7-NAPHTHALENE DISULFONIC ACID;4-AMINO-5-HYDROXY-;P-TOLUENE SULFONATE (ESTER)	ANTHRAQUINONE; 1;5-DIAMINO-4;8-DIHYDROXY-3-(p-METHOXYPHENYL)-
Structure			
Actual Endpoint	Mild	Moderate_Severe	Mild
Predicted Endpoint	Mild	Moderate_Severe	Mild
Distance	0.654	0.772	0.835
Reference	28ZPAK-;125;72	28ZPAK-;194;72	28ZPAK 245;72

Model Prediction

Prediction: Mild

Probability: 0.251

Enrichment: 0.364

Bayesian Score: -8.719

Mahalanobis Distance: 7.566

Mahalanobis Distance p-value: 0.981

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.

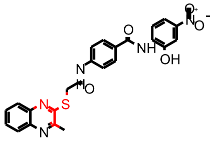
Model Applicability

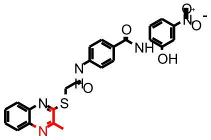
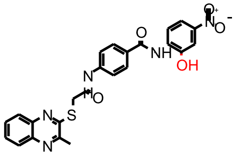
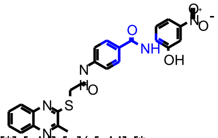
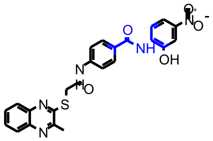
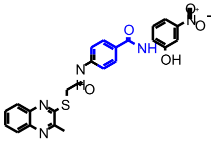
Unknown features are fingerprint features in the query molecule, but not found in the training set.

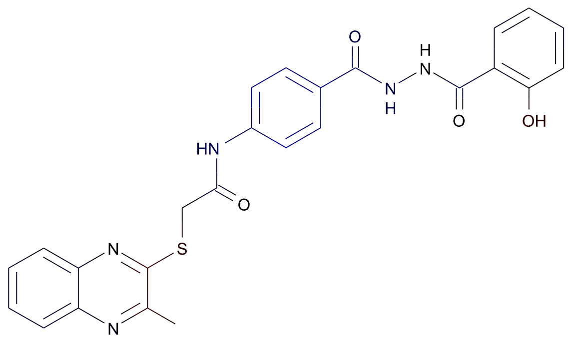
1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_2 feature: -828984032: [*][N+](=[*])[c](:c:[*]):c:[*]
3. Unknown FCFP_2 feature: -1338588315: [*]:[c](:[*])[N+](=O)[O-]
4. Unknown FCFP_2 feature: 1872392852: [*][N+](=O)[*]
5. Unknown FCFP_2 feature: 260476081: [*][N+](=[*])[O-]

Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Moderate_Severe in training set
FCFP_10	-1410049896	 [*]S[c](:n:[*]):[c]([*]):[*]	0.256	2 out of 2

FCFP_10	-1539132615	 [*][c](:[*]):[c](C):n:[*]	0.224	11 out of 13
FCFP_10	7	 [*]O	0.219	117 out of 142
Top Features for negative contribution				
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Moderate_Severe in training set**
FCFP_10	-1925475824	 [*]:[cH]:[c](:[cH]:[*])C(=O)N[c](:[*]):[*]	-1.293	0 out of 4
FCFP_10	1175232969	 [*]:[cH]:[c](NC(=O)[c](:[*]):[*]):[cH]:[*]	-1.293	0 out of 4
FCFP_10	-581879738	 [*]NC(=O)[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	-1.293	0 out of 4



$C_{25}H_{21}N_5O_4S$

Molecular Weight: 487.53034

ALogP: 3.093

Rotatable Bonds: 7

Acceptors: 7

Donors: 4

Model Prediction

Prediction: Mild

Probability: 0.725

Enrichment: 1.053

Bayesian Score: -3.038

Mahalanobis Distance: 8.306

Mahalanobis Distance p-value: 0.84

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.

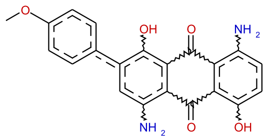
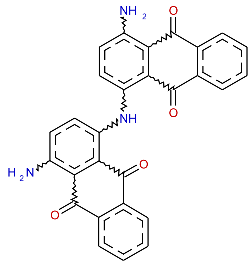
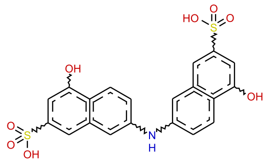
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.

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	**ANTHRAQUINONE; 1;5-DIAMINO-4;8-DIHYDROXY-3-(p-METHOXYPHENYL)-**	**4;4'-DIAMINO-1;1'-DIANTHRIMIDE**	**2-Naphthalenesulfonic acid; 5;6'-iminobis(1-hydroxy-**
Structure			
Actual Endpoint	Mild	Mild	Mild
Predicted Endpoint	Mild	Mild	Mild
Distance	0.710	0.741	0.809
Reference	28ZPAK 245;72	28ZPAK-;125;72	Prehled Prumyslove Toxikologie; Organicke Latky; Marhold; J. pp 1065;86

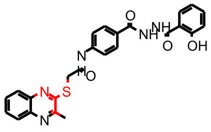
Model Applicability

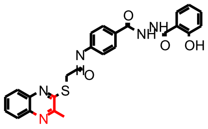
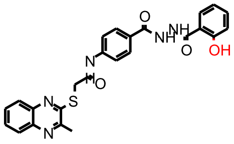
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_2 feature: -885461129: [*]NNC(=[*])[*]

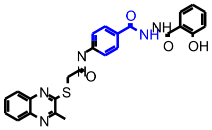
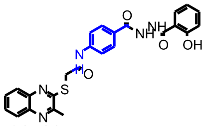
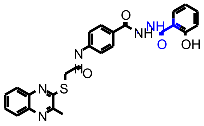
Feature Contribution

Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Moderate_Severe in training set**
FCFP_10	-1410049896	 [*]S[c](:n:[*]):[c]([*]):[*]	0.256	2 out of 2

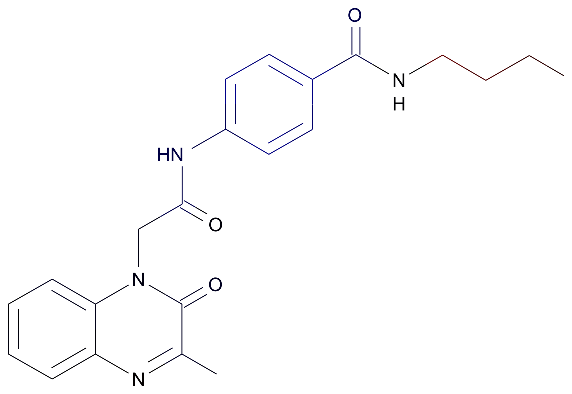
FCFP_10	-1539132615	 [*][c](:[*]):[c](C):n:[*]	0.224	11 out of 13
FCFP_10	7	 [*]O	0.219	117 out of 142

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Moderate_Severe in training set**
FCFP_10	-581879738	 [*]NC(=O)[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	-1.293	0 out of 4
FCFP_10	-790336137	 [*]C(=[*])N[c]1:[cH]:[cH]:[c](:[cH]:[cH]:1)C(=[*])[*]	-0.507	0 out of 1
FCFP_10	-1549103449	 [*]NC(=O)[c](:[*]):[*]	-0.504	2 out of 6

17a

TOPKAT_Ocular_Irritancy_Mild_vs_Moderate_Severe



$C_{22}H_{24}N_4O_3$

Molecular Weight: 392.45096

ALogP: 2.212

Rotatable Bonds: 7

Acceptors: 4

Donors: 2

Model Prediction

Prediction: Mild

Probability: 0.747

Enrichment: 1.085

Bayesian Score: -2.561

Mahalanobis Distance: 12.173

Mahalanobis Distance p-value: 2.7e-005

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.

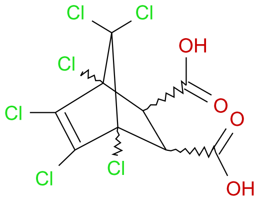
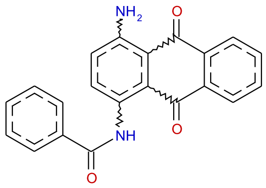
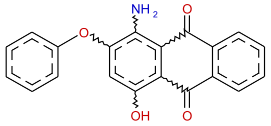
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.

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	**5-NORBORNENE-2;3-DICARBOXYLIC ACID; 1;4;5;6;7;7-HEXACHLORO-**	**1-AMINO-4-BENZOYLAMINO-ANTHRAQUINONE**	**ANTHRAQUINONE; 1-AMINO-4-HYDROXY-2-PHENOXY-**
Structure			
Actual Endpoint	Moderate_Severe	Mild	Mild
Predicted Endpoint	Moderate_Severe	Mild	Mild
Distance	0.640	0.641	0.706
Reference	28ZPAK-;92;72	28ZPAK-;124;72	28ZPAK 239;72

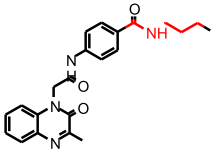
Model Applicability

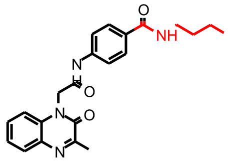
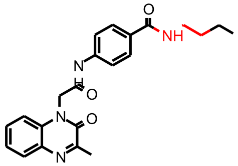
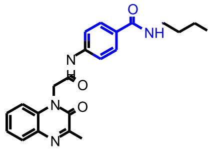
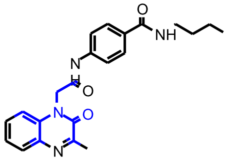
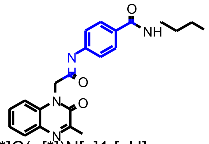
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

Feature Contribution

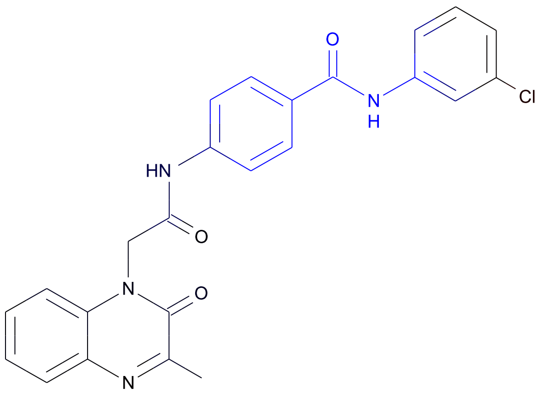
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Moderate_Severe in training set
FCFP_10	159404153	 [*]CCCNC(=[*])[*]	0.324	15 out of 16

<table>
<tr><td>FCFP_10</td><td>-1106976657</td><td>
[*]C(=[*])NCCCC</td><td>0.317</td><td>4 out of 4</td></tr>
<tr><td>FCFP_10</td><td>-1272709286</td><td>
[*]CCN[*]</td><td>0.285</td><td>234 out of 266</td></tr>
<tr><th colspan="5">Top Features for negative contribution</th></tr>
<tr><th>Fingerprint</th><th>Bit/Smiles</th><th>Feature Structure</th><th>Score</th><th>Moderate_Severe in training set</th></tr>
<tr><td>FCFP_10</td><td>-581879738</td><td>
[*]NC(=O)[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1</td><td>-1.293</td><td>0 out of 4</td></tr>
<tr><td>FCFP_10</td><td>-70433048</td><td>
[*]C(=[*])CN1C(=O)C(=[*][c](:[*]):[c]1:[cH]:[*])[*]</td><td>-0.507</td><td>0 out of 1</td></tr>
<tr><td>FCFP_10</td><td>-790336137</td><td>
[*]C(=[*])N[c]1:[cH]:[cH]:[c](:[cH]:[cH]:1)C(=[*])[*]</td><td>-0.507</td><td>0 out of 1</td></tr>
</table>

17b

TOPKAT_Ocular_Irritancy_Mild_vs_Moderate_Severe



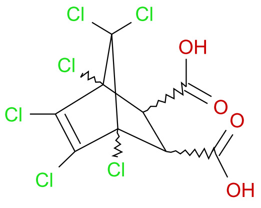
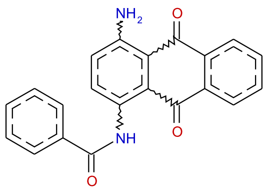
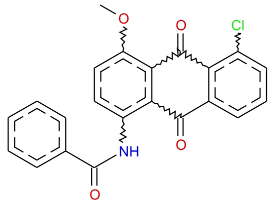
$C_{24}H_{19}ClN_4O_3$
Molecular Weight: 446.88566
ALogP: 3.125
Rotatable Bonds: 5
Acceptors: 4
Donors: 2

Model Prediction

Prediction: Mild
Probability: 0.188
Enrichment: 0.273
Bayesian Score: -9.460
Mahalanobis Distance: 11.685
Mahalanobis Distance p-value: 0.000287

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	**5-NORBORNENE-2;3-DICARBOXYLIC ACID; 1;4;5;6;7;7-HEXACHLORO-**	**1-AMINO-4-BENZOYLAMINO-ANTHRAQUINONE**	**1-BENZOYLAMINO-4-METHOXY-5-CHLORANTHRAQUINONE**
Structure			
Actual Endpoint	Moderate_Severe	Mild	Mild
Predicted Endpoint	Moderate_Severe	Mild	Mild
Distance	0.640	0.650	0.693
Reference	28ZPAK-;92;72	28ZPAK-;124;72	28ZPAK-;90;72

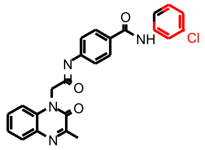
Model Applicability

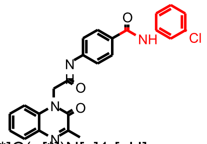
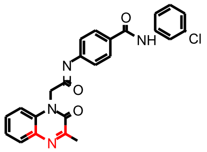
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

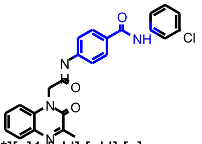
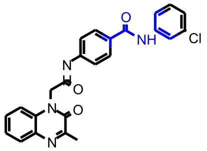
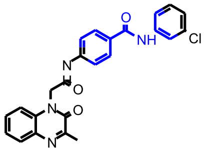
Feature Contribution

Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Moderate_Severe in training set**
FCFP_10	-745491832	 [*][c]1:[*]:[cH]:[cH]:[c](Cl):[cH]:1	0.304	29 out of 32

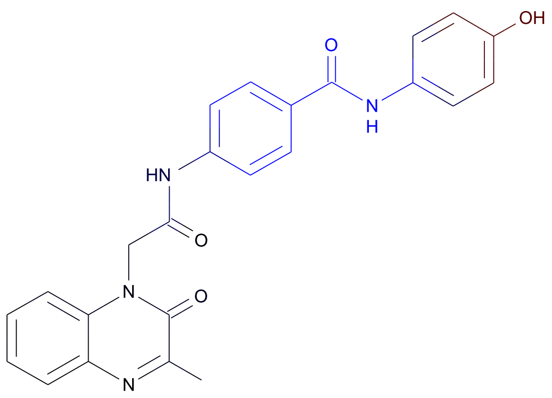
FCFP_10	-1513589583	 [*]C(=[*])N[c]1:[cH]:[cH]:[cH]:[c](Cl):[cH]:1	0.256	2 out of 2
FCFP_10	580453787	 [*]C(=N[c](:[*]):[*])[*]	0.256	2 out of 2

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Moderate_Severe in training set
FCFP_10	384221478	 [*][c]1:[cH]:[cH]:[c](:[cH]:[cH]:1)C(=O)N[c](:[*]):[*]	-1.293	0 out of 4
FCFP_10	1175232969	 [*]:[cH]:[c](NC(=O)[c](:[*]):[*]):[cH]:[*]	-1.293	0 out of 4
FCFP_10	241406177	 [*]:[cH]:[c](NC(=O)[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1):[cH]:[*]	-1.293	0 out of 4

17c

TOPKAT_Ocular_Irritancy_Mild_vs_Moderate_Severe



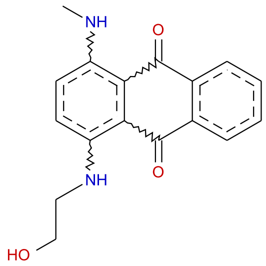
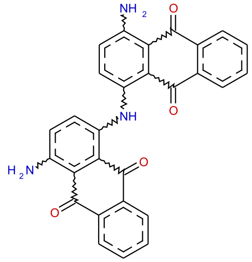
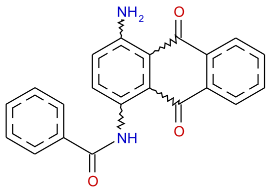
$C_{24}H_{20}N_4O_4$
Molecular Weight: 428.44
ALogP: 2.218
Rotatable Bonds: 5
Acceptors: 5
Donors: 3

Model Prediction

Prediction: Mild
Probability: 0.327
Enrichment: 0.474
Bayesian Score: -7.924
Mahalanobis Distance: 11.804
Mahalanobis Distance p-value: 0.000165

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	**ANTHRAQUINONE; 1-((2-HYDROXYETHYL)AMINO)-4-(METHYLAMINO)-**	**4;4'-DIAMINO-1;1'-DIANTHRIMIDE**	**1-AMINO-4-BENZOYLAMINO-ANTHRAQUINONE**
Structure			
Actual Endpoint	Mild	Mild	Mild
Predicted Endpoint	Mild	Mild	Mild
Distance	0.684	0.734	0.736
Reference	28ZPAK 245;72	28ZPAK-;125;72	28ZPAK-;124;72

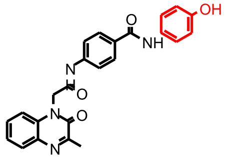
Model Applicability

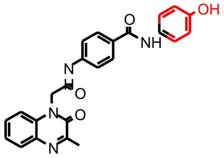
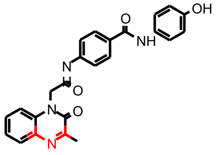
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

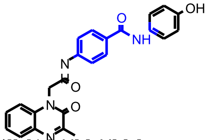
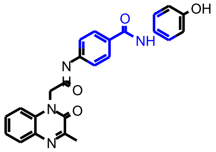
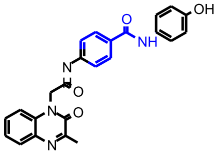
Feature Contribution

Top features for positive contribution

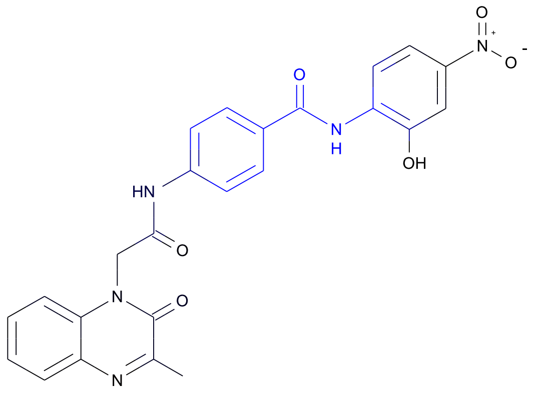
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Moderate_Severe in training set**
FCFP_10	-1066794953	 [*][c]1:[cH]:[cH]:[c](O):[cH]:[cH]:1	0.378	13 out of 13

FCFP_10	-158888774	 O[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	0.356	24 out of 25
FCFP_10	580453787	 [*]C(=N[c](:[*]):[*])[*]	0.256	2 out of 2

Top Features for negative contribution

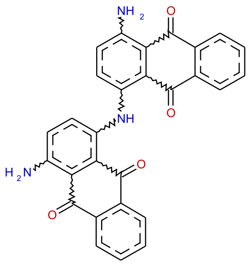
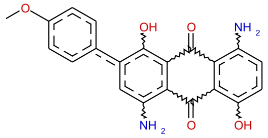
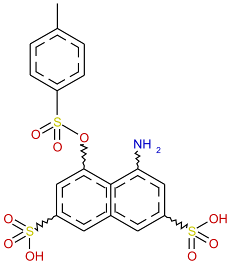
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Moderate_Severe in training set**
FCFP_10	384221478	 [*][c]1:[cH]:[cH]:[c](:[cH]:[cH]:1)C(=O)N[c](:[*]):[*]	-1.293	0 out of 4
FCFP_10	241406177	 [*]:[cH]:[c](NC(=O)[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1):[cH]:[*]	-1.293	0 out of 4
FCFP_10	-581879738	 [*]NC(=O)[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	-1.293	0 out of 4

17d TOPKAT_Ocular_Irritancy_Mild_vs_Moderate_Severe



$C_{24}H_{19}N_5O_6$
Molecular Weight: 473.43756
ALogP: 2.113
Rotatable Bonds: 6
Acceptors: 7
Donors: 3

Structural Similar Compounds

Name	4;4'-DIAMINO-1;1'-DIANTHRIMIDE	ANTHRAQUINONE; 1;5-DIAMINO-4;8-DIHYDROXY-3-(p-METHOXYPHENYL)-	2;7-NAPHTHALENE DISULFONIC ACID;4-AMINO-5-HYDROXY-;P-TOLUENE SULFONATE (ESTER)
Structure			
Actual Endpoint	Mild	Mild	Moderate_Severe
Predicted Endpoint	Mild	Mild	Moderate_Severe
Distance	0.674	0.767	0.826
Reference	28ZPAK-;125;72	28ZPAK 245;72	28ZPAK-;194;72

Model Prediction

Prediction: Mild

Probability: 0.178

Enrichment: 0.258

Bayesian Score: -9.595

Mahalanobis Distance: 11.681

Mahalanobis Distance p-value: 0.000291

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.

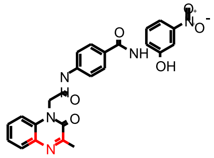
Model Applicability

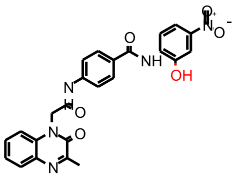
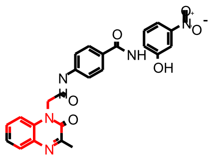
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_2 feature: -828984032: [*][N+](=[*])[c](:c:[*]):c:[*]
3. Unknown FCFP_2 feature: -1338588315: [*]:[c](:[*])[N+](=O)[O-]
4. Unknown FCFP_2 feature: 1872392852: [*][N+](=O)[*]
5. Unknown FCFP_2 feature: 260476081: [*][N+](=[*])[O-]

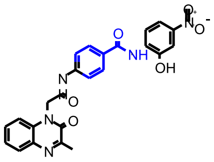
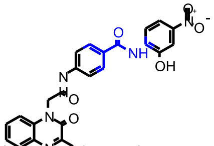
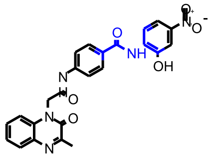
Feature Contribution

Top features for positive contribution

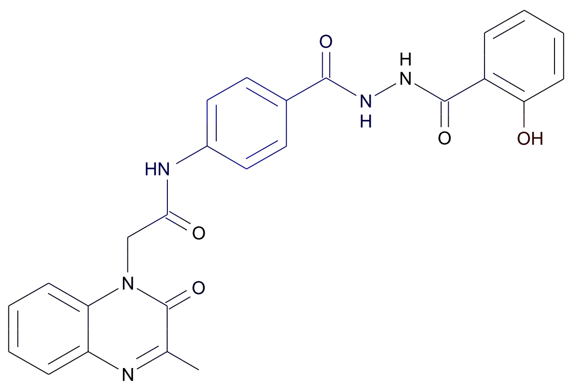
Fingerprint	Bit/Smiles	Feature Structure	Score	Moderate_Severe in training set
FCFP_10	580453787	 [*]C(=N[c](:[*]):[*]) [*]	0.256	2 out of 2

FCFP_10	7	 [*]O	0.219	117 out of 142
FCFP_10	-2007573814	 [*]CN1C(=[*])[*]=N[c]2:[cH]:[*]:[cH]:[cH]:[c]1:2	0.186	1 out of 1

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Moderate_Severe in training set**
FCFP_10	-581879738	 [*]NC(=O)[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	-1.293	0 out of 4
FCFP_10	-1925475824	 [*]:[cH]:[c](:[cH]:[*])C(=O)N[c](:[*]):[*]	-1.293	0 out of 4
FCFP_10	1175232969	 [*]:[cH]:[c](NC(=O)[c](:[*]):[*]):[cH]:[*]	-1.293	0 out of 4

TOPKAT_Ocular_Irritancy_Mild_vs_Moderate_Severe



$C_{25}H_{21}N_5O_5$
Molecular Weight: 471.46474
ALogP: 1.475
Rotatable Bonds: 6
Acceptors: 6
Donors: 4

Model Prediction

Prediction: Mild

Probability: 0.694

Enrichment: 1.008

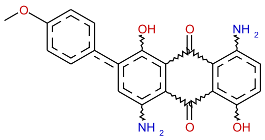
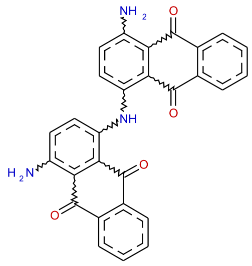
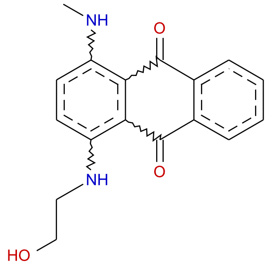
Bayesian Score: -3.623

Mahalanobis Distance: 12.635

Mahalanobis Distance p-value: 2.21e-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.

Structural Similar Compounds

Name	ANTHRAQUINONE; 1;5-DIAMINO-4;8-DIHYDROXY-3-(p-METHOXYPHENYL)-	4;4'-DIAMINO-1;1'-DIANTHRIMIDE	ANTHRAQUINONE; 1-((2-HYDROXYETHYL)AMINO)-4-(METHYLAMINO)-
Structure			
Actual Endpoint	Mild	Mild	Mild
Predicted Endpoint	Mild	Mild	Mild
Distance	0.704	0.818	0.850
Reference	28ZPAK 245;72	28ZPAK-;125;72	28ZPAK 245;72

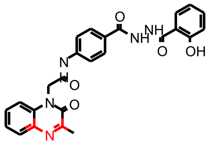
Model Applicability

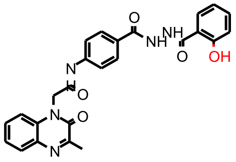
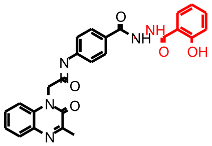
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_2 feature: -885461129: [*]NNC(=[*])[*]

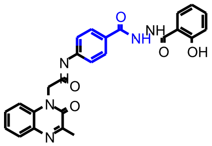
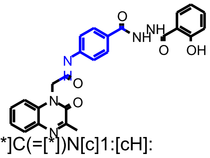
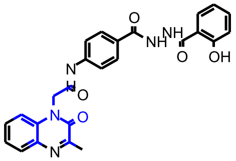
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Moderate_Severe in training set
FCFP_10	580453787	 [*]C(=N[c](:[*]):[*]) [*]	0.256	2 out of 2

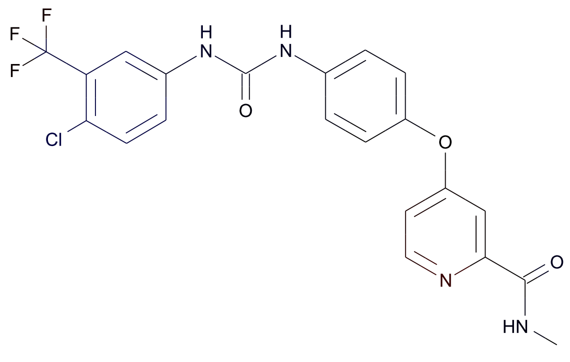
FCFP_10	7	 [*]O	0.219	117 out of 142
FCFP_10	1396546223	 [*]NC(=O)[c]1:[cH]:[cH]:[cH]:[cH]:[c]:1O	0.186	1 out of 1

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Moderate_Severe in training set
FCFP_10	-581879738	 [*]NC(=O)[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	-1.293	0 out of 4
FCFP_10	-790336137	 [*]C(=[*])N[c]1:[cH]:[cH]:[c](:[cH]:[cH]:1)C(=[*])[*]	-0.507	0 out of 1
FCFP_10	-70433048	 [*]C(=[*])CN1C(=O)C(=[*][c](:[*]):[c]1:[cH]:[*])[*]	-0.507	0 out of 1

Sorafenib

TOPKAT_Ocular_Irritancy_Mild_vs_Moderate_Severe



$C_{21}H_{16}ClF_3N_4O_3$

Molecular Weight: 464.82494

ALogP: 4.175

Rotatable Bonds: 6

Acceptors: 4

Donors: 3

Model Prediction

Prediction: Mild

Probability: 0.776

Enrichment: 1.127

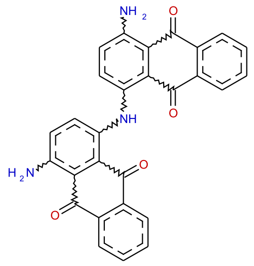
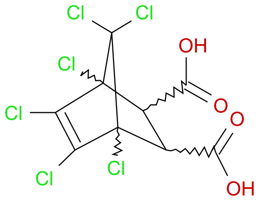
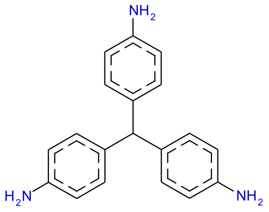
Bayesian Score: -1.801

Mahalanobis Distance: 8.954

Mahalanobis Distance p-value: 0.537

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	**4;4'-DIAMINO-1;1'-DIANTHRIMIDE**	**5-NORBORNENE-2;3-DICARBOXYLIC ACID; 1;4;5;6;7;7-HEXACHLORO-**	**METHANE;TRIS(4-AMINOPHENYL)-**
Structure			
Actual Endpoint	Mild	Moderate_Severe	Moderate_Severe
Predicted Endpoint	Mild	Moderate_Severe	Moderate_Severe
Distance	0.799	0.816	0.827
Reference	28ZPAK-;125;72	28ZPAK-;92;72	28ZPAK-;73;72

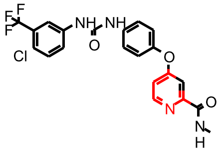
Model Applicability

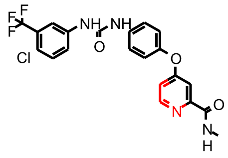
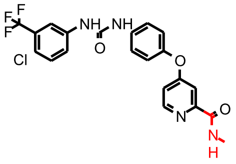
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

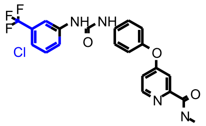
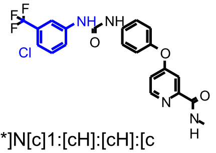
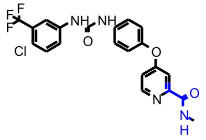
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Moderate_Severe in training set
FCFP_10	-1695756380	 [*][c]1:[*]:[c]([*]):n:[cH]:[cH]:1	0.285	10 out of 11

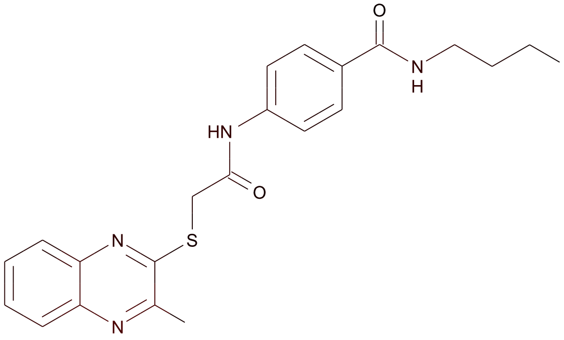
FCFP_10	-124655670	 [*]:[cH]:[cH]:n:[*]	0.259	14 out of 16
FCFP_10	-885550502	 [*]CNC(=[*])[*]	0.239	54 out of 64

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Moderate_Severe in training set**
FCFP_10	2104062943	 [*]C([*])([*])[c]1:[cH]:[*]:[cH]:[cH]:[c]:1Cl	-0.745	7 out of 24
FCFP_10	-174293376	 [*]N[c]1:[cH]:[cH]:[c](Cl):[c](:[cH]:1)C([*])([*])[*]	-0.507	0 out of 1
FCFP_10	-1549103449	 [*]NC(=O)[c](:[*]):[*]	-0.504	2 out of 6

15a

TOPKAT_Ocular_Irritancy_None_vs_Irritant



$C_{22}H_{24}N_4O_2S$

Molecular Weight: 408.51656

ALogP: 3.83

Rotatable Bonds: 8

Acceptors: 5

Donors: 2

Model Prediction

Prediction: Irritant

Probability: 1.000

Enrichment: 1.176

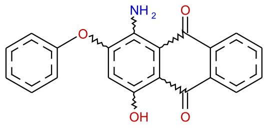
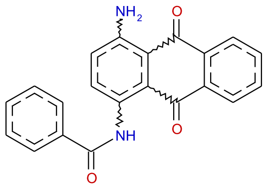
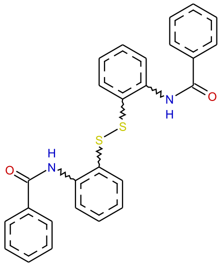
Bayesian Score: 3.672

Mahalanobis Distance: 8.675

Mahalanobis Distance p-value: 0.684

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	**ANTHRAQUINONE; 1-AMINO-4-HYDROXY-2-PHENOXY-**	**1-AMINO-4-BENZOYLAMINO-ANTHRAQUINONE**	**BENZANILIDE;2';2'''-DITHIOBIS-**
Structure			
Actual Endpoint	Irritant	Irritant	Non-Irritant
Predicted Endpoint	Irritant	Irritant	Non-Irritant
Distance	0.679	0.681	0.686
Reference	28ZPAK 239;72	28ZPAK-;124;72	28ZPAK-;173;72

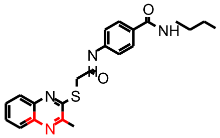
Model Applicability

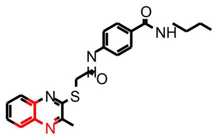
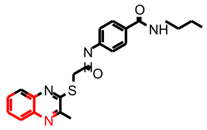
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

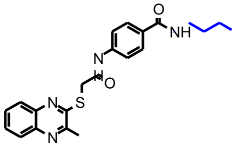
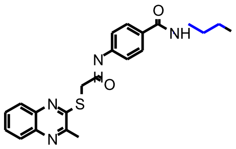
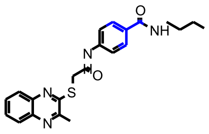
Feature Contribution

Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Irritant in training set**
FCFP_12	1747237384	 [*][c](:[*]):n:[c](:[*]):[*]	0.208	44 out of 44

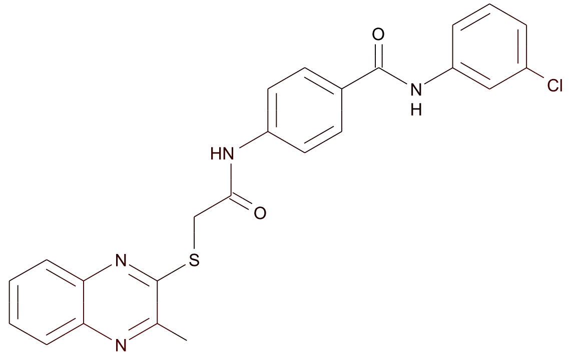
FCFP_12	178336375	 [*]:[cH]:[c](:n:[*]):[c](:[*]):[*]	0.202	19 out of 19
FCFP_12	713358128	 [*]:n:[c]1:[cH]:[cH]:[cH]:[*]:[c]:1:[*]	0.200	17 out of 17

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	-1577600103	 [*]CCCC	-0.090	217 out of 294
FCFP_12	-1272798659	 [*]CCC[*]	0.000	517 out of 643
FCFP_12	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	0.000	319 out of 382

15b

TOPKAT_Ocular_Irritancy_None_vs_Irritant



$C_{24}H_{19}ClN_4O_2S$

Molecular Weight: 462.95126

ALogP: 4.743

Rotatable Bonds: 6

Acceptors: 5

Donors: 2

Model Prediction

Prediction: Irritant

Probability: 1.000

Enrichment: 1.176

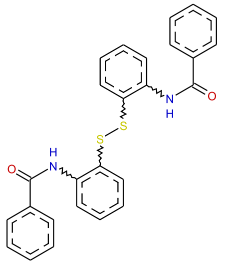
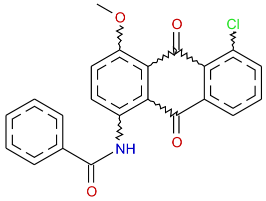
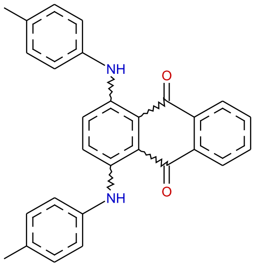
Bayesian Score: 4.245

Mahalanobis Distance: 6.333

Mahalanobis Distance p-value: 1

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	**BENZANILIDE;2';2'''-DITHIOBIS-**	**1-BENZOYLAMINO-4-METHOXY-5-CHLORANTHRAQUINONE**	**ANTHRAQUINONE; 1;4-BIS(p-TOLYLAMINO)-**
Structure			
Actual Endpoint	Non-Irritant	Irritant	Irritant
Predicted Endpoint	Non-Irritant	Irritant	Non-Irritant
Distance	0.587	0.721	0.727
Reference	28ZPAK-;173;72	28ZPAK-;90;72	28ZPAK -;124;72

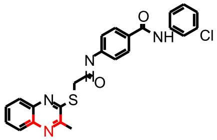
Model Applicability

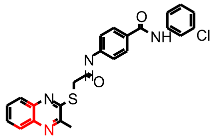
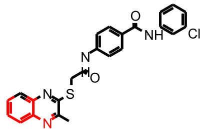
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

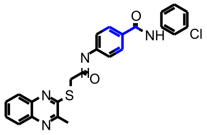
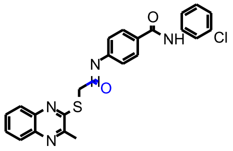
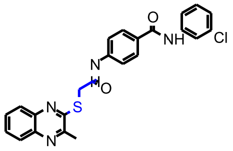
Feature Contribution

Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Irritant in training set**
FCFP_12	1747237384	 [*][c](:[*]):n:[c](:[*]):[*]	0.208	44 out of 44

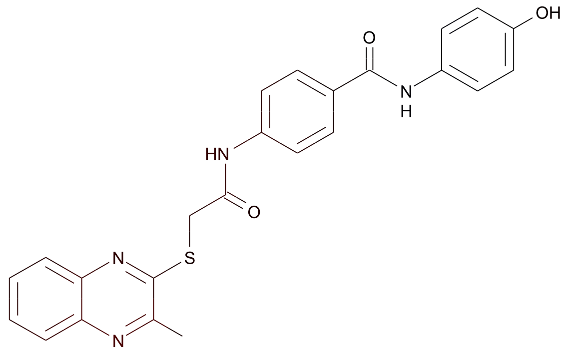
FCFP_12	178336375	 [*]:[cH]:[c](:n:[*]):[c](:[*]):[*]	0.202	19 out of 19
FCFP_12	713358128	 [*]:n:[c]1:[cH]:[cH]:[cH]:[*]:[c]:1:[*]	0.200	17 out of 17

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	0.000	319 out of 382
FCFP_12	1872154524	 [*]C(=O)[*]	0.000	563 out of 690
FCFP_12	-1272768868	 [*]SCC(=[*])[*]	0.000	396 out of 514

15c

TOPKAT_Ocular_Irritancy_None_vs_Irritant



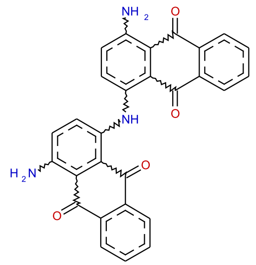
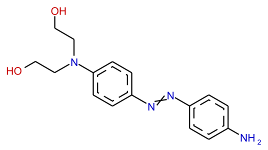
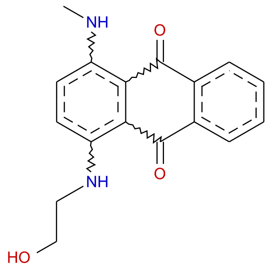
$C_{24}H_{20}N_4O_3S$
Molecular Weight: 444.5056
ALogP: 3.836
Rotatable Bonds: 6
Acceptors: 6
Donors: 3

Model Prediction

Prediction: Irritant
Probability: 1.000
Enrichment: 1.176
Bayesian Score: 3.311
Mahalanobis Distance: 6.595
Mahalanobis Distance p-value: 1

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	4;4'-DIAMINO-1;1'-DIANTHRIMIDE	Disperse Black 9	ANTHRAQUINONE; 1-((2-HYDROXYETHYL)AMINO)-4-(METHYLAMINO)-
Structure			
Actual Endpoint	Irritant	Non-Irritant	Irritant
Predicted Endpoint	Irritant	Non-Irritant	Irritant
Distance	0.603	0.737	0.776
Reference	28ZPAK-;125;72	J. Am. Coll. Toxicol. 5(3):205;1986	28ZPAK 245;72

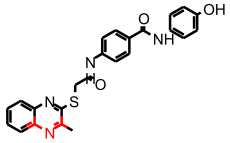
Model Applicability

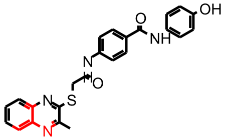
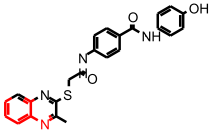
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

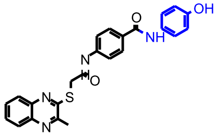
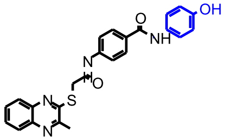
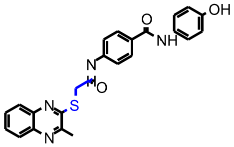
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	1747237384	 [*][c](:[*]):n:[c](:[*]):[*]	0.208	44 out of 44

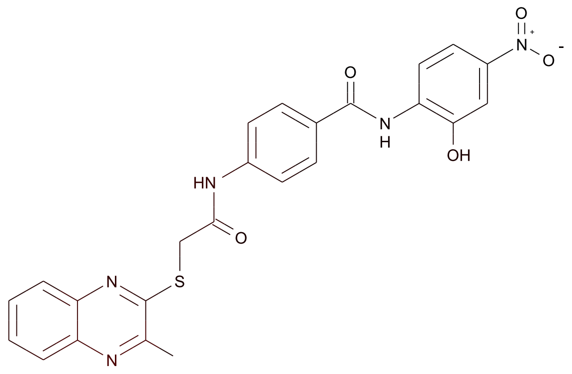
FCFP_12	178336375	 [*]:[cH]:[c](:n:[*]):[c](:[*]):[*]	0.202	19 out of 19
FCFP_12	713358128	 [*]:n:[c]1:[cH]:[cH]:[cH]:[*]:[c]:1:[*]	0.200	17 out of 17

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	-1849978862	 [*]N[c]1:[cH]:[cH]:[c](O):[cH]:[cH]:1	-0.132	2 out of 3
FCFP_12	-1066794953	 [*][c]1:[cH]:[cH]:[c](O):[cH]:[cH]:1	-0.051	13 out of 17
FCFP_12	-1272768868	 [*]SCC(=[*])[*]	0.000	396 out of 514

15d

TOPKAT_Ocular_Irritancy_None_vs_Irritant



$C_{24}H_{19}N_5O_5S$

Molecular Weight: 489.50316

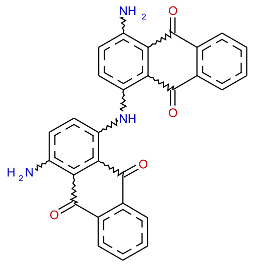
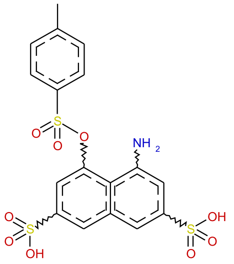
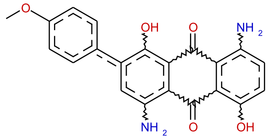
ALogP: 3.731

Rotatable Bonds: 7

Acceptors: 8

Donors: 3

Structural Similar Compounds

Name	4;4'-DIAMINO-1;1'-DIANTHRIMIDE	2;7-NAPHTHALENE DISULFONIC ACID;4-AMINO-5-HYDROXY-;P-TOLUENE SULFONATE (ESTER)	ANTHRAQUINONE; 1;5-DIAMINO-4;8-DIHYDROXY-3-(p-METHOXYPHENYL)-
Structure			
Actual Endpoint	Irritant	Irritant	Irritant
Predicted Endpoint	Irritant	Irritant	Irritant
Distance	0.646	0.763	0.821
Reference	28ZPAK-;125;72	28ZPAK-;194;72	28ZPAK 245;72

Model Prediction

Prediction: Irritant

Probability: 1.000

Enrichment: 1.176

Bayesian Score: 3.642

Mahalanobis Distance: 6.798

Mahalanobis Distance p-value: 0.999

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.

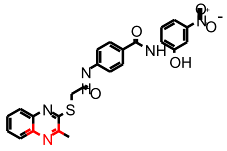
Model Applicability

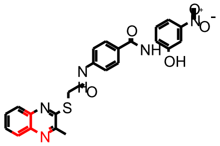
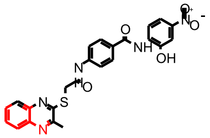
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_2 feature: -828984032: [*][N+](=[*])[c](:c:[*]):c:[*]
3. Unknown FCFP_2 feature: -1338588315: [*]:[c](:[*])[N+](=O)[O-]
4. Unknown FCFP_2 feature: 1872392852: [*][N+](=O)[*]
5. Unknown FCFP_2 feature: 260476081: [*][N+](=[*])[O-]

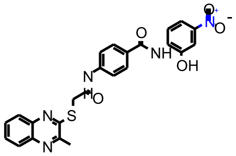
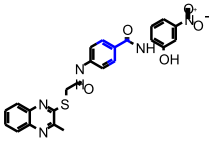
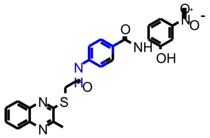
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	1747237384	 [*][c](:[*]):n:[c](:[*]):[*]	0.208	44 out of 44

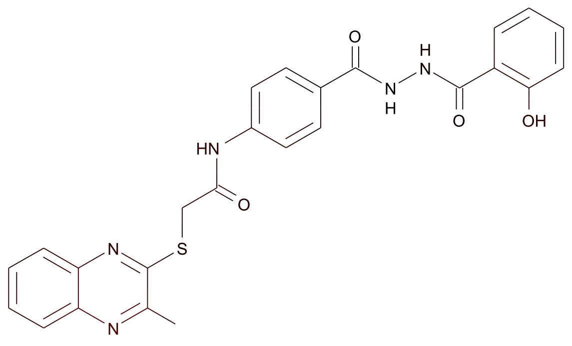
FCFP_12	178336375	 [*]:[cH]:[c](:n:[*]):[c](:[*]):[*]	0.202	19 out of 19
FCFP_12	713358128	 [*]:n:[c]1:[cH]:[cH]:[cH]:[*]:[c]:1:[*]	0.200	17 out of 17

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	8	 [*][N+](=[*])[*]	-0.056	3 out of 4
FCFP_12	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	0.000	319 out of 382
FCFP_12	-773983804	 [*]N[c]1:[cH]:[*]:[c]([*]):[cH]:[cH]:1	0.000	102 out of 121

16

TOPKAT_Ocular_Irritancy_None_vs_Irritant



$C_{25}H_{21}N_5O_4S$

Molecular Weight: 487.53034

ALogP: 3.093

Rotatable Bonds: 7

Acceptors: 7

Donors: 4

Model Prediction

Prediction: Irritant

Probability: 1.000

Enrichment: 1.176

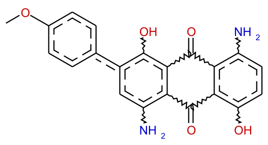
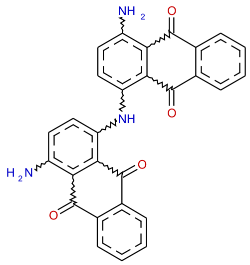
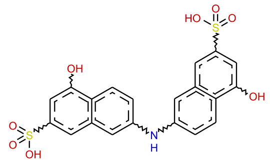
Bayesian Score: 3.768

Mahalanobis Distance: 7.692

Mahalanobis Distance p-value: 0.97

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	**ANTHRAQUINONE; 1;5-DIAMINO-4;8-DIHYDROXY-3-(p-METHOXYPHENYL)-**	**4;4'-DIAMINO-1;1'-DIANTHRIMIDE**	**2-Naphthalenesulfonic acid; 5;6'-iminobis(1-hydroxy-**
Structure			
Actual Endpoint	Irritant	Irritant	Irritant
Predicted Endpoint	Irritant	Irritant	Irritant
Distance	0.697	0.736	0.808
Reference	28ZPAK 245;72	28ZPAK-;125;72	Prehled Prumyslove Toxikologie; Organicke Latky; Marhold; J. pp 1065;86

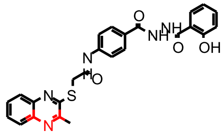
Model Applicability

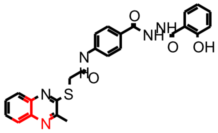
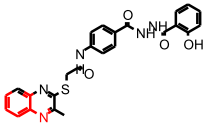
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_2 feature: -885461129: [*]NNC(=[*])[*]

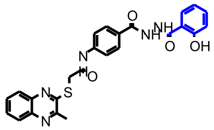
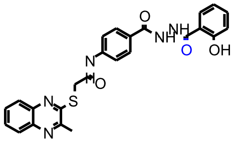
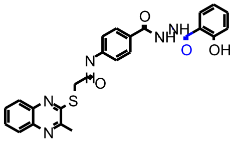
Feature Contribution

Top features for positive contribution

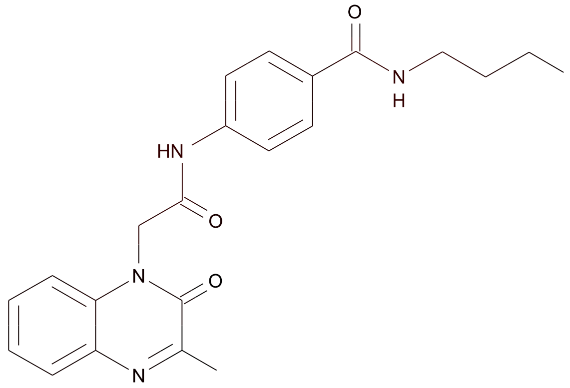
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Irritant in training set**
FCFP_12	1747237384	 [*][c](:[*]):n:[c](:[*]):[*]	0.208	44 out of 44

FCFP_12	178336375	 [*]:[cH]:[c](:n:[*]):[c](:[*]):[*]	0.202	19 out of 19
FCFP_12	713358128	 [*]:n:[c]1:[cH]:[cH]:[cH]:[*]:[c]:1:[*]	0.200	17 out of 17

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Irritant in training set**
FCFP_12	-1698724694	 [*]C(=[*])[c]1:[cH]:[cH]:[cH]:[cH]:[c]:1[*]	-0.096	107 out of 146
FCFP_12	1	 [*]=O	0.000	872 out of 1051
FCFP_12	1872154524	 [*]C(=O)[*]	0.000	563 out of 690

TOPKAT_Ocular_Irritancy_None_vs_Irritant



$C_{22}H_{24}N_4O_3$

Molecular Weight: 392.45096

ALogP: 2.212

Rotatable Bonds: 7

Acceptors: 4

Donors: 2

Model Prediction

Prediction: Irritant

Probability: 1.000

Enrichment: 1.176

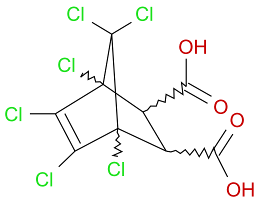
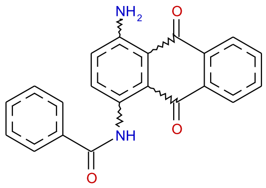
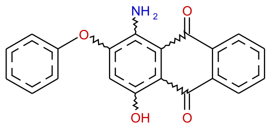
Bayesian Score: 2.098

Mahalanobis Distance: 10.517

Mahalanobis Distance p-value: 0.0231

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	**5-NORBORNENE-2;3-DICARBOXYLIC ACID; 1;4;5;6;7;7-HEXACHLORO-**	**1-AMINO-4-BENZOYLAMINO-ANTHRAQUINONE**	**ANTHRAQUINONE; 1-AMINO-4-HYDROXY-2-PHENOXY-**
Structure			
Actual Endpoint	Irritant	Irritant	Irritant
Predicted Endpoint	Irritant	Irritant	Irritant
Distance	0.627	0.631	0.692
Reference	28ZPAK-;92;72	28ZPAK-;124;72	28ZPAK 239;72

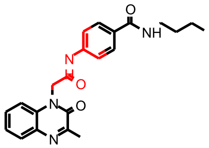
Model Applicability

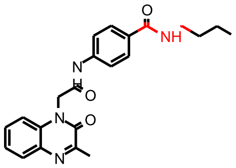
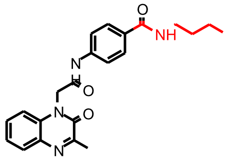
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

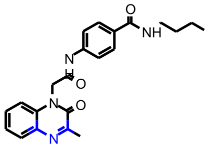
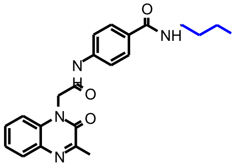
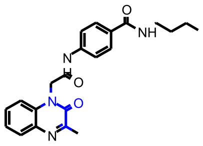
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	1175665944	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	0.198	14 out of 14

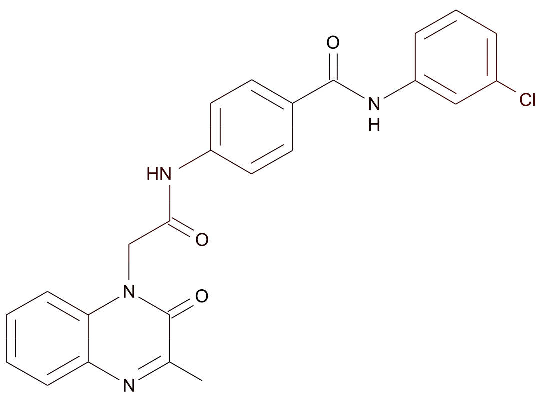
FCFP_12	-885550502	 [*]CNC(=[*])[*]	0.180	64 out of 66
FCFP_12	-1106976657	 [*]C(=[*])NCCCC	0.167	4 out of 4

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	580453787	 [*]C(=N[c](:[*]):[*]) [*]	-0.132	2 out of 3
FCFP_12	-1577600103	 [*]CCCC	-0.090	217 out of 294
FCFP_12	565998553	 [*]N([*])C(=O)C(=[*]) [*]	-0.066	198 out of 262

17b

TOPKAT_Ocular_Irritancy_None_vs_Irritant



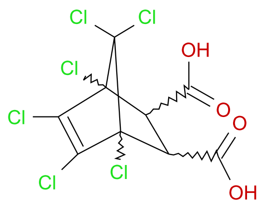
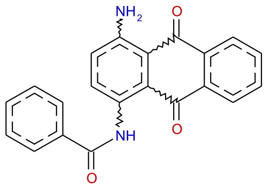
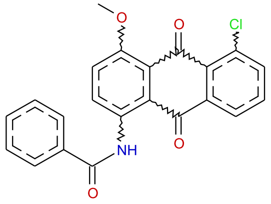
$C_{24}H_{19}ClN_4O_3$
Molecular Weight: 446.88566
ALogP: 3.125
Rotatable Bonds: 5
Acceptors: 4
Donors: 2

Model Prediction

Prediction: Irritant
Probability: 1.000
Enrichment: 1.176
Bayesian Score: 2.492
Mahalanobis Distance: 8.833
Mahalanobis Distance p-value: 0.602

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	**5-NORBORNENE-2;3-DICARBOXYLIC ACID; 1;4;5;6;7;7-HEXACHLORO-**	**1-AMINO-4-BENZOYLAMINO-ANTHRAQUINONE**	**1-BENZOYLAMINO-4-METHOXY-5-CHLORANTHRAQUINONE**
Structure			
Actual Endpoint	Irritant	Irritant	Irritant
Predicted Endpoint	Irritant	Irritant	Irritant
Distance	0.630	0.645	0.692
Reference	28ZPAK-;92;72	28ZPAK-;124;72	28ZPAK-;90;72

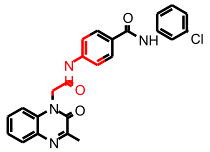
Model Applicability

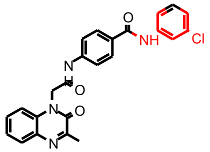
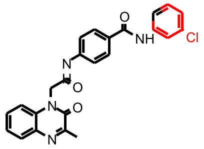
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

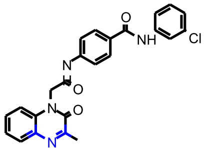
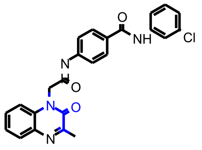
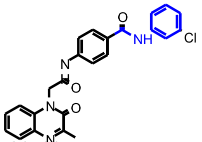
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	1175665944	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	0.198	14 out of 14

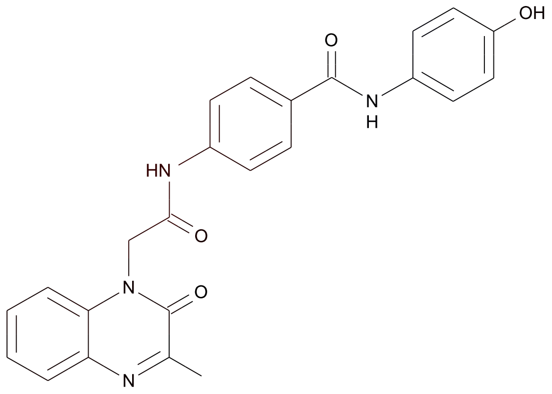
FCFP_12	-770645118	 [*]N[c]1:[cH]:[*]:[cH]:[c](Cl):[cH]:1	0.184	7 out of 7
FCFP_12	-745491832	 [*][c]1:[*]:[cH]:[cH]:[c](Cl):[cH]:1	0.177	32 out of 33

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	580453787	 [*]C(=N[c](:[*]):[*])[*]	-0.132	2 out of 3
FCFP_12	565998553	 [*]N([*])C(=O)C(=[*])[*]	-0.066	198 out of 262
FCFP_12	-792685140	 [*]C(=[*])N[c]1:[cH]:[cH]:[cH]:[c]([*]):[cH]:1	0.000	5 out of 6

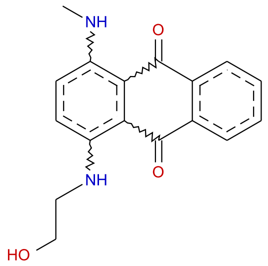
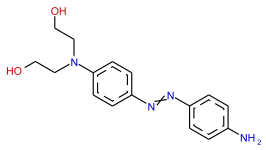
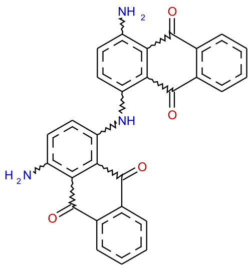
17c

TOPKAT_Ocular_Irritancy_None_vs_Irritant



$C_{24}H_{20}N_4O_4$
Molecular Weight: 428.44
ALogP: 2.218
Rotatable Bonds: 5
Acceptors: 5
Donors: 3

Structural Similar Compounds

Name	ANTHRAQUINONE; 1-((2-HYDROXYETHYL)AMINO)-4-(METHYLAMINO)-	Disperse Black 9	4;4'-DIAMINO-1;1'-DIANTHRIMIDE
Structure			
Actual Endpoint	Irritant	Non-Irritant	Irritant
Predicted Endpoint	Irritant	Non-Irritant	Irritant
Distance	0.675	0.699	0.719
Reference	28ZPAK 245;72	J. Am. Coll. Toxicol. 5(3):205;1986	28ZPAK-;125;72

Model Prediction

Prediction: Irritant
Probability: 1.000
Enrichment: 1.176
Bayesian Score: 1.646
Mahalanobis Distance: 9.030
Mahalanobis Distance p-value: 0.495

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.

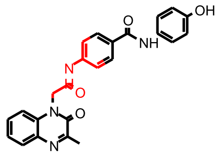
Model Applicability

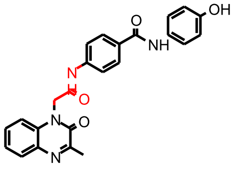
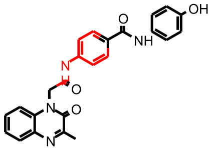
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

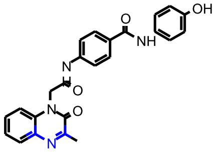
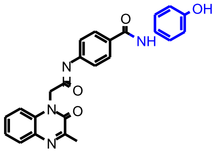
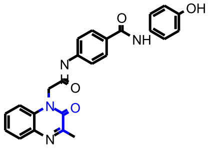
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	1175665944	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	0.198	14 out of 14

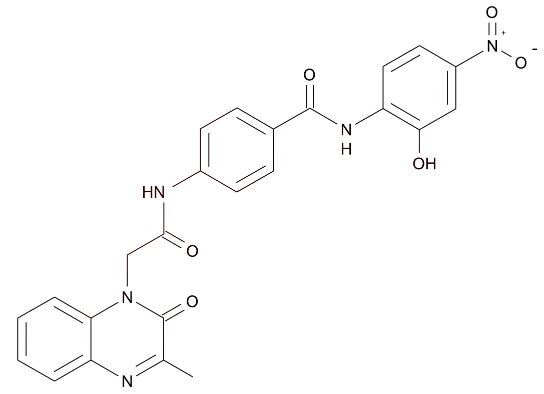
FCFP_12	566058135	 [*]CC(=O)N[*]	0.163	23 out of 24
FCFP_12	-1838187238	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	0.147	17 out of 18

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Irritant in training set**
FCFP_12	580453787	 [*]C(=N[c](:[*]):[*])[*]	-0.132	2 out of 3
FCFP_12	-1849978862	 [*]N[c]1:[cH]:[cH]:[c](O):[cH]:[cH]:1	-0.132	2 out of 3
FCFP_12	565998553	 [*]N([*])C(=O)C(=[*])[*]	-0.066	198 out of 262

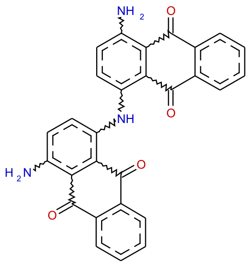
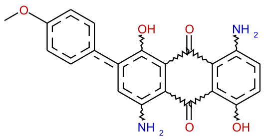
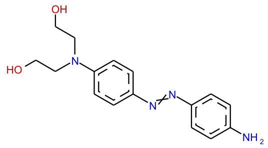
17d

TOPKAT_Ocular_Irritancy_None_vs_Irritant



$C_{24}H_{19}N_5O_6$
Molecular Weight: 473.43756
ALogP: 2.113
Rotatable Bonds: 6
Acceptors: 7
Donors: 3

Structural Similar Compounds

Name	**4;4'-DIAMINO-1;1'-DIANTHRIMIDE**	**ANTHRAQUINONE; 1;5-DIAMINO-4;8-DIHYDROXY-3-(p-METHOXYPHENYL)-**	**Disperse Black 9**
Structure			
Actual Endpoint	Irritant	Irritant	Non-Irritant
Predicted Endpoint	Irritant	Irritant	Non-Irritant
Distance	0.654	0.762	0.773
Reference	28ZPAK-;125;72	28ZPAK 245;72	J. Am. Coll. Toxicol. 5(3):205;1986

Model Prediction

Prediction: Irritant
Probability: 1.000
Enrichment: 1.176
Bayesian Score: 2.016
Mahalanobis Distance: 8.881
Mahalanobis Distance p-value: 0.576

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.

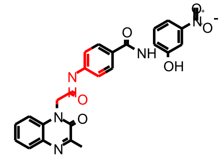
Model Applicability

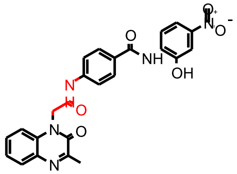
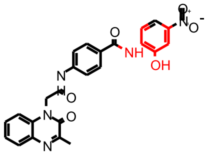
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_2 feature: -828984032: [*][N+](=[*])[c](:c:[*]):c:[*]
3. Unknown FCFP_2 feature: -1338588315: [*]:[c](:[*])[N+](=O)[O-]
4. Unknown FCFP_2 feature: 1872392852: [*][N+](=O)[*]
5. Unknown FCFP_2 feature: 260476081: [*][N+](=[*])[O-]

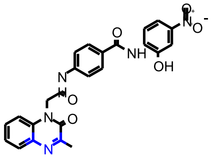
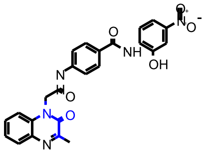
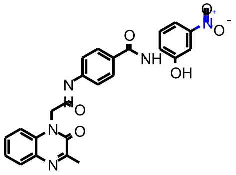
Feature Contribution

Top features for positive contribution

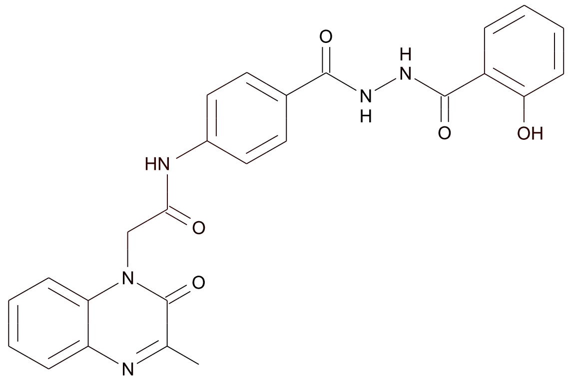
Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	1175665944	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	0.198	14 out of 14

FCFP_12	566058135	 [*]CC(=O)N[*]	0.163	23 out of 24
FCFP_12	-1924607822	 [*]N[c]1:[cH]:[*]:[c]([*]):[cH]:[c]:1O	0.156	3 out of 3

Top Features for negative contribution

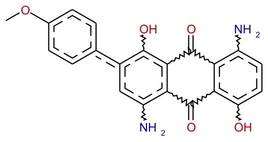
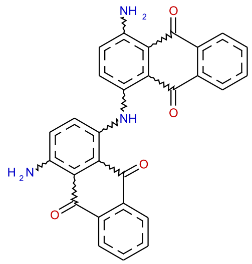
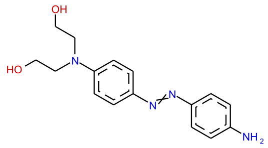
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Irritant in training set**
FCFP_12	580453787	 [*]C(=N[c](:[*]):[*])[*]	-0.132	2 out of 3
FCFP_12	565998553	 [*]N([*])C(=O)C(=[*])[*]	-0.066	198 out of 262
FCFP_12	8	 [*][N+](=[*])[*]	-0.056	3 out of 4

18 TOPKAT_Ocular_Irritancy_None_vs_Irritant



$C_{25}H_{21}N_5O_5$
Molecular Weight: 471.46474
ALogP: 1.475
Rotatable Bonds: 6
Acceptors: 6
Donors: 4

Structural Similar Compounds

Name	ANTHRAQUINONE; 1;5-DIAMINO-4;8-DIHYDROXY-3-(p-METHOXYPHENYL)-	4;4'-DIAMINO-1;1'-DIANTHRIMIDE	Disperse Black 9
Structure			
Actual Endpoint	Irritant	Irritant	Non-Irritant
Predicted Endpoint	Irritant	Irritant	Non-Irritant
Distance	0.694	0.798	0.838
Reference	28ZPAK 245;72	28ZPAK-;125;72	J. Am. Coll. Toxicol. 5(3):205;1986

Model Prediction

Prediction: Irritant
Probability: 1.000
Enrichment: 1.176
Bayesian Score: 2.145
Mahalanobis Distance: 10.269
Mahalanobis Distance p-value: 0.0474

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.

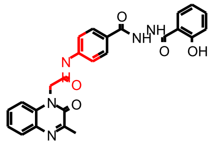
Model Applicability

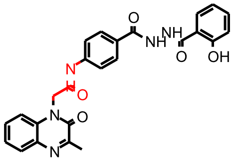
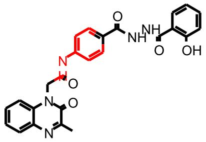
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_2 feature: -885461129: [*]NNC(=[*])[*]

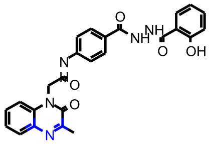
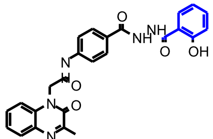
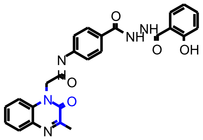
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	1175665944	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	0.198	14 out of 14

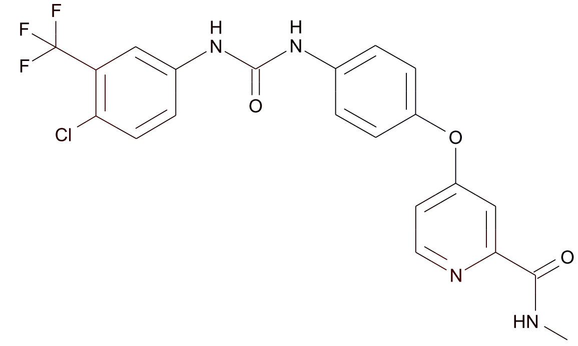
FCFP_12	566058135	 [*]CC(=O)N[*]	0.163	23 out of 24
FCFP_12	-1838187238	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	0.147	17 out of 18

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	580453787	 [*]C(=N[c](:[*]):[*])[*]	-0.132	2 out of 3
FCFP_12	-1698724694	 [*]C(=[*])[c]1:[cH]:[cH]:[cH]:[cH]:[c]:1[*]	-0.096	107 out of 146
FCFP_12	565998553	 [*]N([*])C(=O)C(=[*])[*]	-0.066	198 out of 262

Sorafenib

TOPKAT_Ocular_Irritancy_None_vs_Irritant



$C_{21}H_{16}ClF_3N_4O_3$

Molecular Weight: 464.82494

ALogP: 4.175

Rotatable Bonds: 6

Acceptors: 4

Donors: 3

Model Prediction

Prediction: Irritant

Probability: 1.000

Enrichment: 1.176

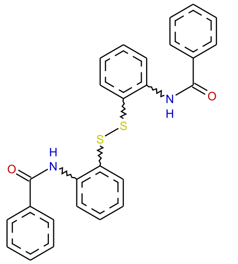
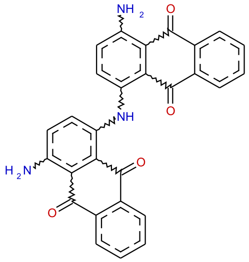
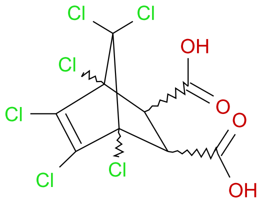
Bayesian Score: 3.039

Mahalanobis Distance: 6.281

Mahalanobis Distance p-value: 1

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	BENZANILIDE;2';2'''-DITHIOBIS-	4;4'-DIAMINO-1;1'-DIANTHRIMIDE	5-NORBORNENE-2;3-DICARBOXYLIC ACID; 1;4;5;6;7;7-HEXACHLORO-
Structure			
Actual Endpoint	Non-Irritant	Irritant	Irritant
Predicted Endpoint	Non-Irritant	Irritant	Irritant
Distance	0.743	0.791	0.801
Reference	28ZPAK-;173;72	28ZPAK-;125;72	28ZPAK-;92;72

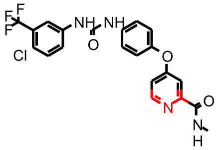
Model Applicability

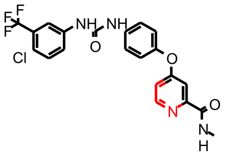
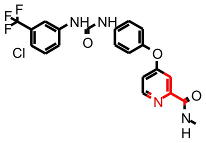
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

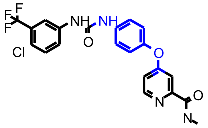
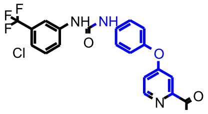
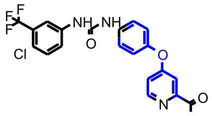
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	1747237384	 [*][c](:[*]):n:[c](:[*]):[*]	0.208	44 out of 44

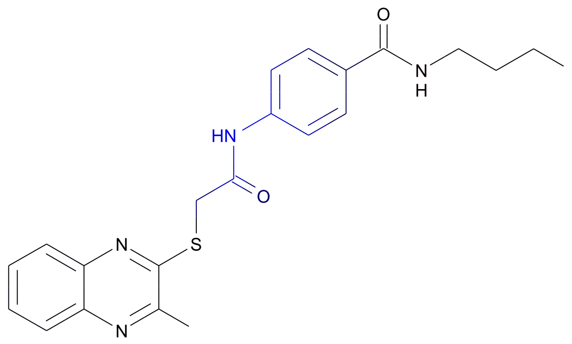
FCFP_12	-124655670	 [*]:[cH]:[cH]:n:[*]	0.200	16 out of 16
FCFP_12	-1539132615	 [*][c](:[*]):[c](C):n:[*]	0.197	13 out of 13

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	-747629521	 [*]N[c]1:[cH]:[cH]:[c](O[c](:[*]):[*]):[cH]:[cH]:1	-0.268	1 out of 2
FCFP_12	702861189	 [*]N[c]1:[cH]:[cH]:[c](O[c]2:[cH]:[cH]:[*]:[c]([*]):[cH]:2):[cH]:[cH]:1	-0.268	1 out of 2
FCFP_12	-215363676	 [*][c]1:[*]:[cH]:[cH]:[c](O[c]2:[cH]:[cH]:[*]:[cH]:[cH]:2):[cH]:1	0.000	4 out of 5

15a

TOPKAT_Skin_Irritancy_None_vs_Irritant



$C_{22}H_{24}N_4O_2S$

Molecular Weight: 408.51656

ALogP: 3.83

Rotatable Bonds: 8

Acceptors: 5

Donors: 2

Model Prediction

Prediction: Non-Irritant

Probability: 0.352

Enrichment: 0.382

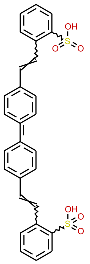
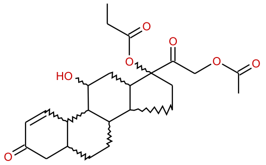
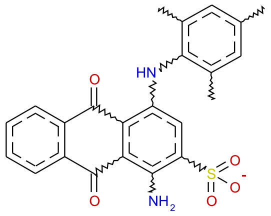
Bayesian Score: -4.947

Mahalanobis Distance: 9.995

Mahalanobis Distance p-value: 0.0635

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	**Benzenesulfonic acid, 2,2'-(4,4'-biphenylylenedivinylene)di-, disod ium salt**	**Pregna-1,4-diene-3,20-dione, 21-(acetyloxy)-11-hydroxy-6-methyl-17- (1-oxopropoxy)-, (6-alpha,11-beta)-**	**2-Anthracenesulfonic acid, 1-amino-9,10-dihydro-9,10-dioxo-4-(2,4,6 -trimethylanilino)-, monosodium salt**
Structure			
Actual Endpoint	Irritant	Irritant	Irritant
Predicted Endpoint	Non-Irritant	Irritant	Non-Irritant
Distance	0.801	0.815	0.849
Reference	MVCRB3 MVC-Report. (Stockholm, Sweden) No.1-2, 1972-73. Discontinued. Volume(issue)/page/year: 2,193,1973	YACHDS Yakuri to Chiryo. Pharmacology and Therapeutics. (Raifu Saiensu Shup pan K.K., 2-5-13, Yaesu, Chuo-ku, Tokyo 104, Japan) V.1-1972- Volume(issue) /page/year: 19,3103,1991	85JCAE "Prehled Prumyslove Toxikologie; Organicke Latky," Marhold, J., Prague , Czechoslovakia, Avicenum, 1986 Volume(issue)/page/year: -,1327,1986

Model Applicability

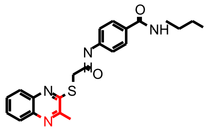
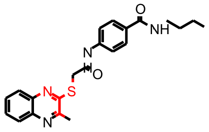
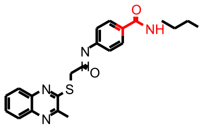
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

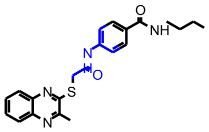
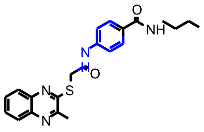
Feature Contribution

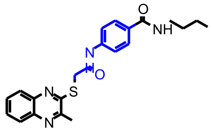
Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Irritant in training set**

FCFP_12	-1539132615	 [*][c](:[*]):[c](C):n :[*]	0.079	9 out of 9
FCFP_12	-1410049896	 [*]S[c](:n:[*]):[c]([*]):[*]	0.073	5 out of 5
FCFP_12	-1549103449	 [*]NC(=O)[c](:[*]):[*]	0.073	5 out of 5

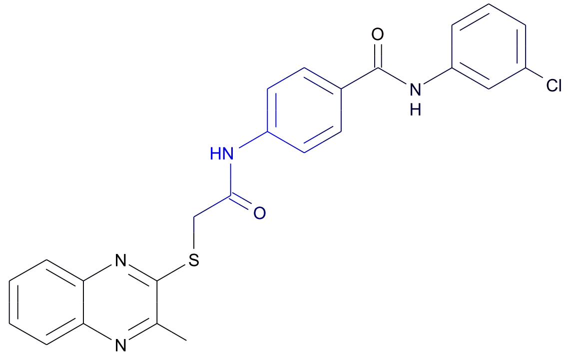
Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	1175665944	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	-1.020	2 out of 8
FCFP_12	-1838187238	 [*]C(=[*])N[c]1:[cH]: [cH]:[*]:[cH]:[cH]:1	-0.692	5 out of 12

FCFP_12	-451043714	 [*]CC(=O)N[c]1:[cH]:[cH]:[c]([*]):[cH]:[cH]:1	-0.650	0 out of 1

15b

TOPKAT_Skin_Irritancy_None_vs_Irritant



$C_{24}H_{19}ClN_4O_2S$

Molecular Weight: 462.95126

ALogP: 4.743

Rotatable Bonds: 6

Acceptors: 5

Donors: 2

Model Prediction

Prediction: Non-Irritant

Probability: 0.051

Enrichment: 0.055

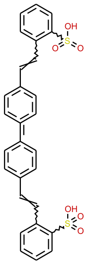
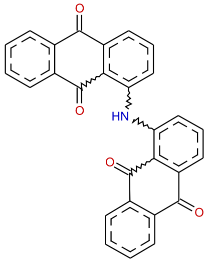
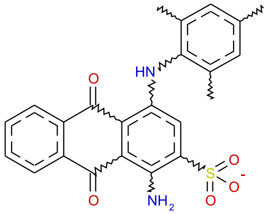
Bayesian Score: -6.256

Mahalanobis Distance: 7.401

Mahalanobis Distance p-value: 0.981

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	Benzenesulfonic acid, 2,2'-(4,4'-biphenylylenedivinylene)di-, disod ium salt	Anthraquinone, 1,1'-iminodi-	2-Anthracenesulfonic acid, 1-amino-9,10-dihydro-9,10-dioxo-4-(2,4,6 -trimethylanilino)-, monosodium salt
Structure			
Actual Endpoint	Irritant	Irritant	Irritant
Predicted Endpoint	Non-Irritant	Non-Irritant	Non-Irritant
Distance	0.683	0.786	0.855
Reference	MVCRB3 MVC-Report. (Stockholm, Sweden) No.1-2, 1972-73. Discontinued. Volume(issue)/page/year: 2,193,1973	85JCAE "Prehled Prumyslove Toxikologie; Organicke Latky," Marhold, J., Prague , Czechoslovakia, Avicenum, 1986 Volume(issue)/page/year: -,735,1986	85JCAE "Prehled Prumyslove Toxikologie; Organicke Latky," Marhold, J., Prague , Czechoslovakia, Avicenum, 1986 Volume(issue)/page/year: -,1327,1986

Model Applicability

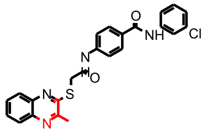
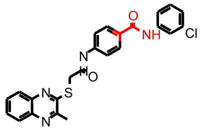
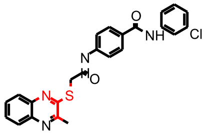
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

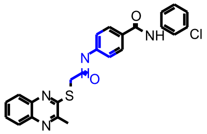
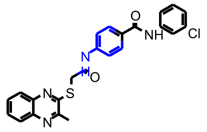
Feature Contribution

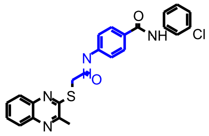
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set

FCFP_12	-1539132615	 [*][c](:[*]):[c](C):n:[*]	0.079	9 out of 9
FCFP_12	-1549103449	 [*]NC(=O)[c](:[*]):[*]	0.073	5 out of 5
FCFP_12	-1410049896	 [*]S[c](:n:[*]):[c]([*]):[*]	0.073	5 out of 5

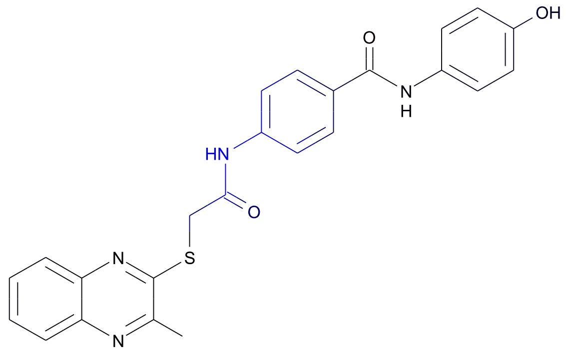
Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	1175665944	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	-1.020	2 out of 8
FCFP_12	-1838187238	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	-0.692	5 out of 12

FCFP_12	-451043714	 [*]CC(=O)N[c]1:[cH]:[cH]:[c]([*]):[cH]:[cH]:1	-0.650	0 out of 1

15c

TOPKAT_Skin_Irritancy_None_vs_Irritant



$C_{24}H_{20}N_4O_3S$

Molecular Weight: 444.5056

ALogP: 3.836

Rotatable Bonds: 6

Acceptors: 6

Donors: 3

Model Prediction

Prediction: Non-Irritant

Probability: 0.128

Enrichment: 0.139

Bayesian Score: -5.715

Mahalanobis Distance: 8.326

Mahalanobis Distance p-value: 0.768

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.

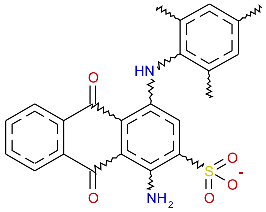
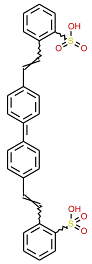
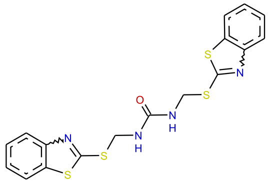
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.

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	**2-Anthracenesulfonic acid, 1-amino-9,10-dihydro-9,10-dioxo-4-(2,4,6 -trimethylanilino)-, monosodium salt**	**Benzenesulfonic acid, 2,2'-(4,4'-biphenylylenedivinylene)di-, disod ium salt**	**Urea, 1,3-bis(2-benzothiazolylthiomethyl)-**
Structure			
Actual Endpoint	Irritant	Irritant	Irritant
Predicted Endpoint	Non-Irritant	Non-Irritant	Irritant
Distance	0.777	0.799	0.845
Reference	85JCAE "Prehled Prumyslove Toxikologie; Organicke Latky," Marhold, J., Prague , Czechoslovakia, Avicenum, 1986 Volume(issue)/page/year: -,1327,1986	MVCRB3 MVC-Report. (Stockholm, Sweden) No.1-2, 1972-73. Discontinued. Volu me(issue)/page/year: 2,193,1973	AMIHBC AMA Archives of Industrial Hygiene and Occupational Medicine. (Chicago , IL) V.2-10, 1950-54. For publisher information, see AEHLAU. Volume(issue)/pag e/year: 5,311,1952

Model Applicability

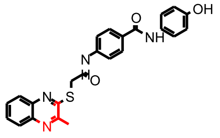
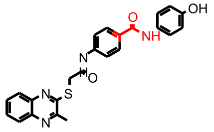
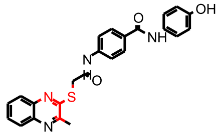
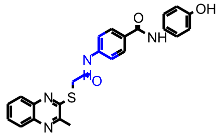
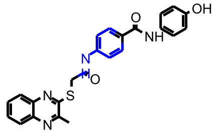
Unknown features are fingerprint features in the query molecule, but not found in the training set.

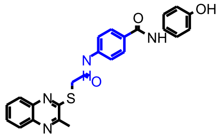
1. All properties and OPS components are within expected ranges.

Feature Contribution

Top features for positive contribution

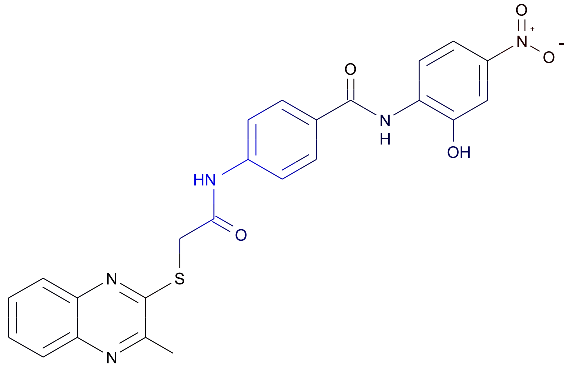
Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set

FCFP_12	-1539132615	 [*][c](:[*]):[c](C):n:[*]	0.079	9 out of 9
FCFP_12	-1549103449	 [*]NC(=O)[c](:[*]):[*]	0.073	5 out of 5
FCFP_12	-1410049896	 [*]S[c](:n:[*]):[c]([*]):[*]	0.073	5 out of 5
Top Features for negative contribution				
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Irritant in training set**
FCFP_12	1175665944	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	-1.020	2 out of 8
FCFP_12	-1838187238	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	-0.692	5 out of 12

FCFP_12	-451043714	 [*]CC(=O)N[c]1:[cH]:[cH]:[c]([*]):[cH]:[cH]:1	-0.650	0 out of 1

15d

TOPKAT_Skin_Irritancy_None_vs_Irritant



$C_{24}H_{19}N_5O_5S$

Molecular Weight: 489.50316

ALogP: 3.731

Rotatable Bonds: 7

Acceptors: 8

Donors: 3

Model Prediction

Prediction: Non-Irritant

Probability: 0.098

Enrichment: 0.106

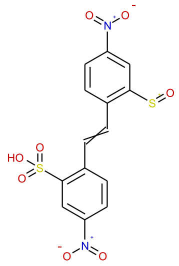
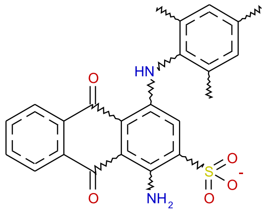
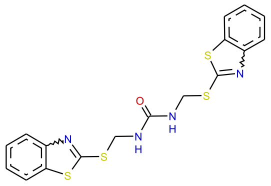
Bayesian Score: -5.883

Mahalanobis Distance: 9.348

Mahalanobis Distance p-value: 0.251

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	**2,2'-Stilbenedisulfonic acid, 4,4'-dinitro-**	**2-Anthracenesulfonic acid, 1-amino-9,10-dihydro-9,10-dioxo-4-(2,4,6 -trimethylanilino)-, monosodium salt**	**Urea, 1,3-bis(2-benzothiazolylthiomethyl)-**
Structure			
Actual Endpoint	Irritant	Irritant	Irritant
Predicted Endpoint	Irritant	Non-Irritant	Irritant
Distance	0.806	0.864	0.868
Reference	85JCAE "Prehled Prumyslove Toxikologie; Organicke Latky," Marhold, J., Prague , Czechoslovakia, Avicenum, 1986 Volume(issue)/page/year: -,1062,1986	85JCAE "Prehled Prumyslove Toxikologie; Organicke Latky," Marhold, J., Prague , Czechoslovakia, Avicenum, 1986 Volume(issue)/page/year: -,1327,1986	AMIHBC AMA Archives of Industrial Hygiene and Occupational Medicine. (Chicago , IL) V.2-10, 1950-54. For publisher information, see AEHLAU. Volume(issue)/pag e/year: 5,311,1952

Model Applicability

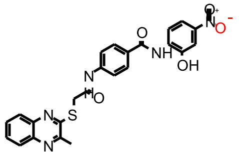
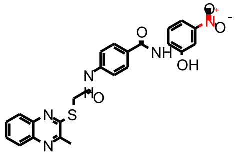
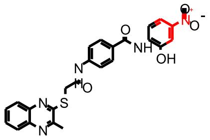
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

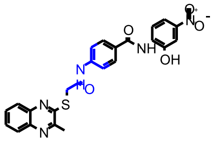
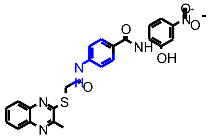
Feature Contribution

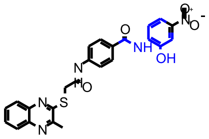
Top features for positive contribution

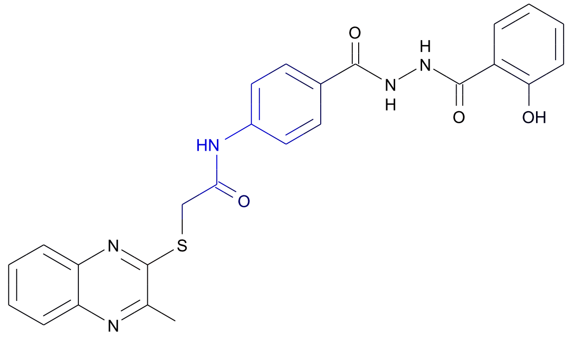
Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set

FCFP_12	5	 [*][O-]	0.085	27 out of 27
FCFP_12	8	 [*][N+](=[*])[*]	0.084	20 out of 20
FCFP_12	-828984032	 [*][N+](=[*])[c](:[cH]:[*]):[cH]:[*]	0.079	9 out of 9

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	1175665944	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	-1.020	2 out of 8
FCFP_12	-1838187238	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	-0.692	5 out of 12

FCFP_12	-1883332927	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[c]:1O	-0.650	0 out of 1



$C_{25}H_{21}N_5O_4S$

Molecular Weight: 487.53034

ALogP: 3.093

Rotatable Bonds: 7

Acceptors: 7

Donors: 4

Model Prediction

Prediction: Non-Irritant

Probability: 0.204

Enrichment: 0.222

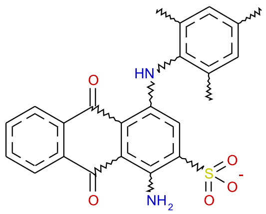
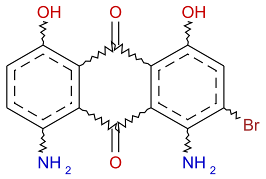
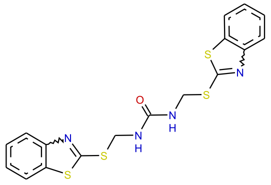
Bayesian Score: -5.394

Mahalanobis Distance: 9.072

Mahalanobis Distance p-value: 0.383

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	**2-Anthracenesulfonic acid, 1-amino-9,10-dihydro-9,10-dioxo-4-(2,4,6 -trimethylanilino)-, monosodium salt**	**Anthraquinone, 2-bromo-1,8-diamino-4,5-dihydroxy-**	**Urea, 1,3-bis(2-benzothiazolylthiomethyl)-**
Structure			
Actual Endpoint	Irritant	Irritant	Irritant
Predicted Endpoint	Non-Irritant	Non-Irritant	Irritant
Distance	0.941	0.965	0.980
Reference	85JCAE "Prehled Prumyslove Toxikologie; Organicke Latky," Marhold, J., Prague , Czechoslovakia, Avicenum, 1986 Volume(issue)/page/year: -,1327,1986	28ZPAK "Sbornik Vysledku Toxixologickeho Vysetreni Latek A Pripravku," Marhol d, J.V., Institut Pro Vychovu Vedoucicn Pracovniku Chemickeho Prumyclu Praha, Cz echoslovakia, 1972 Volume(issue)/page/year: -,244,1	AMIHBC AMA Archives of Industrial Hygiene and Occupational Medicine. (Chicago , IL) V.2-10, 1950-54. For publisher information, see AEHLAU. Volume(issue)/pag e/year: 5,311,1952

Model Applicability

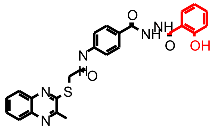
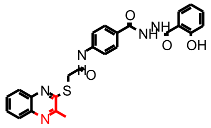
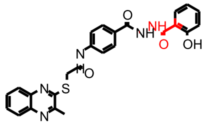
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

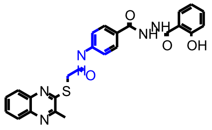
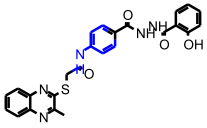
Feature Contribution

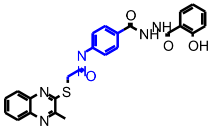
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set

FCFP_12	-309153329	 [*]C(=[*])[c]1:[cH]:[cH]:[cH]:[cH]:[c]:1O	0.082	13 out of 13
FCFP_12	-1539132615	 [*][c](:[*]):[c](C):n:[*]	0.079	9 out of 9
FCFP_12	-1549103449	 [*]NC(=O)[c](:[*]):[*]	0.073	5 out of 5

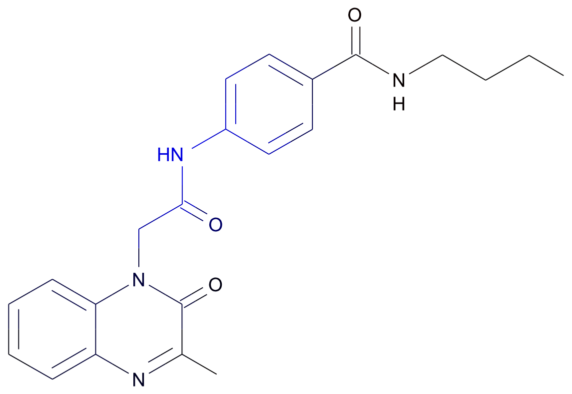
Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	1175665944	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	-1.020	2 out of 8
FCFP_12	-1838187238	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	-0.692	5 out of 12

FCFP_12	-451043714	 [*]CC(=O)N[c]1:[cH]:[cH]:[c]([*]):[cH]:[cH]:1	-0.650	0 out of 1

17a

TOPKAT_Skin_Irritancy_None_vs_Irritant



$C_{22}H_{24}N_4O_3$

Molecular Weight: 392.45096

ALogP: 2.212

Rotatable Bonds: 7

Acceptors: 4

Donors: 2

Model Prediction

Prediction: Non-Irritant

Probability: 0.057

Enrichment: 0.061

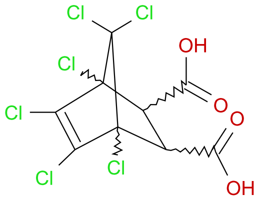
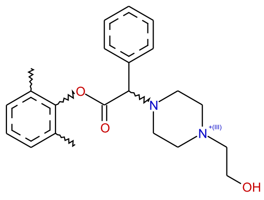
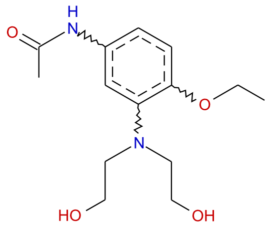
Bayesian Score: -6.199

Mahalanobis Distance: 12.635

Mahalanobis Distance p-value: 7.95e-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	**5-Norbornene-2,3-dicarboxylic acid, 1,4,5,6,7,7-hexachloro-**	**1-Piperazineacetic acid, 4-(2-hydroxyethyl)-alpha-phenyl-, 2,6-xyly l ester, monohydrochloride**	**p-Acetophenetidide, 3'-(bis(2-hydroxyethyl)amino)-**
Structure			
Actual Endpoint	Irritant	Irritant	Non-Irritant
Predicted Endpoint	Irritant	Irritant	Non-Irritant
Distance	0.730	0.792	0.796
Reference	85JCAE "Prehled Prumyslove Toxikologie; Organicke Latky," Marhold, J., Prague , Czechoslovakia, Avicenum, 1986 Volume(issue)/page/year: -,581,1986	BCFAAI Bollettino Chimico Farmaceutico. (Societa Editoriale Farmaceutica, Vi a Ausonio 12, 20123 Milan, Italy) V.33- 1894- Volume(issue)/page/year: 107,3 10,1968	28ZPAK -,100,72

Model Applicability

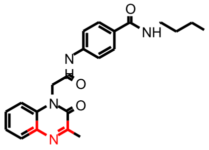
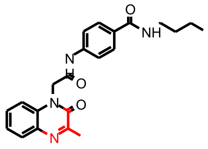
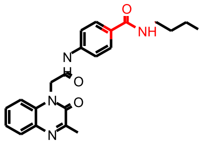
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

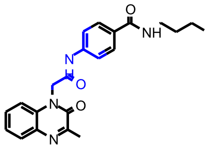
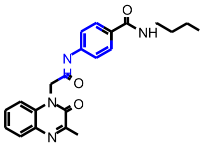
Feature Contribution

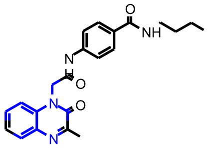
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set

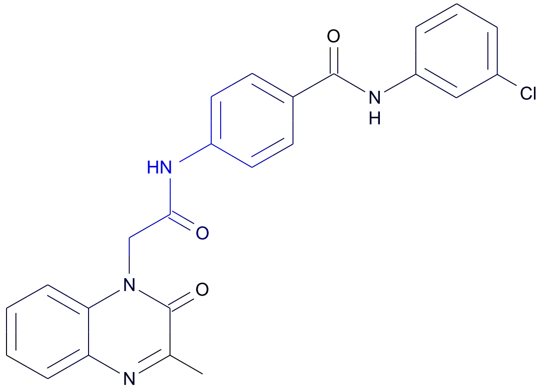
FCFP_12	580453787	 [*]C(=N[c](:[*]):[*]) [*]	0.079	9 out of 9
FCFP_12	565968762	 [*]\N=C(\C)/C(=[*])[*]	0.075	78 out of 79
FCFP_12	-1549103449	 [*]NC(=O)[c](:[*]):[*]	0.073	5 out of 5

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	1175665944	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	-1.020	2 out of 8
FCFP_12	-1838187238	 [*]C(=[*])N[c]1:[cH]: [cH]:[*]:[cH]:[cH]:1	-0.692	5 out of 12

FCFP_12	-2007573814	 [*]CN1C(=[*])[*]=N[c]2:[cH]:[*]:[cH]:[cH]:[c]1:2	-0.650	0 out of 1

TOPKAT_Skin_Irritancy_None_vs_Irritant



$C_{24}H_{19}ClN_4O_3$

Molecular Weight: 446.88566

ALogP: 3.125

Rotatable Bonds: 5

Acceptors: 4

Donors: 2

Model Prediction

Prediction: Non-Irritant

Probability: 0.004

Enrichment: 0.004

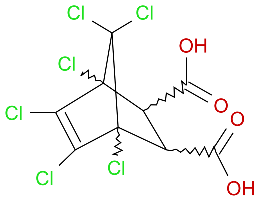
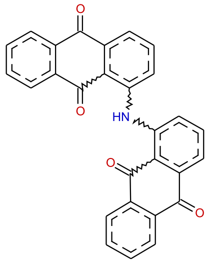
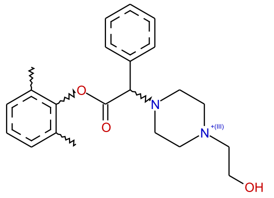
Bayesian Score: -7.507

Mahalanobis Distance: 10.918

Mahalanobis Distance p-value: 0.00339

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	**5-Norbornene-2,3-dicarboxylic acid, 1,4,5,6,7,7-hexachloro-**	**Anthraquinone, 1,1'-iminodi-**	**1-Piperazineacetic acid, 4-(2-hydroxyethyl)-alpha-phenyl-, 2,6-xyly l ester, monohydrochloride**
Structure			
Actual Endpoint	Irritant	Irritant	Irritant
Predicted Endpoint	Irritant	Non-Irritant	Irritant
Distance	0.698	0.797	0.815
Reference	85JCAE "Prehled Prumyslove Toxikologie; Organicke Latky," Marhold, J., Prague , Czechoslovakia, Avicenum, 1986 Volume(issue)/page/year: -,581,1986	85JCAE "Prehled Prumyslove Toxikologie; Organicke Latky," Marhold, J., Prague , Czechoslovakia, Avicenum, 1986 Volume(issue)/page/year: -,735,1986	BCFAAI Bollettino Chimico Farmaceutico. (Societa Editoriale Farmaceutica, Vi a Ausonio 12, 20123 Milan, Italy) V.33- 1894- Volume(issue)/page/year: 107,3 10,1968

Model Applicability

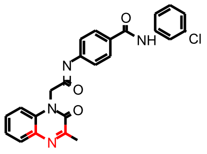
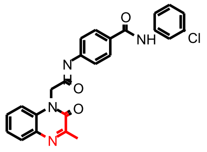
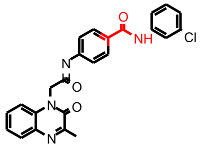
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

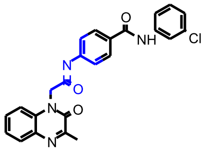
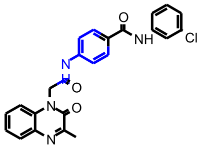
Feature Contribution

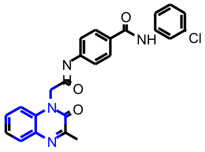
Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Irritant in training set**

FCFP_12	580453787	 [*]C(=N[c](:[*]):[*])[*]	0.079	9 out of 9
FCFP_12	565968762	 [*]\N=C(\C)/C(=[*])[*]	0.075	78 out of 79
FCFP_12	-1549103449	 [*]NC(=O)[c](:[*]):[*]	0.073	5 out of 5

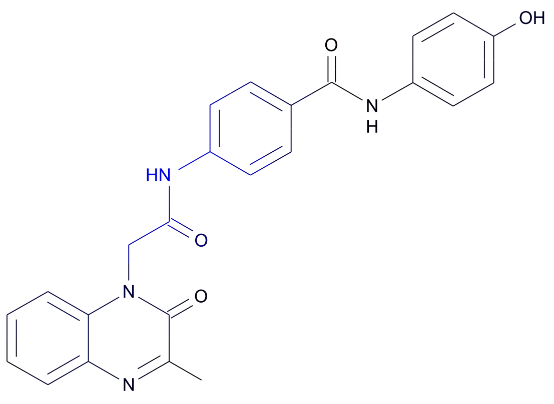
Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	1175665944	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	-1.020	2 out of 8
FCFP_12	-1838187238	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	-0.692	5 out of 12

FCFP_12	-2007573814	 [*]CN1C(=[*])[*]=N[c]2:[cH]:[*]:[cH]:[cH]:[c]1:2	-0.650	0 out of 1

17c

TOPKAT_Skin_Irritancy_None_vs_Irritant



$C_{24}H_{20}N_4O_4$

Molecular Weight: 428.44

ALogP: 2.218

Rotatable Bonds: 5

Acceptors: 5

Donors: 3

Model Prediction

Prediction: Non-Irritant

Probability: 0.021

Enrichment: 0.023

Bayesian Score: -6.700

Mahalanobis Distance: 11.932

Mahalanobis Distance p-value: 3.78e-005

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.

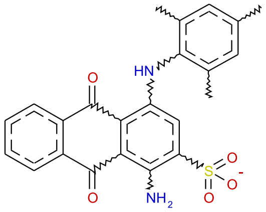
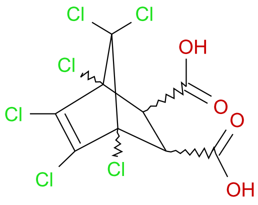
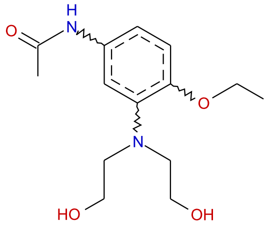
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.

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	**2-Anthracenesulfonic acid, 1-amino-9,10-dihydro-9,10-dioxo-4-(2,4,6 -trimethylanilino)-, monosodium salt**	**5-Norbornene-2,3-dicarboxylic acid, 1,4,5,6,7,7-hexachloro-**	**p-Acetophenetidide, 3'-(bis(2-hydroxyethyl)amino)-**
Structure			
Actual Endpoint	Irritant	Irritant	Non-Irritant
Predicted Endpoint	Non-Irritant	Irritant	Non-Irritant
Distance	0.808	0.820	0.834
Reference	85JCAE "Prehled Prumyslove Toxikologie; Organicke Latky," Marhold, J., Prague , Czechoslovakia, Avicenum, 1986 Volume(issue)/page/year: -,1327,1986	85JCAE "Prehled Prumyslove Toxikologie; Organicke Latky," Marhold, J., Prague , Czechoslovakia, Avicenum, 1986 Volume(issue)/page/year: -,581,1986	28ZPAK -,100,72

Model Applicability

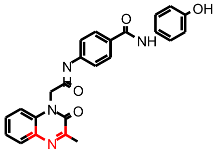
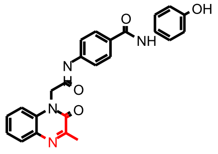
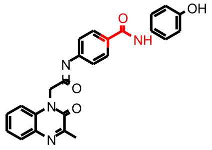
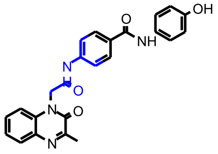
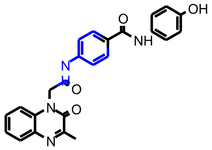
Unknown features are fingerprint features in the query molecule, but not found in the training set.

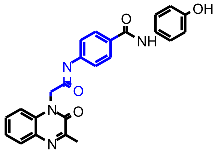
1. All properties and OPS components are within expected ranges.

Feature Contribution

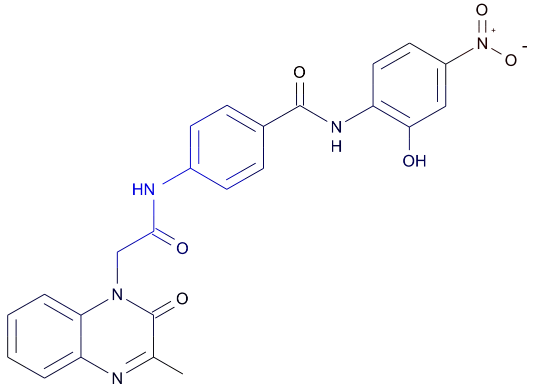
Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Irritant in training set**

FCFP_12	580453787	 [*]C(=N[c](:[*]):[*])[*]	0.079	9 out of 9
FCFP_12	565968762	 [*]\N=C(\C)/C(=[*])[*]	0.075	78 out of 79
FCFP_12	-1549103449	 [*]NC(=O)[c](:[*]):[*]	0.073	5 out of 5
Top Features for negative contribution				
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Irritant in training set**
FCFP_12	1175665944	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	-1.020	2 out of 8
FCFP_12	-1838187238	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	-0.692	5 out of 12

FCFP_12	-451043714	 [*]CC(=O)N[c]1:[cH]:[cH]:[c]([*]):[cH]:[cH]:1	-0.650	0 out of 1

TOPKAT_Skin_Irritancy_None_vs_Irritant



$C_{24}H_{19}N_5O_6$

Molecular Weight: 473.43756

ALogP: 2.113

Rotatable Bonds: 6

Acceptors: 7

Donors: 3

Model Prediction

Prediction: Non-Irritant

Probability: 0.012

Enrichment: 0.013

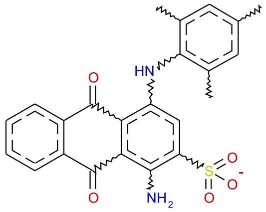
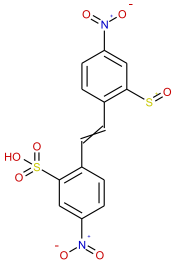
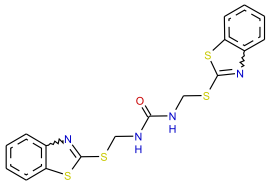
Bayesian Score: -6.975

Mahalanobis Distance: 12.655

Mahalanobis Distance p-value: 7.07e-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	**2-Anthracenesulfonic acid, 1-amino-9,10-dihydro-9,10-dioxo-4-(2,4,6 -trimethylanilino)-, monosodium salt**	**2,2'-Stilbenedisulfonic acid, 4,4'-dinitro-**	**Urea, 1,3-bis(2-benzothiazolylthiomethyl)-**
Structure			
Actual Endpoint	Irritant	Irritant	Irritant
Predicted Endpoint	Non-Irritant	Irritant	Irritant
Distance	0.743	0.858	0.912
Reference	85JCAE "Prehled Prumyslove Toxikologie; Organicke Latky," Marhold, J., Prague , Czechoslovakia, Avicenum, 1986 Volume(issue)/page/year: -,1327,1986	85JCAE "Prehled Prumyslove Toxikologie; Organicke Latky," Marhold, J., Prague , Czechoslovakia, Avicenum, 1986 Volume(issue)/page/year: -,1062,1986	AMIHBC AMA Archives of Industrial Hygiene and Occupational Medicine. (Chicago , IL) V.2-10, 1950-54. For publisher information, see AEHLAU. Volume(issue)/pag e/year: 5,311,1952

Model Applicability

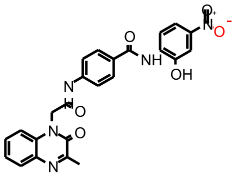
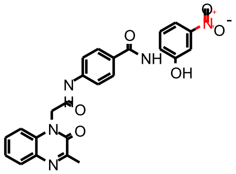
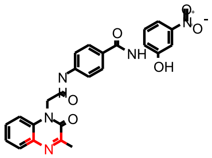
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

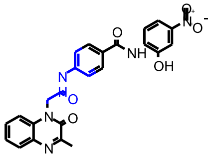
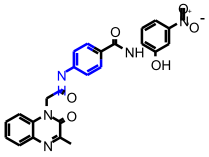
Feature Contribution

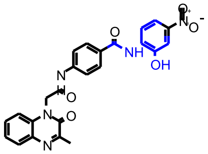
Top features for positive contribution

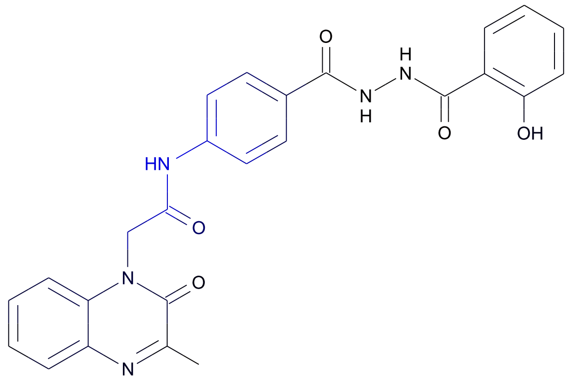
Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set

FCFP_12	5	 [*][O-]	0.085	27 out of 27
FCFP_12	8	 [*][N+](=[*])[*]	0.084	20 out of 20
FCFP_12	580453787	 [*]C(=N[c](:[*]):[*])[*]	0.079	9 out of 9

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set
FCFP_12	1175665944	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	-1.020	2 out of 8
FCFP_12	-1838187238	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	-0.692	5 out of 12

FCFP_12	-1883332927	 [*]C(=[*])N[c]1:[cH]: [cH]:[*]:[cH]:[c]:1O	-0.650	0 out of 1



$C_{25}H_{21}N_5O_5$

Molecular Weight: 471.46474

ALogP: 1.475

Rotatable Bonds: 6

Acceptors: 6

Donors: 4

Model Prediction

Prediction: Non-Irritant

Probability: 0.029

Enrichment: 0.032

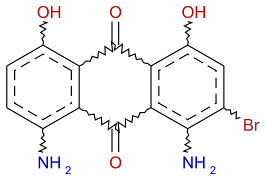
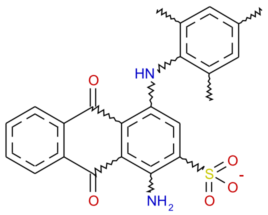
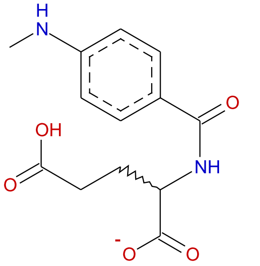
Bayesian Score: -6.546

Mahalanobis Distance: 12.616

Mahalanobis Distance p-value: 8.89e-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	Anthraquinone, 2-bromo-1,8-diamino-4,5-dihydroxy-	2-Anthracenesulfonic acid, 1-amino-9,10-dihydro-9,10-dioxo-4-(2,4,6 -trimethylanilino)-, monosodium salt	Glutamic acid, N-(p-(methylamino)benzoyl)-, sodium salt
Structure			
Actual Endpoint	Irritant	Irritant	Irritant
Predicted Endpoint	Non-Irritant	Non-Irritant	Non-Irritant
Distance	0.876	0.916	0.972
Reference	28ZPAK "Sbornik Vysledku Toxixologickeho Vysetreni Latek A Pripravku," Marhol d, J.V., Institut Pro Vychovu Vedoucicn Pracovniku Chemickeho Prumyclu Praha, Cz echoslovakia, 1972 Volume(issue)/page/year: -,244,1	85JCAE "Prehled Prumyslove Toxikologie; Organicke Latky," Marhold, J., Prague , Czechoslovakia, Avicenum, 1986 Volume(issue)/page/year: -,1327,1986	FCTOD7 Food and Chemical Toxicology. (Pergamon Press Inc., Maxwell House, Fa irview Park, Elmsford, NY 10523) V.20- 1982- Volume(issue)/page/year: 20,563 ,1982

Model Applicability

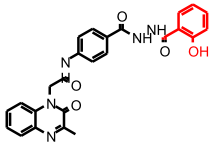
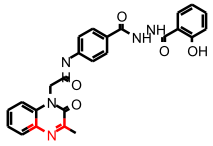
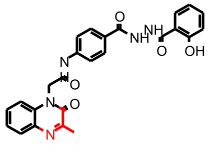
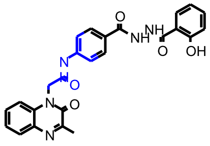
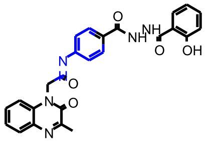
Unknown features are fingerprint features in the query molecule, but not found in the training set.

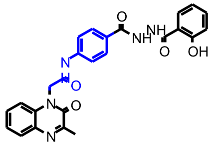
1. All properties and OPS components are within expected ranges.

Feature Contribution

Top features for positive contribution

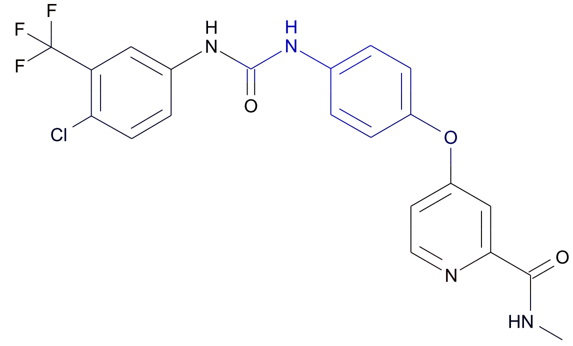
Fingerprint	Bit/Smiles	Feature Structure	Score	Irritant in training set

FCFP_12	-309153329	 [*]C(=[*])[c]1:[cH]:[cH]:[cH]:[cH]:[c]:1O	0.082	13 out of 13
FCFP_12	580453787	 [*]C(=N[c](:[*]):[*])[*]	0.079	9 out of 9
FCFP_12	565968762	 [*]\N=C(\C)/C(=[*])[*]	0.075	78 out of 79
Top Features for negative contribution				
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Irritant in training set**
FCFP_12	1175665944	 [*]CC(=O)N[c](:[cH]:[*]):[cH]:[*]	-1.020	2 out of 8
FCFP_12	-1838187238	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	-0.692	5 out of 12

FCFP_12	-451043714	 [*]CC(=O)N[c]1:[cH]:[cH]:[c]([*]):[cH]:[cH]:1	-0.650	0 out of 1

Sorafenib

TOPKAT_Skin_Irritancy_None_vs_Irritant



$C_{21}H_{16}ClF_3N_4O_3$

Molecular Weight: 464.82494

ALogP: 4.175

Rotatable Bonds: 6

Acceptors: 4

Donors: 3

Model Prediction

Prediction: Non-Irritant

Probability: 0.264

Enrichment: 0.287

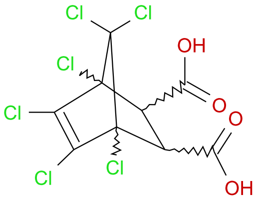
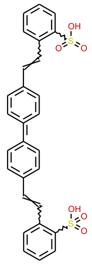
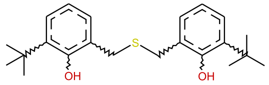
Bayesian Score: -5.229

Mahalanobis Distance: 8.271

Mahalanobis Distance p-value: 0.791

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	**5-Norbornene-2,3-dicarboxylic acid, 1,4,5,6,7,7-hexachloro-**	**Benzenesulfonic acid, 2,2'-(4,4'-biphenylylenedivinylene)di-, disod ium salt**	**Sulfide, bis(4-t-butyl-m-cresyl)-**
Structure			
Actual Endpoint	Irritant	Irritant	Irritant
Predicted Endpoint	Irritant	Non-Irritant	Irritant
Distance	0.844	0.871	0.884
Reference	85JCAE "Prehled Prumyslove Toxikologie; Organicke Latky," Marhold, J., Prague , Czechoslovakia, Avicenum, 1986 Volume(issue)/page/year: -,581,1986	MVCRB3 MVC-Report. (Stockholm, Sweden) No.1-2, 1972-73. Discontinued. Volu me(issue)/page/year: 2,193,1973	AMIHBC AMA Archives of Industrial Hygiene and Occupational Medicine. (Chicago , IL) V.2-10, 1950-54. For publisher information, see AEHLAU. Volume(issue)/pag e/year: 5,311,1952

Model Applicability

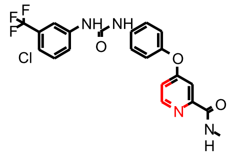
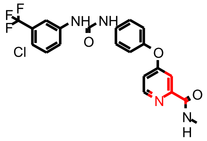
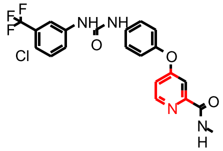
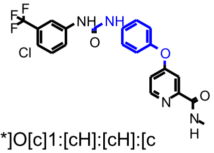
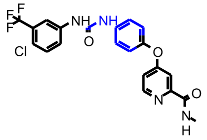
Unknown features are fingerprint features in the query molecule, but not found in the training set.

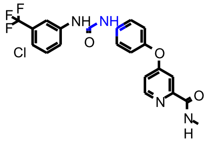
1. All properties and OPS components are within expected ranges.

Feature Contribution

Top features for positive contribution

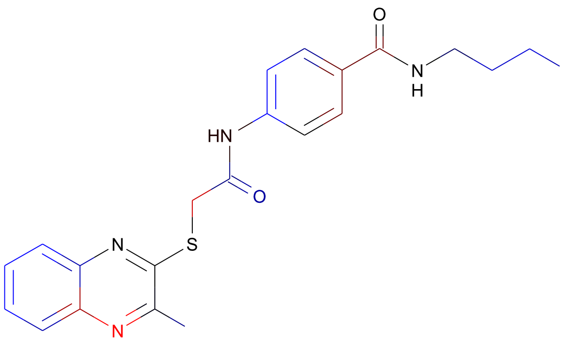
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Irritant in training set**

FCFP_12	-124655670	 [*]:[cH]:[cH]:n:[*]	0.082	13 out of 13
FCFP_12	-1539132615	 [*][c](:[*]):[c](C):n:[*]	0.079	9 out of 9
FCFP_12	-1695756380	 [*][c]1:[*]:[c]([*]):n:[cH]:[cH]:1	0.077	7 out of 7
Top Features for negative contribution				
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**	**Irritant in training set**
FCFP_12	-789307649	 [*]O[c]1:[cH]:[cH]:[c](NC(=[*])[*]):[cH]:[cH]:1	-1.539	0 out of 4
FCFP_12	-1838187238	 [*]C(=[*])N[c]1:[cH]:[cH]:[*]:[cH]:[cH]:1	-0.692	5 out of 12

FCFP_12	1294255210	 [*]C(=[*])N[c](:[*]): [*]	-0.486	12 out of 22

15a

TOPKAT_Carcinogenic_Potency_TD50_Mouse



$C_{22}H_{24}N_4O_2S$

Molecular Weight: 408.51656

ALogP: 3.83

Rotatable Bonds: 8

Acceptors: 5

Donors: 2

Model Prediction

Prediction: 40.565

Unit: mg/kg_body_weight/day

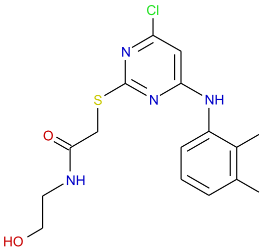
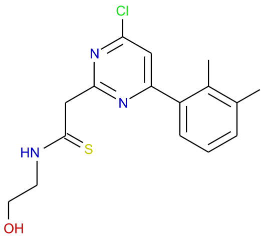
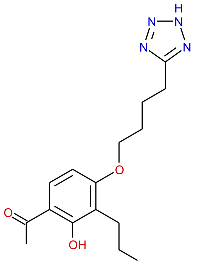
Mahalanobis Distance: 12.772

Mahalanobis Distance p-value: 4.23e-007

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	**4-Chloro-6-(2,3-xylidino)-2-pyri-mi-dinylthio(N-b-hydroxy-ethyl) acetamide**	**455**	**Compound LY171883**
Structure			
Actual Endpoint (-log C)	3.91517	3.87681	3.45372
Predicted Endpoint (-log C)	3.92186	3.77582	2.84749
Distance	0.664	0.694	0.694
Reference	CPDB	CPDB	CPDB

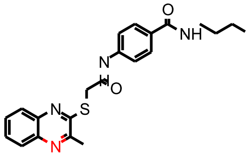
Model Applicability

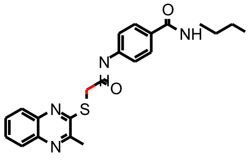
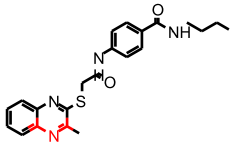
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

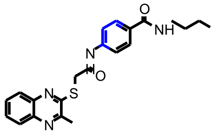
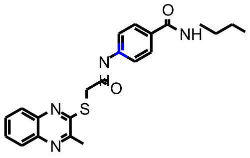
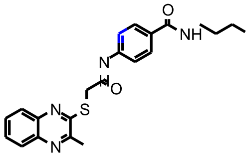
Feature Contribution

Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
ECFP_6	655739385	 [*]:n:[*]	0.229

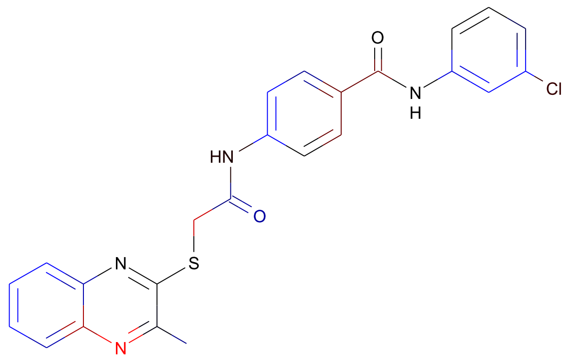
ECFP_6	1559650422	 [*]C[*]	0.203
ECFP_6	834876373	 [*][c](:[*]):n:[c](:[*]):[*]	0.163

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	1996767644	 [*][c](:[*]):[cH]:[cH]:[*]	-0.251
ECFP_6	642810091	 [*][c](:[*]):[*]	-0.247
ECFP_6	-182236392	 [*]:[cH]:[*]	-0.232

15b

TOPKAT_Carcinogenic_Potency_TD50_Mouse



$C_{24}H_{19}ClN_4O_2S$

Molecular Weight: 462.95126

ALogP: 4.743

Rotatable Bonds: 6

Acceptors: 5

Donors: 2

Model Prediction

Prediction: 9.366

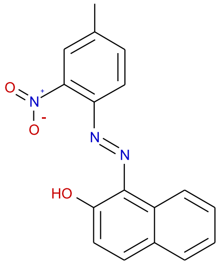
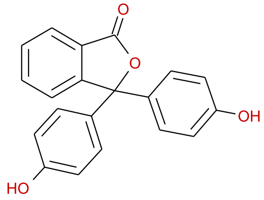
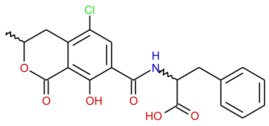
Unit: mg/kg_body_weight/day

Mahalanobis Distance: 14.386

Mahalanobis Distance p-value: 2.76e-011

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	**646**	**Phenolphthalein**	**Ochratoxin A**
Structure			
Actual Endpoint (-log C)	0.937339	2.43468	4.79932
Predicted Endpoint (-log C)	3.26294	3.66084	3.6353
Distance	0.799	0.822	0.835
Reference	CPDB	CPDB	CPDB

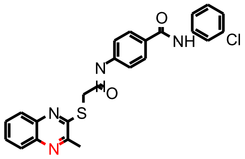
Model Applicability

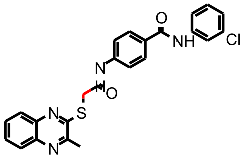
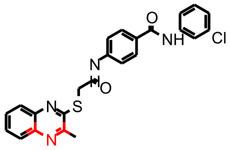
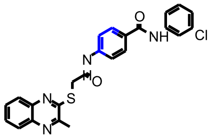
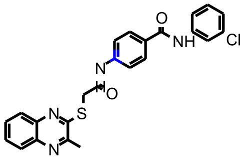
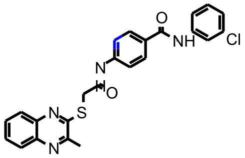
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

Feature Contribution

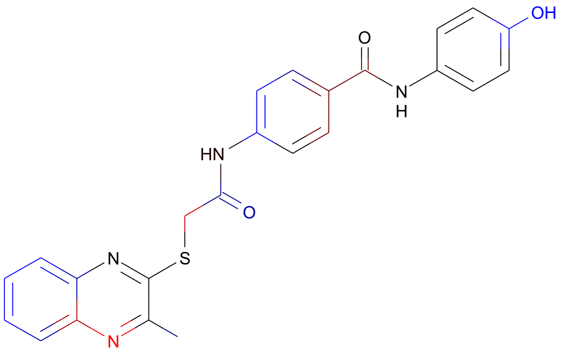
Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
ECFP_6	655739385	 [*]:n:[*]	0.229

ECFP_6	1559650422	 [*]C[*]	0.203
ECFP_6	834876373	 [*][c](:[*]):n:[c](:[*]):[*]	0.163
Top Features for negative contribution			
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
ECFP_6	1996767644	 [*][c](:[*]):[cH]:[cH]:[*]	-0.251
ECFP_6	642810091	 [*][c](:[*]):[*]	-0.247
ECFP_6	-182236392	 [*]:[cH]:[*]	-0.232

15c

TOPKAT_Carcinogenic_Potency_TD50_Mouse



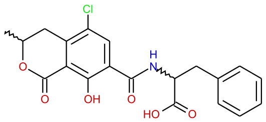
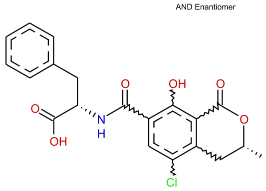
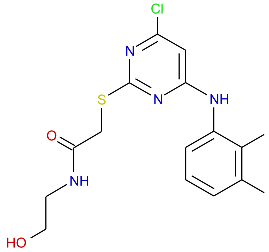
$C_{24}H_{20}N_4O_3S$
Molecular Weight: 444.5056
ALogP: 3.836
Rotatable Bonds: 6
Acceptors: 6
Donors: 3

Model Prediction

Prediction: 17.665
Unit: mg/kg_body_weight/day
Mahalanobis Distance: 13.273
Mahalanobis Distance p-value: 2.6e-008

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	**Ochratoxin A**	**542**	**4-Chloro-6-(2,3-xylidino)-2-pyri-mi-dinylthio(N-b-hydroxy-ethyl) acetamide**
Structure		 AND Enantiomer	
Actual Endpoint (-log C)	4.79932	4.79932	3.91517
Predicted Endpoint (-log C)	3.6353	3.6353	3.92186
Distance	0.793	0.793	0.829
Reference	CPDB	CPDB	CPDB

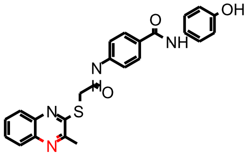
Model Applicability

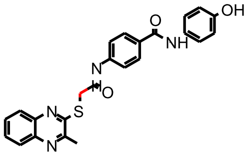
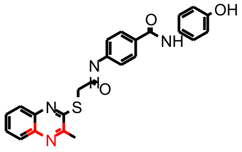
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

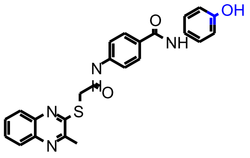
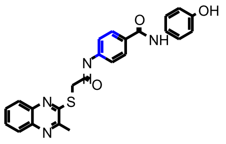
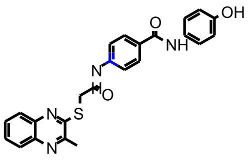
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	655739385	 [*]:n:[*]	0.229

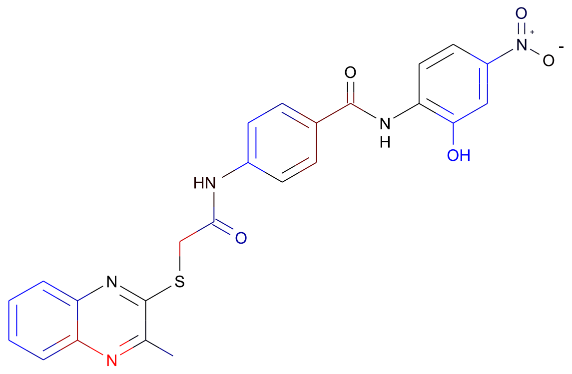
ECFP_6	1559650422	 [*]C[*]	0.203
ECFP_6	834876373	 [*][c](:[*]):n:[c](:[*]):[*]	0.163

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	2019062761	 [*]:[c](:[*])O	-0.258
ECFP_6	1996767644	 [*][c](:[*]):[cH]:[cH]:[*]	-0.251
ECFP_6	642810091	 [*][c](:[*]):[*]	-0.247

15d

TOPKAT_Carcinogenic_Potency_TD50_Mouse



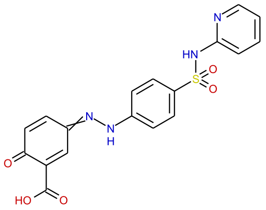
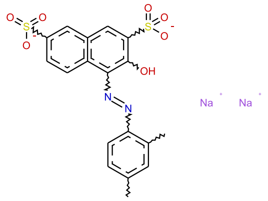
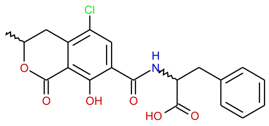
$C_{24}H_{19}N_5O_5S$
Molecular Weight: 489.50316
ALogP: 3.731
Rotatable Bonds: 7
Acceptors: 8
Donors: 3

Model Prediction

Prediction: 31.468
Unit: mg/kg_body_weight/day
Mahalanobis Distance: 14.067
Mahalanobis Distance p-value: 2.14e-010

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	Salicylazosulfapyridine	420	Ochratoxin A
Structure			
Actual Endpoint (-log C)	2.5034	2.78302	4.79932
Predicted Endpoint (-log C)	3.54214	3.31546	3.6353
Distance	0.848	0.901	0.913
Reference	CPDB	CPDB	CPDB

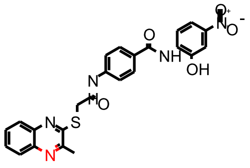
Model Applicability

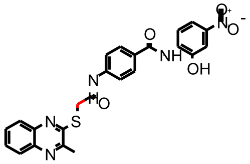
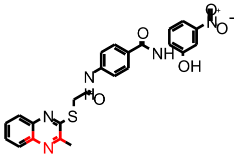
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

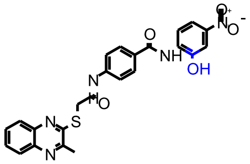
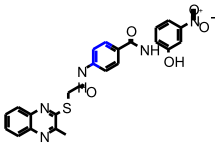
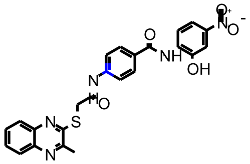
Feature Contribution

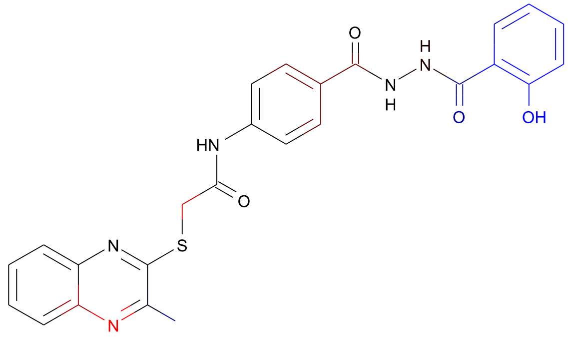
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	655739385	 [*]:n:[*]	0.229

ECFP_6	1559650422	 [*]C[*]	0.203
ECFP_6	834876373	 [*][c](:[*]):n:[c](:[*]):[*]	0.163

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	2019062761	 [*]:[c](:[*])O	-0.258
ECFP_6	1996767644	 [*][c](:[*]):[cH]:[cH]:[*]	-0.251
ECFP_6	642810091	 [*][c](:[*]):[*]	-0.247



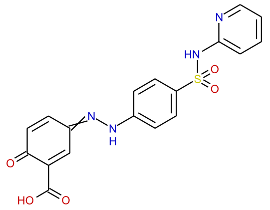
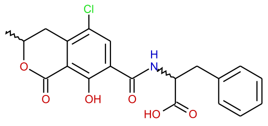
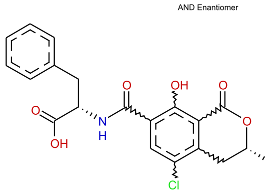
$C_{25}H_{21}N_5O_4S$
Molecular Weight: 487.53034
ALogP: 3.093
Rotatable Bonds: 7
Acceptors: 7
Donors: 4

Model Prediction

Prediction: 16.061
Unit: mg/kg_body_weight/day
Mahalanobis Distance: 13.288
Mahalanobis Distance p-value: 2.39e-008

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	Salicylazosulfapyridine	Ochratoxin A	542
Structure			 AND Enantiomer
Actual Endpoint (-log C)	2.5034	4.79932	4.79932
Predicted Endpoint (-log C)	3.54214	3.6353	3.6353
Distance	0.841	0.870	0.870
Reference	CPDB	CPDB	CPDB

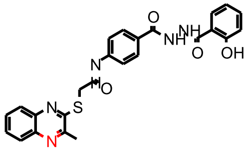
Model Applicability

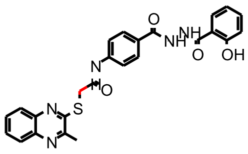
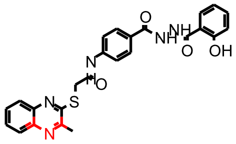
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

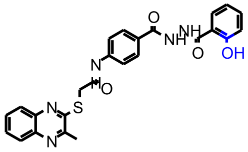
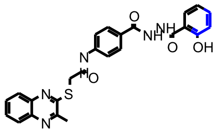
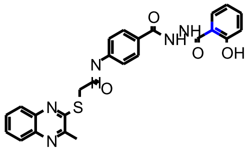
Feature Contribution

Top features for positive contribution

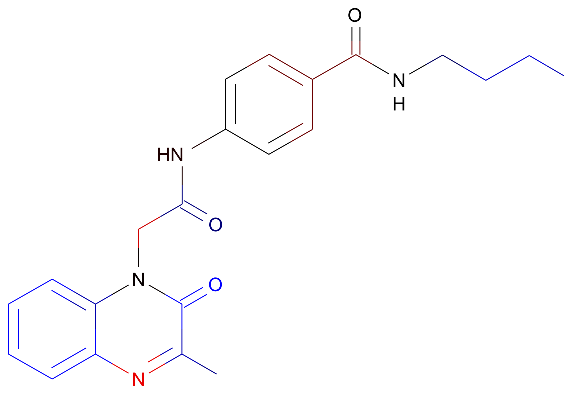
Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	655739385	 [*]:n:[*]	0.229

ECFP_6	1559650422	 [*]C[*]	0.203
ECFP_6	834876373	 [*][c](:[*]):n:[c](:[*]):[*]	0.163

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	2019062761	 [*]:[c](:[*])O	-0.258
ECFP_6	1996767644	 [*][c](:[*]):[cH]:[cH]:[*]	-0.251
ECFP_6	642810091	 [*][c](:[*]):[*]	-0.247

17a



$C_{22}H_{24}N_4O_3$

Molecular Weight: 392.45096

ALogP: 2.212

Rotatable Bonds: 7

Acceptors: 4

Donors: 2

Model Prediction

Prediction: 142.906

Unit: mg/kg_body_weight/day

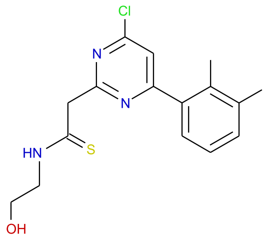
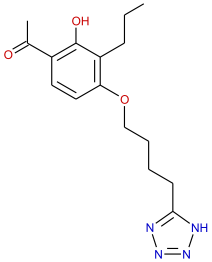
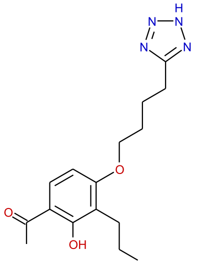
Mahalanobis Distance: 12.732

Mahalanobis Distance p-value: 5.24e-007

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	455	832	Compound LY171883
Structure			
Actual Endpoint (-log C)	3.87681	3.45372	3.45372
Predicted Endpoint (-log C)	3.77582	2.80429	2.84749
Distance	0.635	0.676	0.682
Reference	CPDB	CPDB	CPDB

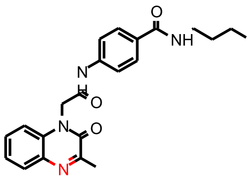
Model Applicability

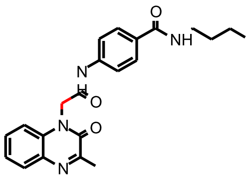
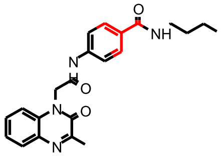
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. OPS PC20 out of range. Value: 3.8775. Training min, max, SD, explained variance: -4.3384, 3.4394, 1.14, 0.0162.
2. Unknown ECFP_2 feature: -597295171: [*][c](:[*]):[c](N=[*]):c:[*]
3. Unknown ECFP_2 feature: -1236953626: [*]N([*])[c](:[c]([*]):[*]):c:[*]
4. Unknown ECFP_2 feature: 1483289300: [*]\N=C(\C)/C(=[*])[*]
5. Unknown ECFP_2 feature: -661097313: [*]CN(C(=[*])[*])[c](:[*]):[*]
6. Unknown ECFP_2 feature: -37698365: [*]N([*])CC(=[*])[*]

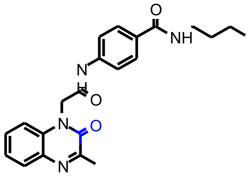
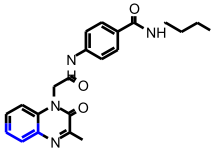
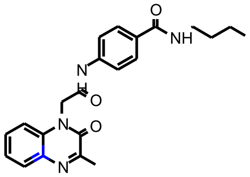
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	655739385	 [*]:n:[*]	0.229

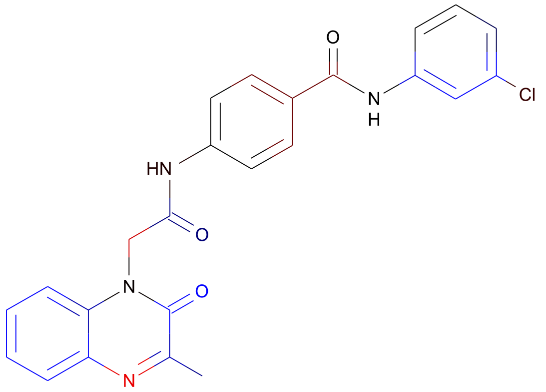
ECFP_6	1559650422	 [*]C[*]	0.203
ECFP_6	-175146122	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	0.107

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	2106656448	 [*]C(=O)[*]	-0.275
ECFP_6	1996767644	 [*][c](:[*]):[cH]:[cH]:[*]	-0.251
ECFP_6	642810091	 [*][c](:[*]):[*]	-0.247

17b

TOPKAT_Carcinogenic_Potency_TD50_Mouse



$C_{24}H_{19}ClN_4O_3$

Molecular Weight: 446.88566

ALogP: 3.125

Rotatable Bonds: 5

Acceptors: 4

Donors: 2

Model Prediction

Prediction: 33.153

Unit: mg/kg_body_weight/day

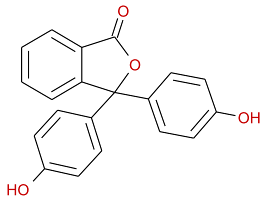
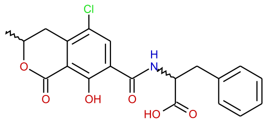
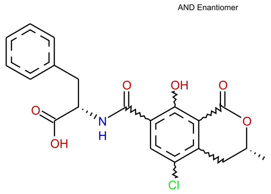
Mahalanobis Distance: 14.301

Mahalanobis Distance p-value: 4.8e-011

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	Phenolphthalein	Ochratoxin A	542
Structure			 AND Enantiomer
Actual Endpoint (-log C)	2.43468	4.79932	4.79932
Predicted Endpoint (-log C)	3.66084	3.6353	3.6353
Distance	0.700	0.746	0.746
Reference	CPDB	CPDB	CPDB

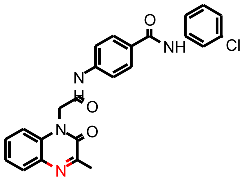
Model Applicability

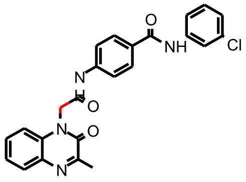
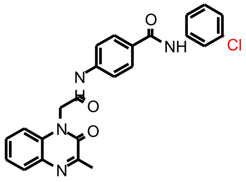
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. OPS PC20 out of range. Value: 4.0307. Training min, max, SD, explained variance: -4.3384, 3.4394, 1.14, 0.0162.
2. Unknown ECFP_2 feature: -597295171: [*][c](:[*]):[c](N=[*]):c:[*]
3. Unknown ECFP_2 feature: -1236953626: [*]N([*])[c](:[c]([*]):[*]):c:[*]
4. Unknown ECFP_2 feature: 1483289300: [*]\N=C(\C)/C(=[*])[*]
5. Unknown ECFP_2 feature: -661097313: [*]CN(C(=[*])[*])[c](:[*]):[*]
6. Unknown ECFP_2 feature: -37698365: [*]N([*])CC(=[*])[*]

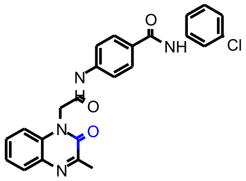
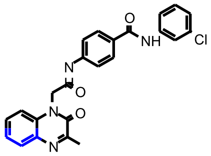
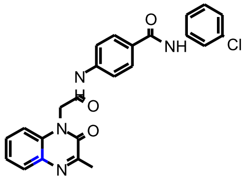
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	655739385	 [*]:n:[*]	0.229

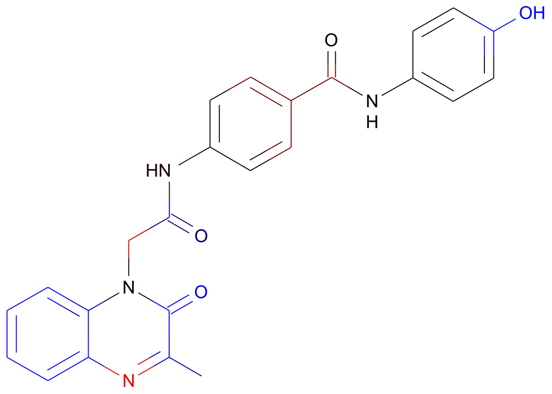
ECFP_6	1559650422	 [*]C[*]	0.203
ECFP_6	-817402818	 [*]Cl	0.129

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	2106656448	 [*]C(=O)[*]	-0.275
ECFP_6	1996767644	 [*][c](:[*]):[cH]:[cH]:[*]	-0.251
ECFP_6	642810091	 [*][c](:[*]):[*]	-0.247

17c

TOPKAT_Carcinogenic_Potency_TD50_Mouse



$C_{24}H_{20}N_{4}O_{4}$

Molecular Weight: 428.44

ALogP: 2.218

Rotatable Bonds: 5

Acceptors: 5

Donors: 3

Model Prediction

Prediction: 62.436

Unit: mg/kg_body_weight/day

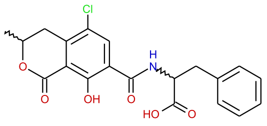
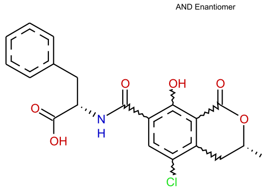
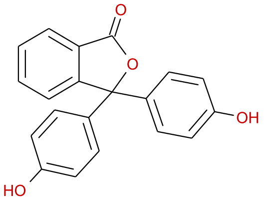
Mahalanobis Distance: 13.204

Mahalanobis Distance p-value: 3.87e-008

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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

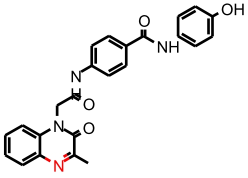
Name	Ochratoxin A	542	Phenolphthalein
Structure		 AND Enantiomer	
Actual Endpoint (-log C)	4.79932	4.79932	2.43468
Predicted Endpoint (-log C)	3.6353	3.6353	3.66084
Distance	0.668	0.668	0.755
Reference	CPDB	CPDB	CPDB

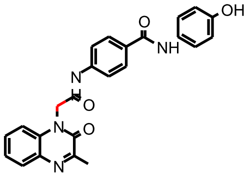
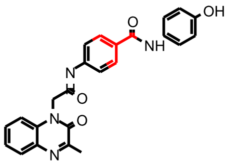
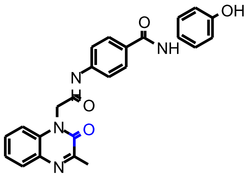
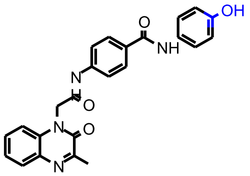
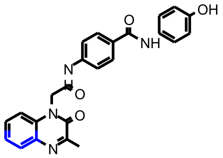
Model Applicability

Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_2 feature: -597295171: [*][c](:[*]):[c](N=[*]):c:[*]
3. Unknown ECFP_2 feature: -1236953626: [*]N([*])[c](:[c]([*]):[*]):c:[*]
4. Unknown ECFP_2 feature: 1483289300: [*]\N=C(\C)/C(=[*])[*]
5. Unknown ECFP_2 feature: -661097313: [*]CN(C(=[*])[*])[c](:[*]):[*]
6. Unknown ECFP_2 feature: -37698365: [*]N([*])CC(=[*])[*]

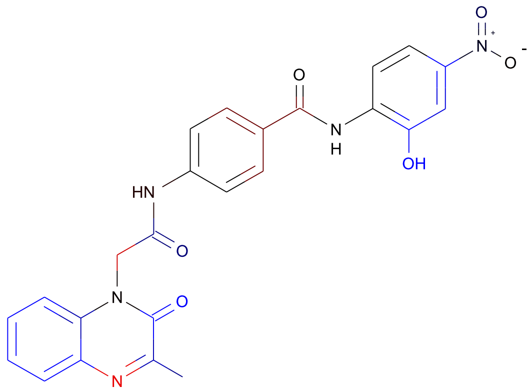
Feature Contribution

Top features for positive contribution			
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
ECFP_6	655739385	 [*]:n:[*]	0.229

ECFP_6	1559650422	 [*]C[*]	0.203
ECFP_6	-175146122	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	0.107
Top Features for negative contribution			
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
ECFP_6	2106656448	 [*]C(=O)[*]	-0.275
ECFP_6	2019062761	 [*]:[c](:[*])O	-0.258
ECFP_6	1996767644	 [*][c](:[*]):[cH]:[cH]:[*]	-0.251

17d

TOPKAT_Carcinogenic_Potency_TD50_Mouse



$C_{24}H_{19}N_5O_6$

Molecular Weight: 473.43756

ALogP: 2.113

Rotatable Bonds: 6

Acceptors: 7

Donors: 3

Model Prediction

Prediction: 111.609

Unit: mg/kg_body_weight/day

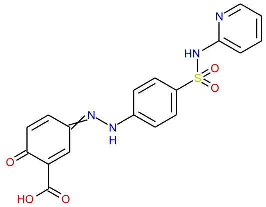
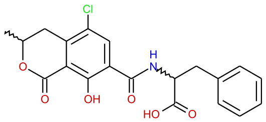
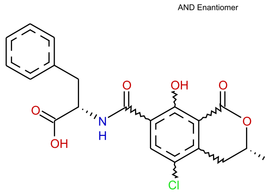
Mahalanobis Distance: 14.728

Mahalanobis Distance p-value: 2.85e-012

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	Salicylazosulfapyridine	Ochratoxin A	542
Structure			 AND Enantiomer
Actual Endpoint (-log C)	2.5034	4.79932	4.79932
Predicted Endpoint (-log C)	3.54214	3.6353	3.6353
Distance	0.678	0.743	0.743
Reference	CPDB	CPDB	CPDB

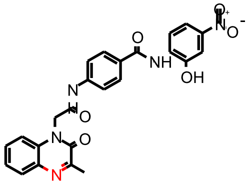
Model Applicability

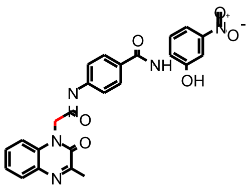
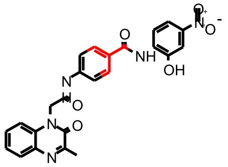
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_2 feature: -597295171: [*][c](:[*]):[c](N=[*]):c:[*]
3. Unknown ECFP_2 feature: -1236953626: [*]N([*])[c](:[c]([*]):[*]):c:[*]
4. Unknown ECFP_2 feature: 1483289300: [*]\N=C(\C)/C(=[*])[*]
5. Unknown ECFP_2 feature: -661097313: [*]CN(C(=[*])[*])[c](:[*]):[*]
6. Unknown ECFP_2 feature: -37698365: [*]N([*])CC(=[*])[*]

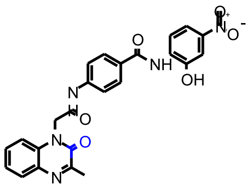
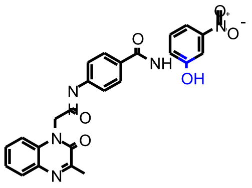
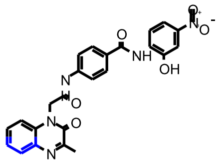
Feature Contribution

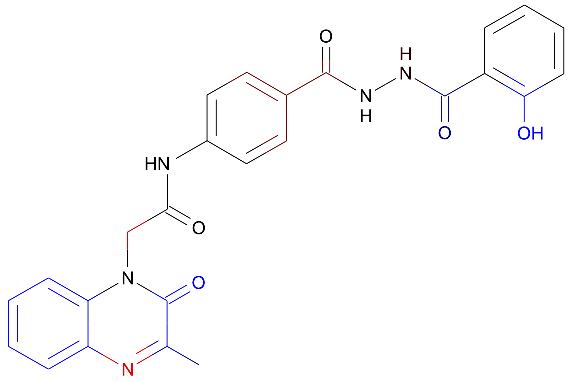
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	655739385	 [*]:n:[*]	0.229

ECFP_6	1559650422	 [*]C[*]	0.203
ECFP_6	-175146122	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	0.107

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	2106656448	 [*]C(=O)[*]	-0.275
ECFP_6	2019062761	 [*]:[c](:[*])O	-0.258
ECFP_6	1996767644	 [*][c](:[*]):[cH]:[cH]:[*]	-0.251



$C_{25}H_{21}N_5O_5$

Molecular Weight: 471.46474

ALogP: 1.475

Rotatable Bonds: 6

Acceptors: 6

Donors: 4

Model Prediction

Prediction: 56.955

Unit: mg/kg_body_weight/day

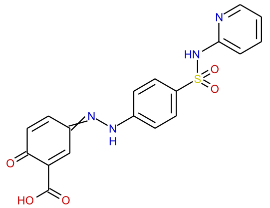
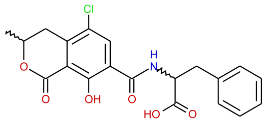
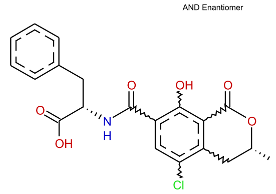
Mahalanobis Distance: 13.274

Mahalanobis Distance p-value: 2.59e-008

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	**Salicylazosulfapyridine**	**Ochratoxin A**	**542**
Structure			 AND Enantiomer
Actual Endpoint (-log C)	2.5034	4.79932	4.79932
Predicted Endpoint (-log C)	3.54214	3.6353	3.6353
Distance	0.732	0.736	0.736
Reference	CPDB	CPDB	CPDB

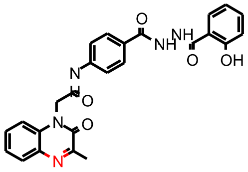
Model Applicability

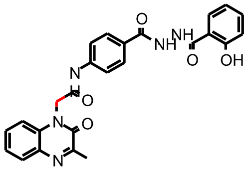
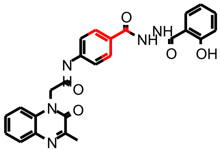
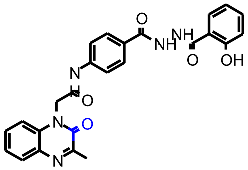
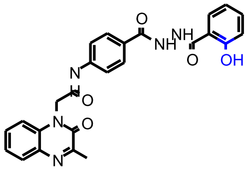
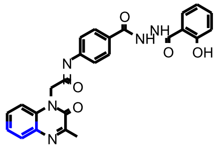
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_2 feature: -597295171: [*][c](:[*]):[c](N=[*]):c:[*]
3. Unknown ECFP_2 feature: -1236953626: [*]N([*])[c](:[c]([*]):[*]):c:[*]
4. Unknown ECFP_2 feature: 1483289300: [*]\N=C(\C)/C(=[*])[*]
5. Unknown ECFP_2 feature: -661097313: [*]CN(C(=[*])[*])[c](:[*]):[*]
6. Unknown ECFP_2 feature: -37698365: [*]N([*])CC(=[*])[*]

Feature Contribution

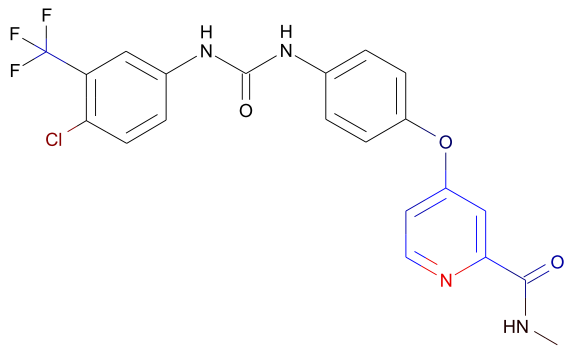
Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
ECFP_6	655739385	 [*]:n:[*]	0.229

<table>
<tr><td>ECFP_6</td><td>1559650422</td><td>
[*]C[*]</td><td>0.203</td></tr>
<tr><td>ECFP_6</td><td>-175146122</td><td>
[*]C(=[*])[c](:[cH]:[*]):[cH]:[*]</td><td>0.107</td></tr>
<tr><th colspan="4">Top Features for negative contribution</th></tr>
<tr><th>Fingerprint</th><th>Bit/Smiles</th><th>Feature Structure</th><th>Score</th></tr>
<tr><td>ECFP_6</td><td>2106656448</td><td>
[*]C(=O)[*]</td><td>-0.275</td></tr>
<tr><td>ECFP_6</td><td>2019062761</td><td>
[*]:[c](:[*])O</td><td>-0.258</td></tr>
<tr><td>ECFP_6</td><td>1996767644</td><td>
[*][c](:[*]):[cH]:[cH]:[*]</td><td>-0.251</td></tr>
</table>

Sorafenib

TOPKAT_Carcinogenic_Potency_TD50_Mouse



$C_{21}H_{16}ClF_3N_4O_3$

Molecular Weight: 464.82494

ALogP: 4.175

Rotatable Bonds: 6

Acceptors: 4

Donors: 3

Model Prediction

Prediction: 19.236

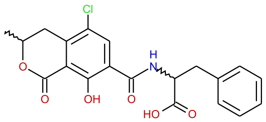
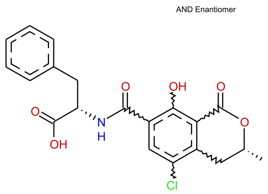
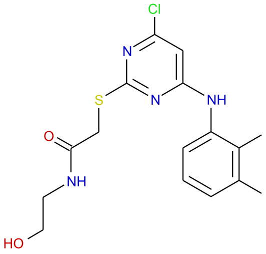
Unit: mg/kg_body_weight/day

Mahalanobis Distance: 12.401

Mahalanobis Distance p-value: 2.94e-006

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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

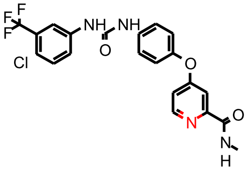
Name	**Ochratoxin A**	**542**	**4-Chloro-6-(2,3-xylidino)-2-pyri-mi-dinylthio(N-b-hydroxy-ethyl) acetamide**
Structure			
Actual Endpoint (-log C)	4.79932	4.79932	3.91517
Predicted Endpoint (-log C)	3.6353	3.6353	3.92186
Distance	0.718	0.718	0.738
Reference	CPDB	CPDB	CPDB

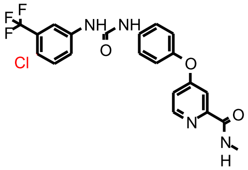
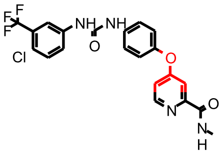
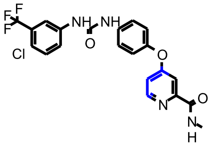
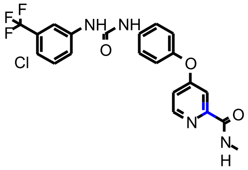
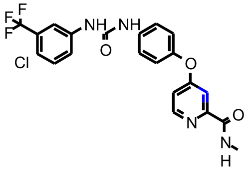
Model Applicability

Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_2 feature: 1413420509: [*]C(=[*])[c](:n:[*]):c:[*]
3. Unknown ECFP_2 feature: 1338334141: [*]C(=[*])NC

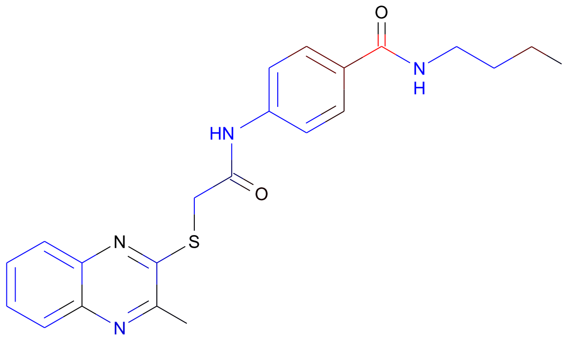
Feature Contribution

Top features for positive contribution			
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
ECFP_6	655739385	 [*]:n:[*]	0.229

ECFP_6	-817402818	 [*]Cl	0.129
ECFP_6	-176455838	 [*]O[c](:[cH]:[*]):[cH]:[*]	0.082
Top Features for negative contribution			
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
ECFP_6	1996767644	 [*][c](:[*]):[cH]:[cH]:[*]	-0.251
ECFP_6	642810091	 [*][c](:[*]):[*]	-0.247
ECFP_6	-182236392	 [*]:[cH]:[*]	-0.232

15a

TOPKAT_Carcinogenic_Potency_TD50_Rat



$C_{22}H_{24}N_4O_2S$

Molecular Weight: 408.51656

ALogP: 3.83

Rotatable Bonds: 8

Acceptors: 5

Donors: 2

Model Prediction

Prediction: 187.702

Unit: mg/kg_body_weight/day

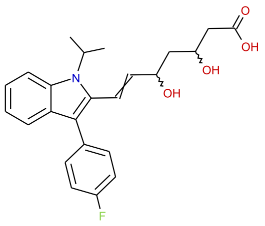
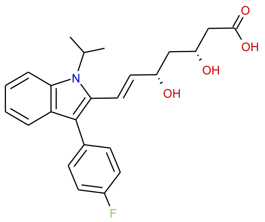
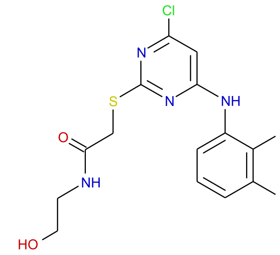
Mahalanobis Distance: 17.700

Mahalanobis Distance p-value: 3.71e-020

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	Fluvastatin	913	4-Chloro-6-(2,3-xylidino)-2-pyri-mi-dinylthio(N-b-hydroxy-ethyl) acetamide
Structure			
Actual Endpoint (-log C)	3.51742	3.51742	4.75226
Predicted Endpoint (-log C)	5.41573	5.41573	3.29421
Distance	0.593	0.593	0.633
Reference	CPDB	CPDB	CPDB

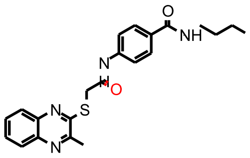
Model Applicability

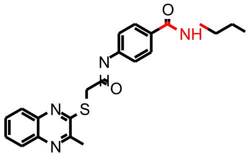
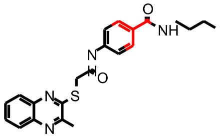
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

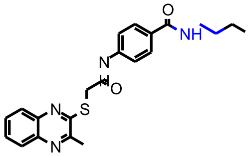
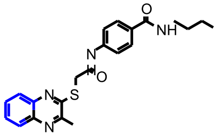
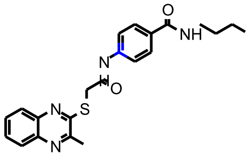
Feature Contribution

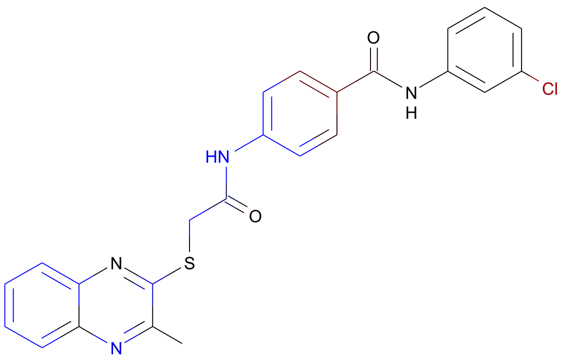
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	1	 [*]=O	0.234

FCFP_6	-885550502	 [*]CNC(=[*])[*]	0.229
FCFP_6	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	0.137

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_6	-1272709286	 [*]CCN[*]	-0.526
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.422
FCFP_6	16	 [*][c](:[*]):[*]	-0.354



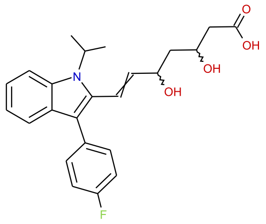
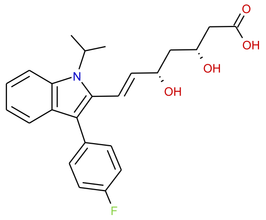
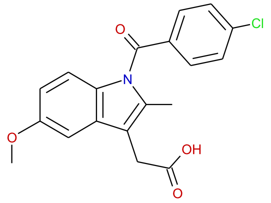
$C_{24}H_{19}ClN_4O_2S$
Molecular Weight: 462.95126
ALogP: 4.743
Rotatable Bonds: 6
Acceptors: 5
Donors: 2

Model Prediction

Prediction: 25.425
Unit: mg/kg_body_weight/day
Mahalanobis Distance: 16.964
Mahalanobis Distance p-value: 1.58e-017

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	Fluvastatin	913	Indomethacin
Structure			
Actual Endpoint (-log C)	3.51742	3.51742	5.49293
Predicted Endpoint (-log C)	5.41573	5.41573	4.9569
Distance	0.672	0.672	0.711
Reference	CPDB	CPDB	CPDB

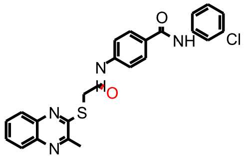
Model Applicability

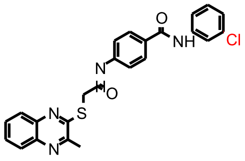
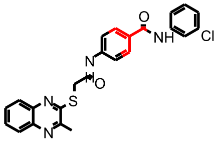
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. OPS PC16 out of range. Value: -3.695. Training min, max, SD, explained variance: -3.6111, 5.345, 1.406, 0.0190.

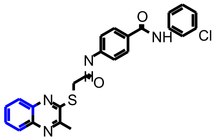
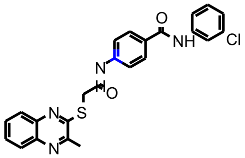
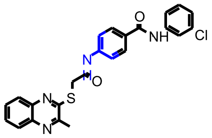
Feature Contribution

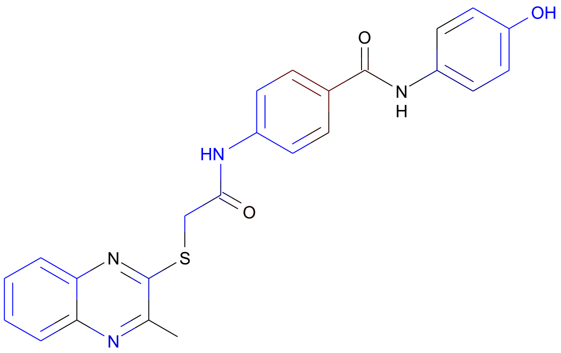
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	1	 [*]=O	0.234

FCFP_6	32	 [*]Cl	0.154
FCFP_6	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	0.137

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.422
FCFP_6	16	 [*][c](:[*]):[*]	-0.354
FCFP_6	590925877	 [*]N[c](:[cH]:[*]):[cH]:[*]	-0.323



$C_{24}H_{20}N_4O_3S$

Molecular Weight: 444.5056

ALogP: 3.836

Rotatable Bonds: 6

Acceptors: 6

Donors: 3

Model Prediction

Prediction: 210.252

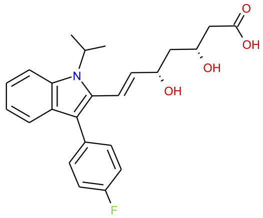
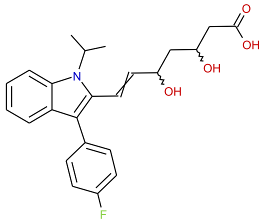
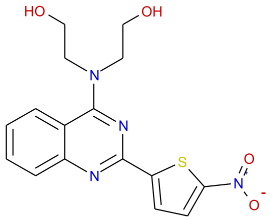
Unit: mg/kg_body_weight/day

Mahalanobis Distance: 16.933

Mahalanobis Distance p-value: 2.03e-017

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	**913**	**Fluvastatin**	**4-Bis(2-hydroxyethyl)amino-2-(5-nitro-2-thienyl)quinazoline**
Structure			
Actual Endpoint (-log C)	3.51742	3.51742	5.05984
Predicted Endpoint (-log C)	5.41573	5.41573	4.23808
Distance	0.709	0.709	0.716
Reference	CPDB	CPDB	CPDB

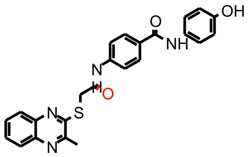
Model Applicability

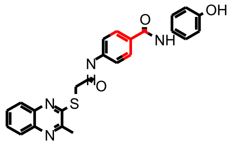
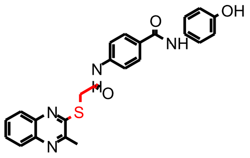
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

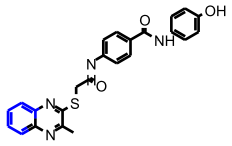
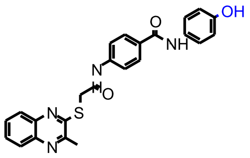
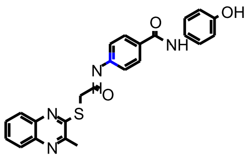
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	1	 [*]=O	0.234

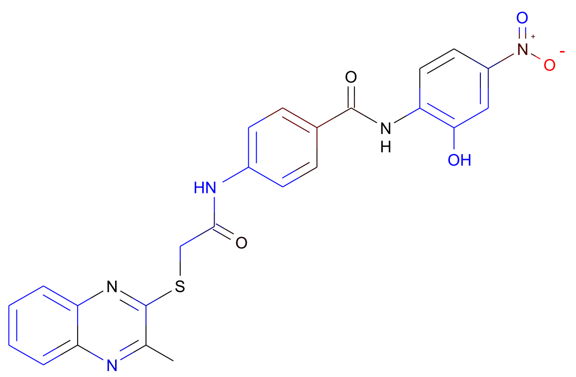
FCFP_6	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	0.137
FCFP_6	-1272768868	 [*]SCC(=[*])[*]	0.127

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.422
FCFP_6	7	 [*]O	-0.372
FCFP_6	16	 [*][c](:[*]):[*]	-0.354

15d

TOPKAT_Carcinogenic_Potency_TD50_Rat



$C_{24}H_{19}N_5O_5S$

Molecular Weight: 489.50316

ALogP: 3.731

Rotatable Bonds: 7

Acceptors: 8

Donors: 3

Model Prediction

Prediction: 76.496

Unit: mg/kg_body_weight/day

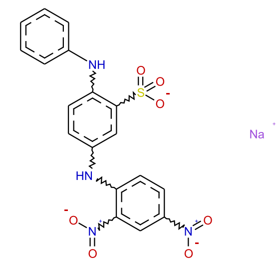
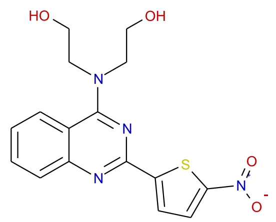
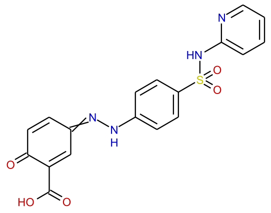
Mahalanobis Distance: 20.119

Mahalanobis Distance p-value: 1.51e-029

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	**623**	**4-Bis(2-hydroxyethyl)amino-2-(5-nitro-2-thienyl)quinazoline**	**Salicylazosulfapyridine**
Structure			
Actual Endpoint (-log C)	2.39985	5.05984	2.39891
Predicted Endpoint (-log C)	3.4177	4.23808	3.17598
Distance	0.673	0.749	0.786
Reference	CPDB	CPDB	CPDB

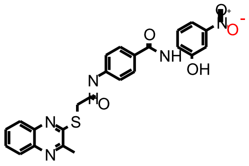
Model Applicability

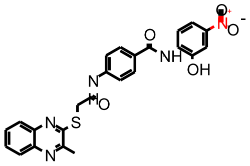
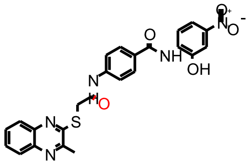
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

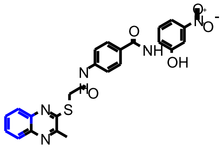
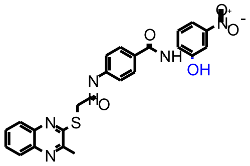
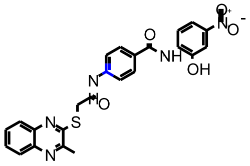
Feature Contribution

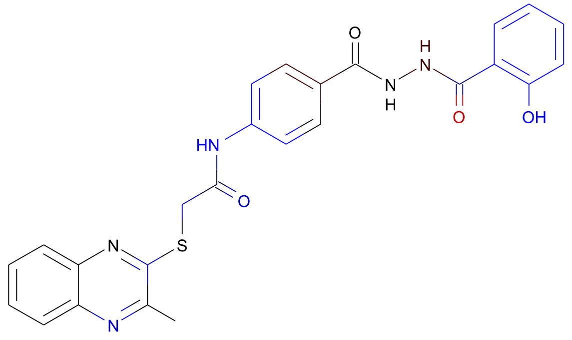
Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_6	5	 [*][O-]	0.431

FCFP_6	8	 [*][N+](=[*])[*]	0.336
FCFP_6	1	 [*]=O	0.234

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.422
FCFP_6	7	 [*]O	-0.372
FCFP_6	16	 [*][c](:[*]):[*]	-0.354



$C_{25}H_{21}N_5O_4S$

Molecular Weight: 487.53034

ALogP: 3.093

Rotatable Bonds: 7

Acceptors: 7

Donors: 4

Model Prediction

Prediction: 249.201

Unit: mg/kg_body_weight/day

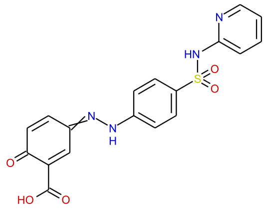
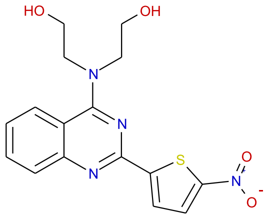
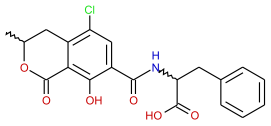
Mahalanobis Distance: 17.086

Mahalanobis Distance p-value: 5.91e-018

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	Salicylazosulfapyridine	4-Bis(2-hydroxyethyl)amino-2-(5-nitro-2-thienyl)quinazoline	Ochratoxin A
Structure			
Actual Endpoint (-log C)	2.39891	5.05984	6.47264
Predicted Endpoint (-log C)	3.17598	4.23808	5.06501
Distance	0.779	0.798	0.814
Reference	CPDB	CPDB	CPDB

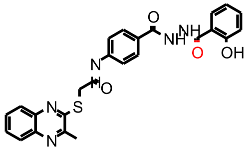
Model Applicability

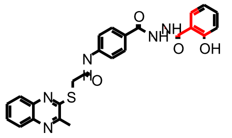
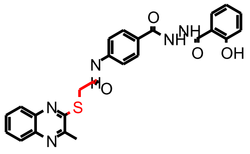
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

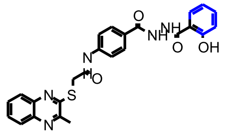
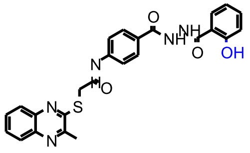
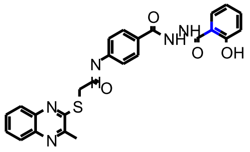
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	1	 [*]=O	0.234

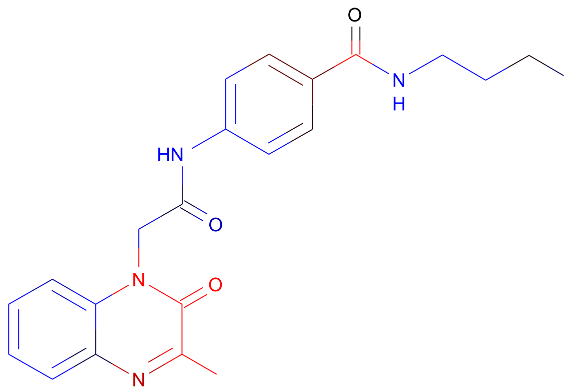
FCFP_6	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	0.137
FCFP_6	-1272768868	 [*]SCC(=[*])[*]	0.127

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.422
FCFP_6	7	 [*]O	-0.372
FCFP_6	16	 [*][c](:[*]):[*]	-0.354

17a

TOPKAT_Carcinogenic_Potency_TD50_Rat



$C_{22}H_{24}N_4O_3$

Molecular Weight: 392.45096

ALogP: 2.212

Rotatable Bonds: 7

Acceptors: 4

Donors: 2

Model Prediction

Prediction: 51.930

Unit: mg/kg_body_weight/day

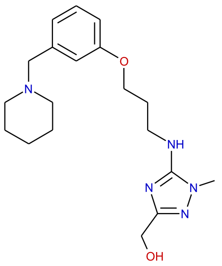
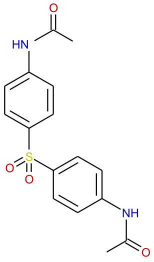
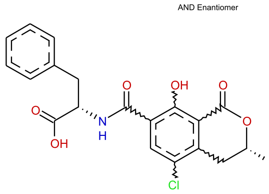
Mahalanobis Distance: 17.563

Mahalanobis Distance p-value: 1.17e-019

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	**Loxtidine**	**4,4´-Sulfonylbisacetanilide**	**542**
Structure			
Actual Endpoint (-log C)	2.87532	3.77655	6.59334
Predicted Endpoint (-log C)	3.63996	3.55337	5.06501
Distance	0.593	0.599	0.663
Reference	CPDB	CPDB	CPDB

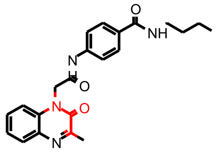
Model Applicability

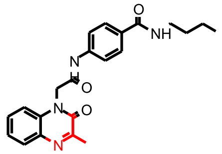
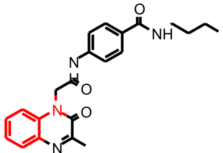
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

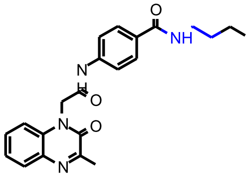
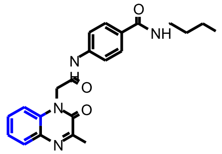
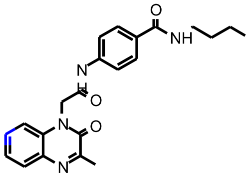
Feature Contribution

Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_6	565998553	 [*]N([*])C(=O)C(=[*]) [*]	0.357

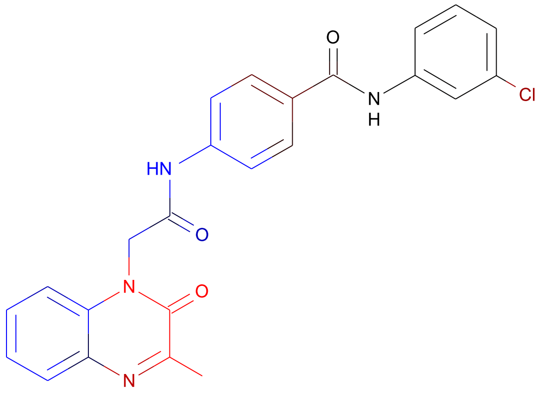
FCFP_6	565968762	 [*]\N=C(\C)/C(=[*])[*]	0.266
FCFP_6	-2090462286	 [*]N([*])[c]1:[cH]:[cH]:[cH]:[cH]:[c]:1[*]	0.245

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_6	-1272709286	 [*]CCN[*]	-0.526
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.422
FCFP_6	16	 [*][c](:[*]):[*]	-0.354

17b

TOPKAT_Carcinogenic_Potency_TD50_Rat



$C_{24}H_{19}ClN_4O_3$

Molecular Weight: 446.88566

ALogP: 3.125

Rotatable Bonds: 5

Acceptors: 4

Donors: 2

Model Prediction

Prediction: 6.023

Unit: mg/kg_body_weight/day

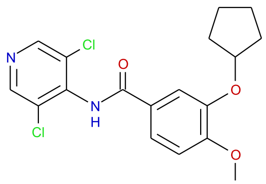
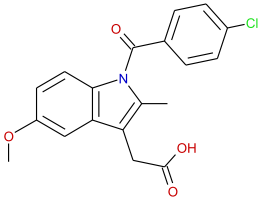
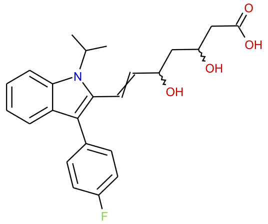
Mahalanobis Distance: 16.376

Mahalanobis Distance p-value: 1.61e-015

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	**3-(Cyclopentyloxy)-N-(3,5-di-chloro-4-pyridyl)-4-methoxy-benzamide**	**Indomethacin**	**Fluvastatin**
Structure			
Actual Endpoint (-log C)	5.39369	5.49293	3.51742
Predicted Endpoint (-log C)	4.27874	4.9569	5.41573
Distance	0.682	0.683	0.687
Reference	CPDB	CPDB	CPDB

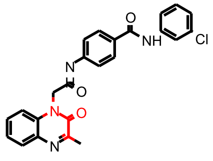
Model Applicability

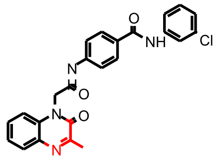
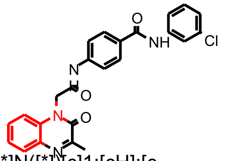
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. OPS PC20 out of range. Value: 6.7989. Training min, max, SD, explained variance: -3.9266, 5.5565, 1.236, 0.0147.

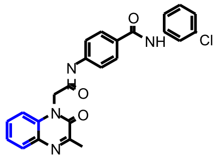
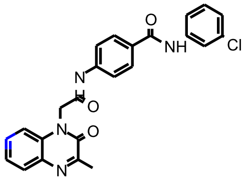
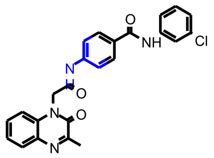
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	565998553	 [*]N([*])C(=O)C(=[*]) [*]	0.357

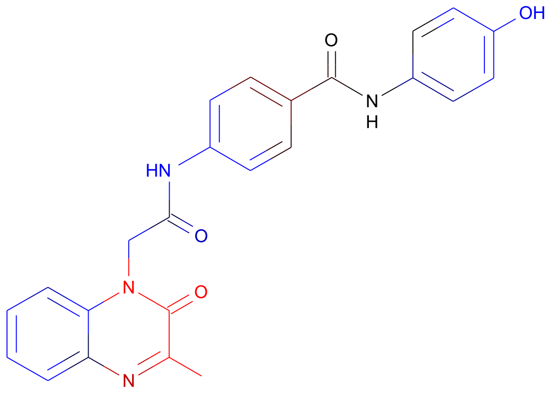
FCFP_6	565968762	 [*]\N=C(\C)/C(=[*])[*]	0.266
FCFP_6	-2090462286	 [*]N([*])[c]1:[cH]:[cH]:[cH]:[cH]:[c]:1[*]	0.245

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.422
FCFP_6	16	 [*][c](:[*]):[*]	-0.354
FCFP_6	590925877	 [*]N[c](:[cH]:[*]):[cH]:[*]	-0.323

17c

TOPKAT_Carcinogenic_Potency_TD50_Rat



$C_{24}H_{20}N_4O_4$
Molecular Weight: 428.44
ALogP: 2.218
Rotatable Bonds: 5
Acceptors: 5
Donors: 3

Model Prediction

Prediction: 49.731

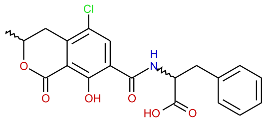
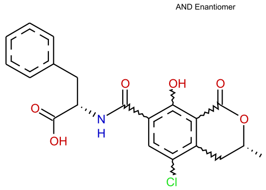
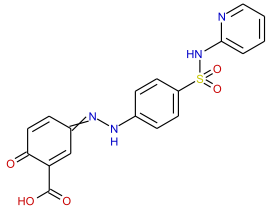
Unit: mg/kg_body_weight/day

Mahalanobis Distance: 16.423

Mahalanobis Distance p-value: 1.13e-015

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	Ochratoxin A	542	Salicylazosulfapyridine
Structure			
Actual Endpoint (-log C)	6.47264	6.59334	2.39891
Predicted Endpoint (-log C)	5.06501	5.06501	3.17598
Distance	0.639	0.639	0.712
Reference	CPDB	CPDB	CPDB

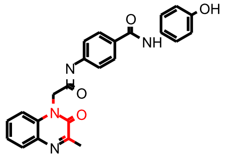
Model Applicability

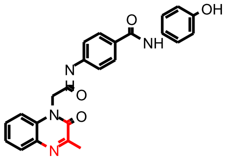
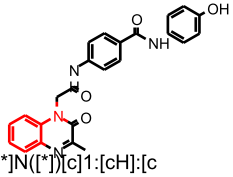
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. OPS PC20 out of range. Value: 6.2813. Training min, max, SD, explained variance: -3.9266, 5.5565, 1.236, 0.0147.

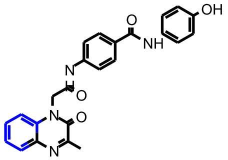
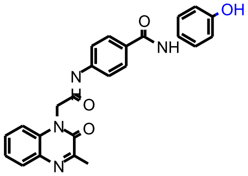
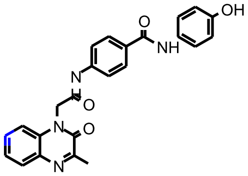
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	565998553	 [*]N([*])C(=O)C(=[*]) [*]	0.357

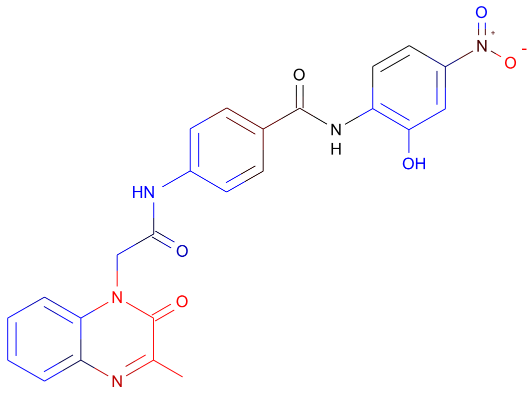
FCFP_6	565968762	 [*]\N=C(\C)/C(=[*])[*]	0.266
FCFP_6	-2090462286	 [*]N([*])[c]1:[cH]:[cH]:[cH]:[cH]:[c]:1[*]	0.245

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.422
FCFP_6	7	 [*]O	-0.372
FCFP_6	16	 [*][c](:[*]):[*]	-0.354

17d

TOPKAT_Carcinogenic_Potency_TD50_Rat



$C_{24}H_{19}N_5O_6$

Molecular Weight: 473.43756

ALogP: 2.113

Rotatable Bonds: 6

Acceptors: 7

Donors: 3

Model Prediction

Prediction: 18.156

Unit: mg/kg_body_weight/day

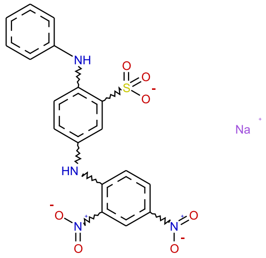
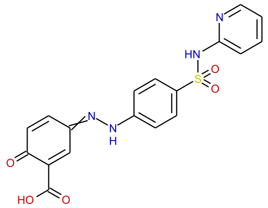
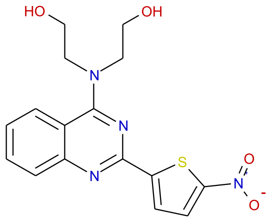
Mahalanobis Distance: 19.562

Mahalanobis Distance p-value: 2.68e-027

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	623	Salicylazosulfapyridine	4-Bis(2-hydroxyethyl)amino-2-(5-nitro-2-thienyl)quinazoline
Structure			
Actual Endpoint (-log C)	2.39985	2.39891	5.05984
Predicted Endpoint (-log C)	3.4177	3.17598	4.23808
Distance	0.641	0.678	0.733
Reference	CPDB	CPDB	CPDB

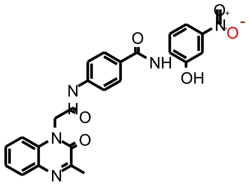
Model Applicability

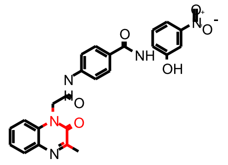
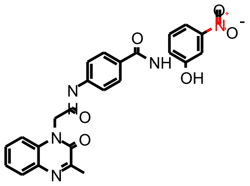
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. OPS PC20 out of range. Value: 6.0917. Training min, max, SD, explained variance: -3.9266, 5.5565, 1.236, 0.0147.

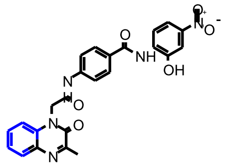
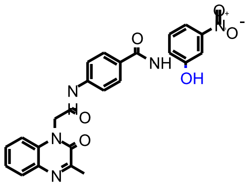
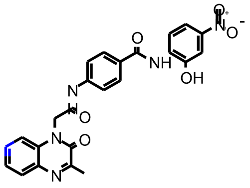
Feature Contribution

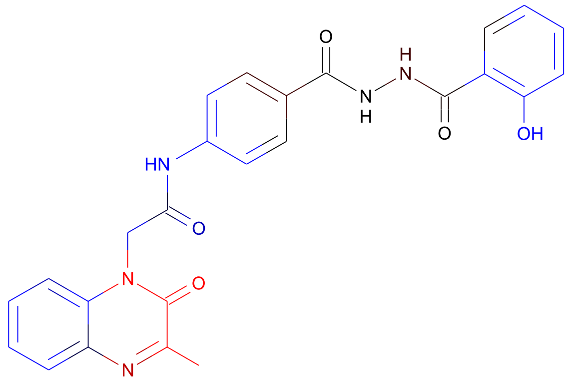
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	5	 [*][O-]	0.431

FCFP_6	565998553	 [*]N([*])C(=O)C(=[*])[*]	0.357
FCFP_6	8	 [*][N+](=[*])[*]	0.336

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.422
FCFP_6	7	 [*]O	-0.372
FCFP_6	16	 [*][c](:[*]):[*]	-0.354



$C_{25}H_{21}N_5O_5$

Molecular Weight: 471.46474

ALogP: 1.475

Rotatable Bonds: 6

Acceptors: 6

Donors: 4

Model Prediction

Prediction: 59.139

Unit: mg/kg_body_weight/day

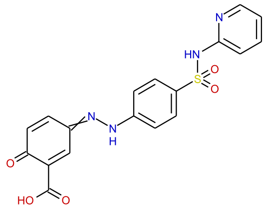
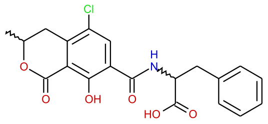
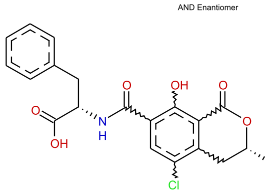
Mahalanobis Distance: 16.869

Mahalanobis Distance p-value: 3.38e-017

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	Salicylazosulfapyridine	Ochratoxin A	542
Structure			 AND Enantiomer
Actual Endpoint (-log C)	2.39891	6.47264	6.59334
Predicted Endpoint (-log C)	3.17598	5.06501	5.06501
Distance	0.706	0.737	0.737
Reference	CPDB	CPDB	CPDB

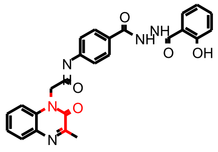
Model Applicability

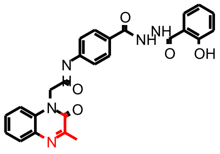
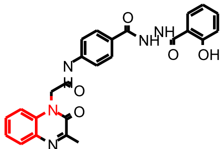
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. OPS PC20 out of range. Value: 5.8079. Training min, max, SD, explained variance: -3.9266, 5.5565, 1.236, 0.0147.

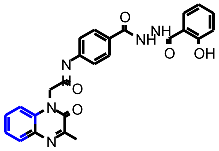
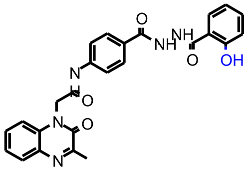
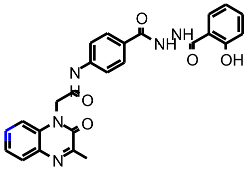
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	565998553	 [*]N([*])C(=O)C(=[*]) [*]	0.357

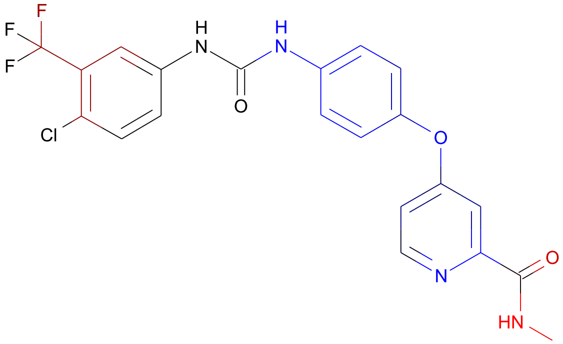
FCFP_6	565968762	 [*]\N=C(\C)/C(=[*])[*]	0.266
FCFP_6	-2090462286	 [*]N([*])[c]1:[cH]:[cH]:[cH]:[cH]:[c]:1[*]	0.245

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.422
FCFP_6	7	 [*]O	-0.372
FCFP_6	16	 [*][c](:[*]):[*]	-0.354

Sorafenib

TOPKAT_Carcinogenic_Potency_TD50_Rat



$C_{21}H_{16}ClF_3N_4O_3$

Molecular Weight: 464.82494

ALogP: 4.175

Rotatable Bonds: 6

Acceptors: 4

Donors: 3

Model Prediction

Prediction: 14.244

Unit: mg/kg_body_weight/day

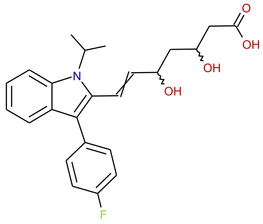
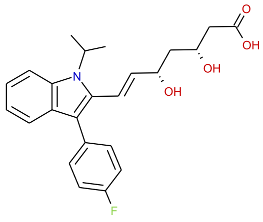
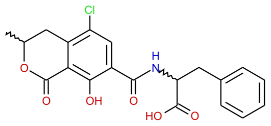
Mahalanobis Distance: 20.410

Mahalanobis Distance p-value: 9.56e-031

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	Fluvastatin	913	Ochratoxin A
Structure			
Actual Endpoint (-log C)	3.51742	3.51742	6.47264
Predicted Endpoint (-log C)	5.41573	5.41573	5.06501
Distance	0.597	0.597	0.666
Reference	CPDB	CPDB	CPDB

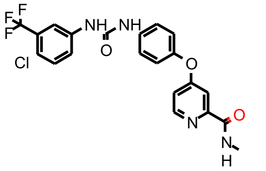
Model Applicability

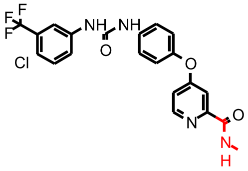
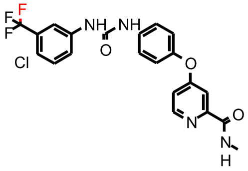
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_2 feature: -1029533685: [*]:[c](:[*])C(F)(F)F

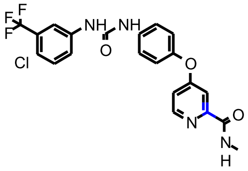
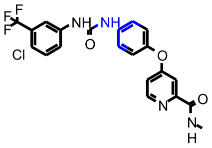
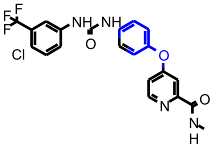
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	1	 [*]=O	0.234

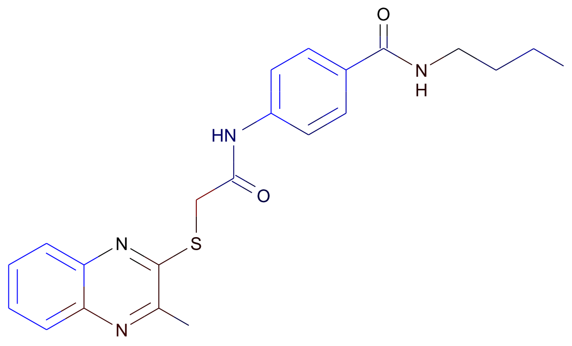
FCFP_6	-885550502	 [*]CNC(=[*])[*]	0.229
FCFP_6	32	 [*]Cl	0.154

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_6	16	 [*][c](:[*]):[*]	-0.354
FCFP_6	590925877	 [*]N[c](:[cH]:[*]):[cH]:[*]	-0.323
FCFP_6	1674451008	 [*][c]1:[*]:[cH]:[cH]:[cH]:[c]:1N=[*]	-0.233

15a

TOPKAT_Chronic_LOAEL



$C_{22}H_{24}N_4O_2S$

Molecular Weight: 408.51656

ALogP: 3.83

Rotatable Bonds: 8

Acceptors: 5

Donors: 2

Model Prediction

Prediction: 0.583

Unit: g/kg_body_weight

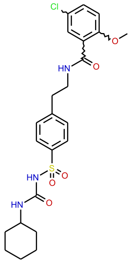
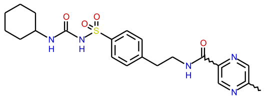
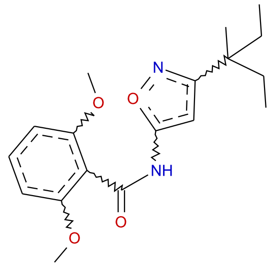
Mahalanobis Distance: 33.972

Mahalanobis Distance p-value: 8.09e-032

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	GLYBURIDE	GLIPIZIDE	ISOXABEN
Structure			
Actual Endpoint (-log C)	4.21661	3.94991	3.81665
Predicted Endpoint (-log C)	4.21035	3.95594	4.42315
Distance	0.657	0.693	0.700
Reference	UPJ-26452	NDA-17583	EPA COVER SHEET 0339;881201;(1)

Model Applicability

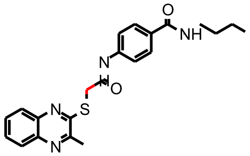
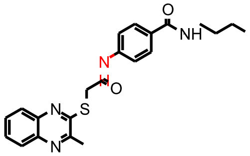
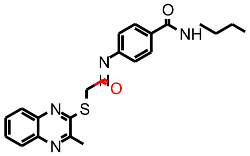
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_6 feature: -830332112: [*]S[*]
3. Unknown ECFP_6 feature: -955816473: [*]SCC(=[*])[*]
4. Unknown ECFP_6 feature: 1731843802: [*]CC(=O)N[*]
5. Unknown ECFP_6 feature: -177077903: [*]N[c](:[cH]:[*]):[cH]:[*]
6. Unknown ECFP_6 feature: -175146122: [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]
7. Unknown ECFP_6 feature: 1430169877: [*]NC(=O)[c](:[*]):[*]
8. Unknown ECFP_6 feature: 497523368: [*]CNC(=[*])[*]
9. Unknown ECFP_6 feature: -1791034651: [*]CCN[*]
10. Unknown ECFP_6 feature: 1997021792: [*]:[cH]:[cH]:[cH]:[*]
11. Unknown ECFP_6 feature: 1410041175: [*]:[cH]:[c](:n:[*]):[c](:[*]):[*]
12. Unknown ECFP_6 feature: 1652635785: [*][c](:[*]):[c](C):n:[*]
13. Unknown ECFP_6 feature: 1050567921: [*]S[c](:n:[*]):[c]([*]):[*]
14. Unknown ECFP_6 feature: 1427820655: [*]CS[c](:[*]):[*]
15. Unknown ECFP_6 feature: -1793471910: [*]CCC

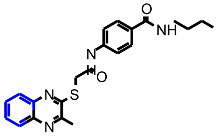
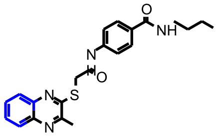
Feature Contribution

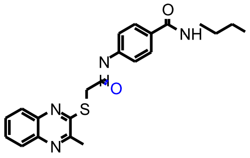
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score

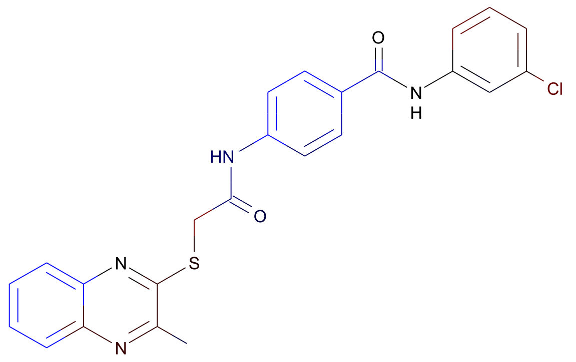
ECFP_6	1559650422	 [*]C[*]	0.129
FCFP_6	3	 [*]N[*]	0.092
ECFP_6	2099970318	 [*]C(=O)[*]	0.077

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.134
ECFP_6	1564392544	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.133

FCFP_6	1	 [*]=O	-0.102

15b



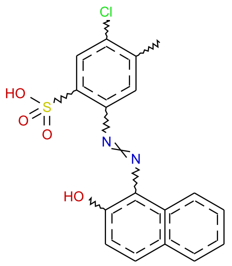
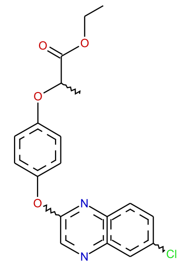
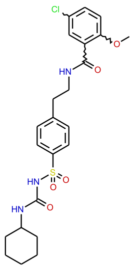
$C_{24}H_{19}ClN_4O_2S$
Molecular Weight: 462.95126
ALogP: 4.743
Rotatable Bonds: 6
Acceptors: 5
Donors: 2

Model Prediction

Prediction: 0.114
Unit: g/kg_body_weight
Mahalanobis Distance: 36.682
Mahalanobis Distance p-value: 1.47e-036

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	D & C RED 9	ASSURE	GLYBURIDE
Structure			
Actual Endpoint (-log C)	3.87715	5.00328	4.21661
Predicted Endpoint (-log C)	3.6546	4.27671	4.21035
Distance	0.707	0.770	0.776
Reference	NTP REPORT # 225	EPA COVER SHEET 0335;891001;(1)	UPJ-26452

Model Applicability

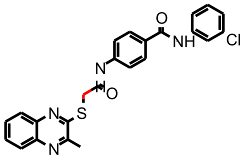
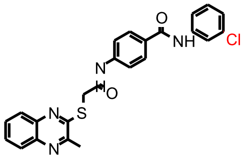
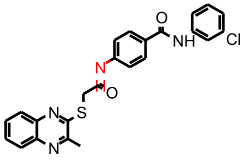
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_6 feature: -830332112: [*]S[*]
3. Unknown ECFP_6 feature: -955816473: [*]SCC(=[*])[*]
4. Unknown ECFP_6 feature: 1731843802: [*]CC(=O)N[*]
5. Unknown ECFP_6 feature: -177077903: [*]N[c](:[cH]:[*]):[cH]:[*]
6. Unknown ECFP_6 feature: -175146122: [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]
7. Unknown ECFP_6 feature: 1430169877: [*]NC(=O)[c](:[*]):[*]
8. Unknown ECFP_6 feature: 1997021792: [*]:[cH]:[cH]:[cH]:[*]
9. Unknown ECFP_6 feature: 1410041175: [*]:[cH]:[c](:n:[*]):[c](:[*]):[*]
10. Unknown ECFP_6 feature: 1652635785: [*][c](:[*]):[c](C):n:[*]
11. Unknown ECFP_6 feature: 1050567921: [*]S[c](:n:[*]):[c]([*]):[*]
12. Unknown ECFP_6 feature: 1427820655: [*]CS[c](:[*]):[*]
13. Unknown ECFP_6 feature: -176494269: [*]:[cH]:[c](Cl):[cH]:[*]
14. Unknown ECFP_6 feature: 99947387: [*]:[c](:[*])Cl

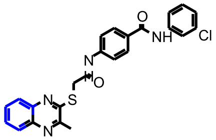
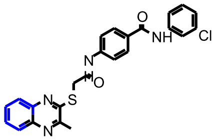
Feature Contribution

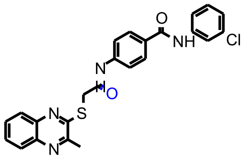
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score

ECFP_6	1559650422	 [*]C[*]	0.129
FCFP_6	32	 [*]Cl	0.101
FCFP_6	3	 [*]N[*]	0.092

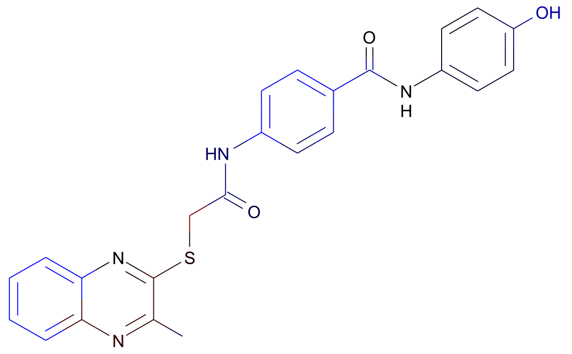
Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.134
ECFP_6	1564392544	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.133

FCFP_6	1	 [*]=O	-0.102

15c

TOPKAT_Chronic_LOAEL



$C_{24}H_{20}N_4O_3S$

Molecular Weight: 444.5056

ALogP: 3.836

Rotatable Bonds: 6

Acceptors: 6

Donors: 3

Model Prediction

Prediction: 0.181

Unit: g/kg_body_weight

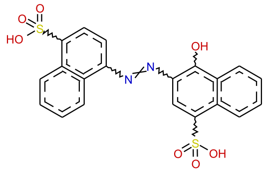
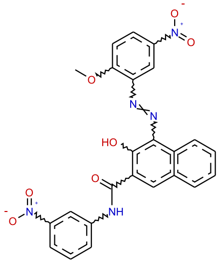
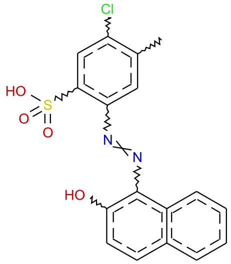
Mahalanobis Distance: 36.988

Mahalanobis Distance p-value: 4.39e-037

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	C.I. ACID RED 14	C.I.PIGMENT RED 23	D & C RED 9
Structure			
Actual Endpoint (-log C)	2.8654	2.28997	3.87715
Predicted Endpoint (-log C)	3.29295	3.52921	3.6546
Distance	0.724	0.739	0.742
Reference	NTP REPORT # 220	NTP 411 146	NTP REPORT # 225

Model Applicability

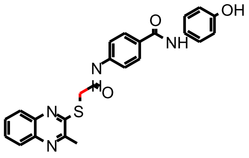
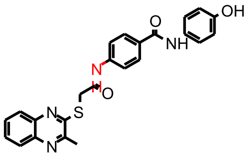
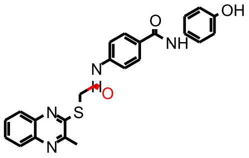
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_6 feature: -830332112: [*]S[*]
3. Unknown ECFP_6 feature: -955816473: [*]SCC(=[*])[*]
4. Unknown ECFP_6 feature: 1731843802: [*]CC(=O)N[*]
5. Unknown ECFP_6 feature: -177077903: [*]N[c](:[cH]:[*]):[cH]:[*]
6. Unknown ECFP_6 feature: -175146122: [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]
7. Unknown ECFP_6 feature: 1430169877: [*]NC(=O)[c](:[*]):[*]
8. Unknown ECFP_6 feature: 1997021792: [*]:[cH]:[cH]:[cH]:[*]
9. Unknown ECFP_6 feature: 1410041175: [*]:[cH]:[c](:n:[*]):[c](:[*]):[*]
10. Unknown ECFP_6 feature: 1652635785: [*][c](:[*]):[c](C):n:[*]
11. Unknown ECFP_6 feature: 1050567921: [*]S[c](:n:[*]):[c]([*]):[*]
12. Unknown ECFP_6 feature: 1427820655: [*]CS[c](:[*]):[*]
13. Unknown ECFP_6 feature: -177786161: [*]:[cH]:[c](O):[cH]:[*]
14. Unknown ECFP_6 feature: 2019062761: [*]:[c](:[*])O

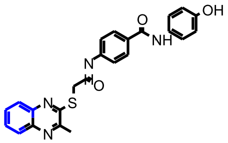
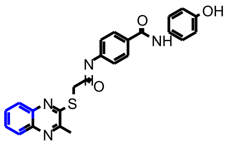
Feature Contribution

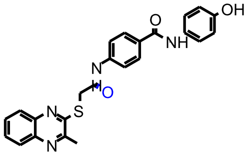
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score

ECFP_6	1559650422	 [*]C[*]	0.129
FCFP_6	3	 [*]N[*]	0.092
ECFP_6	2099970318	 [*]C(=O)[*]	0.077

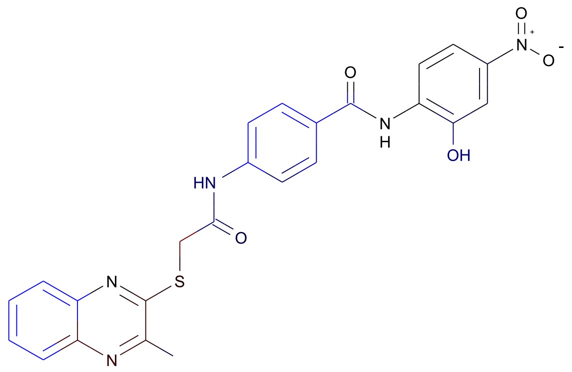
Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.134
ECFP_6	1564392544	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.133

FCFP_6	1	 [*]=O	-0.102

15d

TOPKAT_Chronic_LOAEL



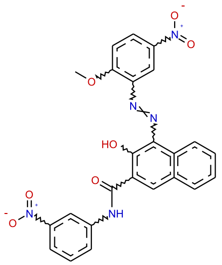
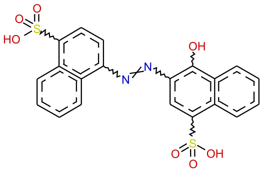
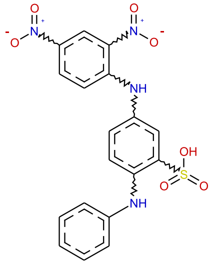
$C_{24}H_{19}N_5O_5S$
Molecular Weight: 489.50316
ALogP: 3.731
Rotatable Bonds: 7
Acceptors: 8
Donors: 3

Model Prediction

Prediction: 0.091
Unit: g/kg_body_weight
Mahalanobis Distance: 42.015
Mahalanobis Distance p-value: 1.87e-045

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	C.I.PIGMENT RED 23	C.I. ACID RED 14	C.I. ACID ORANGE 3
Structure			
Actual Endpoint (-log C)	2.28997	2.8654	3.20573
Predicted Endpoint (-log C)	3.52921	3.29295	3.55956
Distance	0.551	0.589	0.618
Reference	NTP 411 146	NTP REPORT # 220	NTP REPORT # 335

Model Applicability

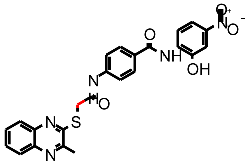
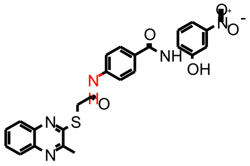
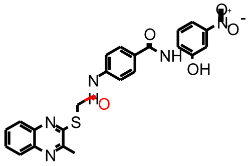
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_2 feature: 5: [*][O-]
3. Unknown FCFP_2 feature: -828984032: [*][N+](=[*])[c](:c:[*]):c:[*]
4. Unknown FCFP_2 feature: -1338588315: [*]:[c](:[*])[N+](=O)[O-]
5. Unknown FCFP_2 feature: 1872392852: [*][N+](=O)[*]
6. Unknown FCFP_2 feature: 260476081: [*][N+](=[*])[O-]
7. Unknown ECFP_6 feature: -830332112: [*]S[*]
8. Unknown ECFP_6 feature: 1043790491: [*][N+](=[*])[*]
9. Unknown ECFP_6 feature: 781519895: [*][O-]
10. Unknown ECFP_6 feature: -955816473: [*]SCC(=[*])[*]
11. Unknown ECFP_6 feature: 1731843802: [*]CC(=O)N[*]
12. Unknown ECFP_6 feature: -177077903: [*]N[c](:[cH]:[*]):[cH]:[*]
13. Unknown ECFP_6 feature: -175146122: [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]
14. Unknown ECFP_6 feature: 1430169877: [*]NC(=O)[c](:[*]):[*]
15. Unknown ECFP_6 feature: 1335108269: [*]N[c](:[cH]:[*]):[c]([*]):[*]
16. Unknown ECFP_6 feature: 1997021792: [*]:[cH]:[cH]:[cH]:[*]
17. Unknown ECFP_6 feature: 1410041175: [*]:[cH]:[c](:n:[*]):[c](:[*]):[*]
18. Unknown ECFP_6 feature: 1652635785: [*][c](:[*]):[c](C):n:[*]
19. Unknown ECFP_6 feature: 1050567921: [*]S[c](:n:[*]):[c]([*]):[*]
20. Unknown ECFP_6 feature: 1427820655: [*]CS[c](:[*]):[*]
21. Unknown ECFP_6 feature: -179073144: [*][N+](=[*])[c](:[cH]:[*]):[cH]:[*]

22. Unknown ECFP_6 feature: 2019062761: [*]:[c](:[*])O
23. Unknown ECFP_6 feature: -215026467: [*]:[c](:[*])[N+](=O)[O-]
24. Unknown ECFP_6 feature: 2104376220: [*][N+](=O)[*]
25. Unknown ECFP_6 feature: -659271057: [*][N+](=[*])[O-]

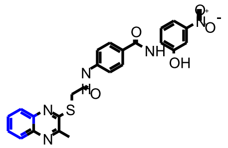
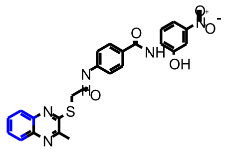
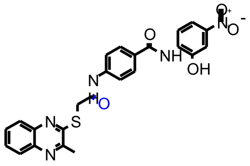
Feature Contribution

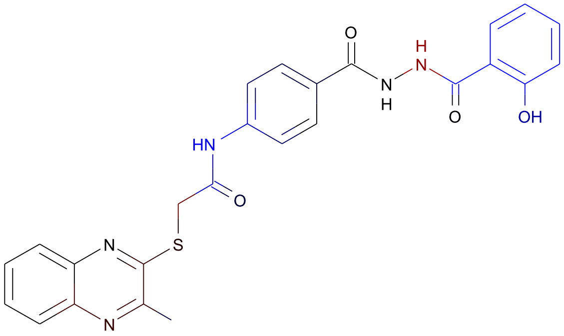
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	1559650422	 [*]C[*]	0.129
FCFP_6	3	 [*]N[*]	0.092
ECFP_6	2099970318	 [*]C(=O)[*]	0.077

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score

FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.134
ECFP_6	1564392544	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.133
FCFP_6	1	 [*]=O	-0.102



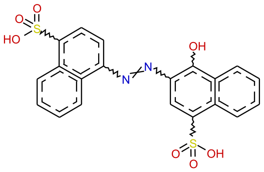
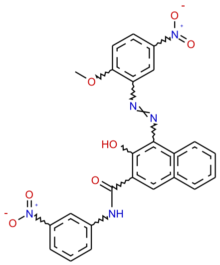
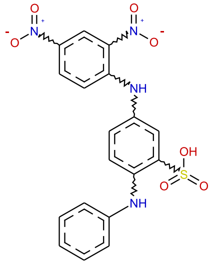
$C_{25}H_{21}N_5O_4S$
Molecular Weight: 487.53034
ALogP: 3.093
Rotatable Bonds: 7
Acceptors: 7
Donors: 4

Model Prediction

Prediction: 0.382
Unit: g/kg_body_weight
Mahalanobis Distance: 39.016
Mahalanobis Distance p-value: 1.58e-040

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	C.I. ACID RED 14	C.I.PIGMENT RED 23	C.I. ACID ORANGE 3
Structure			
Actual Endpoint (-log C)	2.8654	2.28997	3.20573
Predicted Endpoint (-log C)	3.29295	3.52921	3.55956
Distance	0.681	0.716	0.727
Reference	NTP REPORT # 220	NTP 411 146	NTP REPORT # 335

Model Applicability

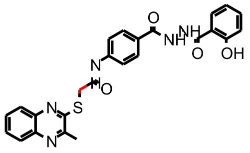
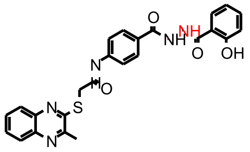
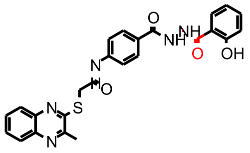
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_6 feature: -830332112: [*]S[*]
3. Unknown ECFP_6 feature: 1635339976: [*]NNC(=[*])[*]
4. Unknown ECFP_6 feature: 1430169877: [*]NC(=O)[c](:[*]):[*]
5. Unknown ECFP_6 feature: 1997021792: [*]:[cH]:[cH]:[cH]:[*]
6. Unknown ECFP_6 feature: 2019062761: [*]:[c](:[*])O
7. Unknown ECFP_6 feature: -955816473: [*]SCC(=[*])[*]
8. Unknown ECFP_6 feature: 1731843802: [*]CC(=O)N[*]
9. Unknown ECFP_6 feature: -177077903: [*]N[c](:[cH]:[*]):[cH]:[*]
10. Unknown ECFP_6 feature: -175146122: [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]
11. Unknown ECFP_6 feature: 1410041175: [*]:[cH]:[c](:n:[*]):[c](:[*]):[*]
12. Unknown ECFP_6 feature: 1652635785: [*][c](:[*]):[c](C):n:[*]
13. Unknown ECFP_6 feature: 1050567921: [*]S[c](:n:[*]):[c]([*]):[*]
14. Unknown ECFP_6 feature: 1427820655: [*]CS[c](:[*]):[*]

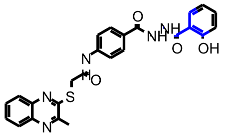
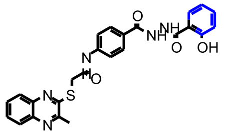
Feature Contribution

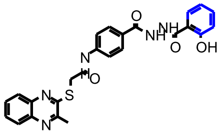
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score

ECFP_6	1559650422	 [*]C[*]	0.129
FCFP_6	3	 [*]N[*]	0.092
ECFP_6	2099970318	 [*]C(=O)[*]	0.077

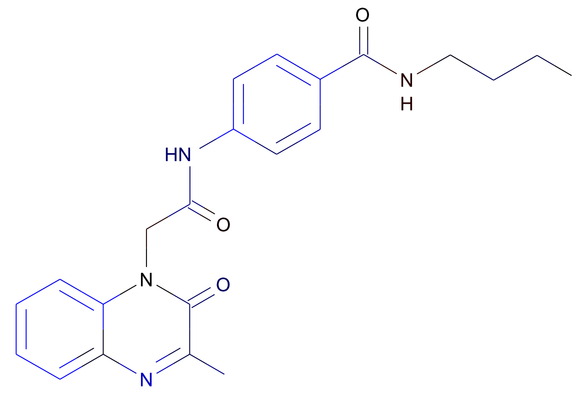
Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
ECFP_6	1337040050	 [*]C(=[*])[c](:[cH]:[*]):[c]([*]):[*]	-0.158
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.134

ECFP_6	1564392544	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.133

17a

TOPKAT_Chronic_LOAEL



$C_{22}H_{24}N_4O_3$

Molecular Weight: 392.45096

ALogP: 2.212

Rotatable Bonds: 7

Acceptors: 4

Donors: 2

Model Prediction

Prediction: 0.334

Unit: g/kg_body_weight

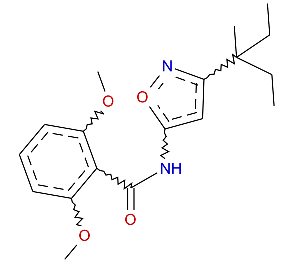
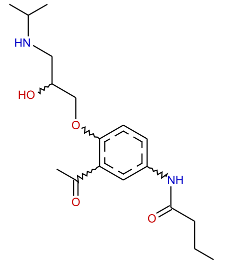
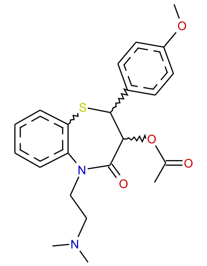
Mahalanobis Distance: 31.967

Mahalanobis Distance p-value: 3.12e-028

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	**ISOXABEN**	**ACEBUTOLOL**	**DILTIAZEM**
Structure			
Actual Endpoint (-log C)	3.81665	3.04978	4.21961
Predicted Endpoint (-log C)	4.42315	3.08045	4.005
Distance	0.623	0.641	0.654
Reference	EPA COVER SHEET 0339;881201;(1)	NDA-18917	NDA-18602

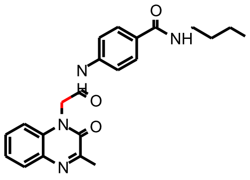
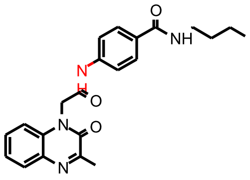
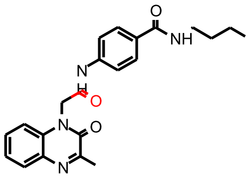
Model Applicability

Unknown features are fingerprint features in the query molecule, but not found in the training set.

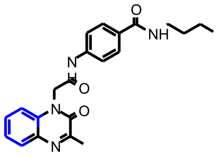
1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_6 feature: 1997021792: [*]:[cH]:[cH]:[cH]:[*]
3. Unknown ECFP_6 feature: -597295171: [*][c](:[*]):[c](:[cH]:[*])N=[*]
4. Unknown ECFP_6 feature: -1236953626: [*]N([*])[c](:[cH]:[*]):[c]([*]):[*]
5. Unknown ECFP_6 feature: 2085698692: [*]C(=N[c](:[*]):[*])[*]
6. Unknown ECFP_6 feature: 1483289300: [*]\N=C(\C)/C(=[*])[*]
7. Unknown ECFP_6 feature: 1945129186: [*]N([*])C(=O)C(=[*])[*]
8. Unknown ECFP_6 feature: -661097313: [*]CN(C(=[*])[*])[c](:[*]):[*]
9. Unknown ECFP_6 feature: -37698365: [*]N([*])CC(=[*])[*]
10. Unknown ECFP_6 feature: 1731843802: [*]CC(=O)N[*]
11. Unknown ECFP_6 feature: -177077903: [*]N[c](:[cH]:[*]):[cH]:[*]
12. Unknown ECFP_6 feature: -175146122: [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]
13. Unknown ECFP_6 feature: 1430169877: [*]NC(=O)[c](:[*]):[*]
14. Unknown ECFP_6 feature: 497523368: [*]CNC(=[*])[*]
15. Unknown ECFP_6 feature: -1791034651: [*]CCN[*]
16. Unknown ECFP_6 feature: -1793471910: [*]CCC

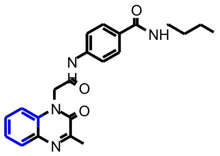
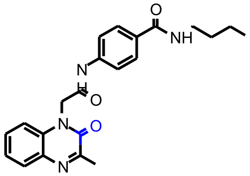
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	1559650422	 [*]C[*]	0.129
FCFP_6	3	 [*]N[*]	0.092
ECFP_6	2099970318	 [*]C(=O)[*]	0.077

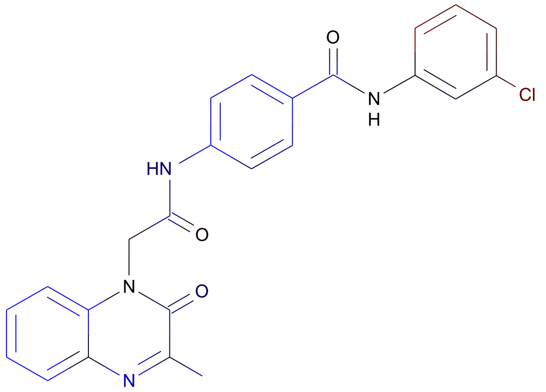
Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.134

ECFP_6	1564392544	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.133
ECFP_6	2106656448	 [*]C(=O)[*]	-0.110

17b

TOPKAT_Chronic_LOAEL



$C_{24}H_{19}ClN_4O_3$

Molecular Weight: 446.88566

ALogP: 3.125

Rotatable Bonds: 5

Acceptors: 4

Donors: 2

Model Prediction

Prediction: 0.072

Unit: g/kg_body_weight

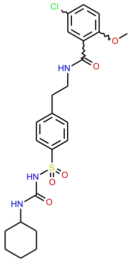
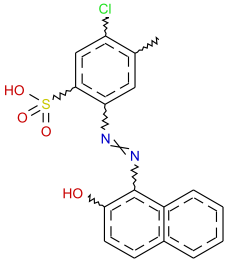
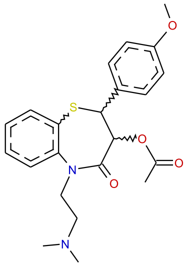
Mahalanobis Distance: 31.457

Mahalanobis Distance p-value: 2.59e-027

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	GLYBURIDE	D & C RED 9	DILTIAZEM
Structure			
Actual Endpoint (-log C)	4.21661	3.87715	4.21961
Predicted Endpoint (-log C)	4.21035	3.6546	4.005
Distance	0.720	0.722	0.723
Reference	UPJ-26452	NTP REPORT # 225	NDA-18602

Model Applicability

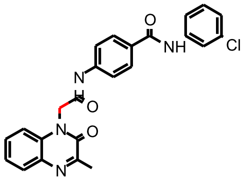
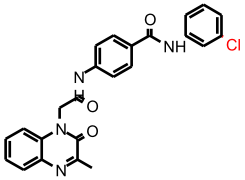
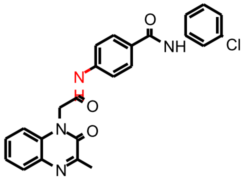
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_6 feature: 1997021792: [*]:[cH]:[cH]:[cH]:[*]
3. Unknown ECFP_6 feature: -597295171: [*][c](:[*]):[c](:[cH]:[*])N=[*]
4. Unknown ECFP_6 feature: -1236953626: [*]N([*])[c](:[cH]:[*]):[c]([*]):[*]
5. Unknown ECFP_6 feature: 2085698692: [*]C(=N[c](:[*]):[*])[*]
6. Unknown ECFP_6 feature: 1483289300: [*]\N=C(\C)/C(=[*])[*]
7. Unknown ECFP_6 feature: 1945129186: [*]N([*])C(=O)C(=[*])[*]
8. Unknown ECFP_6 feature: -661097313: [*]CN(C(=[*])[*])[c](:[*]):[*]
9. Unknown ECFP_6 feature: -37698365: [*]N([*])CC(=[*])[*]
10. Unknown ECFP_6 feature: 1731843802: [*]CC(=O)N[*]
11. Unknown ECFP_6 feature: -177077903: [*]N[c](:[cH]:[*]):[cH]:[*]
12. Unknown ECFP_6 feature: -175146122: [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]
13. Unknown ECFP_6 feature: 1430169877: [*]NC(=O)[c](:[*]):[*]
14. Unknown ECFP_6 feature: -176494269: [*]:[cH]:[c](Cl):[cH]:[*]
15. Unknown ECFP_6 feature: 99947387: [*]:[c](:[*])Cl

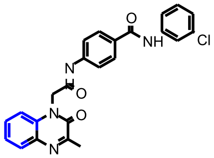
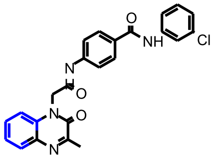
Feature Contribution

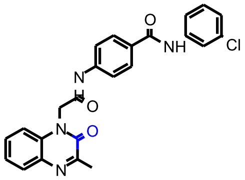
Top features for positive contribution

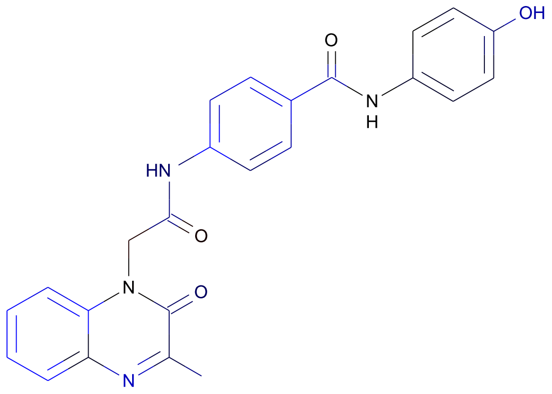
Fingerprint	Bit/Smiles	Feature Structure	Score

ECFP_6	1559650422	 [*]C[*]	0.129
FCFP_6	32	 [*]Cl	0.101
FCFP_6	3	 [*]N[*]	0.092

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.134
ECFP_6	1564392544	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.133

ECFP_6	2106656448	 [*]C(=O)[*]	-0.110



$C_{24}H_{20}N_4O_4$

Molecular Weight: 428.44

ALogP: 2.218

Rotatable Bonds: 5

Acceptors: 5

Donors: 3

Model Prediction

Prediction: 0.132

Unit: g/kg_body_weight

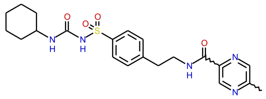
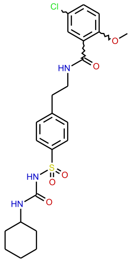
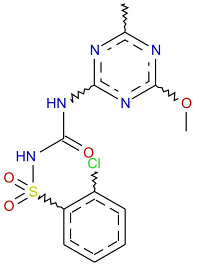
Mahalanobis Distance: 31.083

Mahalanobis Distance p-value: 1.24e-026

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	GLIPIZIDE	GLYBURIDE	CHLORSULFURON
Structure			
Actual Endpoint (-log C)	3.94991	4.21661	4.15566
Predicted Endpoint (-log C)	3.95594	4.21035	3.79771
Distance	0.653	0.701	0.742
Reference	NDA-17583	UPJ-26452	EPA COVER SHEET 0027;880301;(1)

Model Applicability

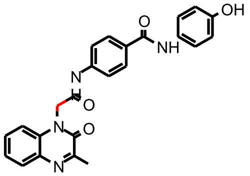
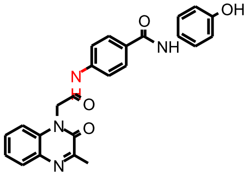
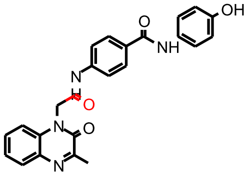
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_6 feature: 1997021792: [*]:[cH]:[cH]:[cH]:[*]
3. Unknown ECFP_6 feature: -597295171: [*][c](:[*]):[c](:[cH]:[*])N=[*]
4. Unknown ECFP_6 feature: -1236953626: [*]N([*])[c](:[cH]:[*]):[c]([*]):[*]
5. Unknown ECFP_6 feature: 2085698692: [*]C(=N[c](:[*]):[*])[*]
6. Unknown ECFP_6 feature: 1483289300: [*]\N=C(\C)/C(=[*])[*]
7. Unknown ECFP_6 feature: 1945129186: [*]N([*])C(=O)C(=[*])[*]
8. Unknown ECFP_6 feature: -661097313: [*]CN(C(=[*])[*])[c](:[*]):[*]
9. Unknown ECFP_6 feature: -37698365: [*]N([*])CC(=[*])[*]
10. Unknown ECFP_6 feature: 1731843802: [*]CC(=O)N[*]
11. Unknown ECFP_6 feature: -177077903: [*]N[c](:[cH]:[*]):[cH]:[*]
12. Unknown ECFP_6 feature: -175146122: [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]
13. Unknown ECFP_6 feature: 1430169877: [*]NC(=O)[c](:[*]):[*]
14. Unknown ECFP_6 feature: -177786161: [*]:[cH]:[c](O):[cH]:[*]
15. Unknown ECFP_6 feature: 2019062761: [*]:[c](:[*])O

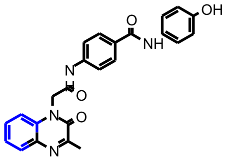
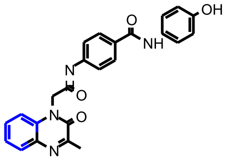
Feature Contribution

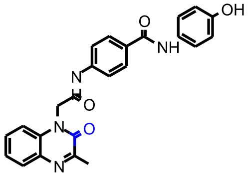
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score

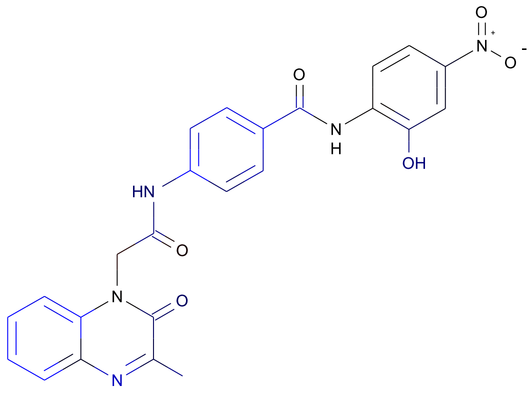
ECFP_6	1559650422	 [*]C[*]	0.129
FCFP_6	3	 [*]N[*]	0.092
ECFP_6	2099970318	 [*]C(=O)[*]	0.077

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.134
ECFP_6	1564392544	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.133

ECFP_6	2106656448	 [*]C(=O)[*]	-0.110

17d



$C_{24}H_{19}N_5O_6$

Molecular Weight: 473.43756

ALogP: 2.113

Rotatable Bonds: 6

Acceptors: 7

Donors: 3

Model Prediction

Prediction: 0.103

Unit: g/kg_body_weight

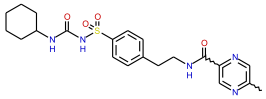
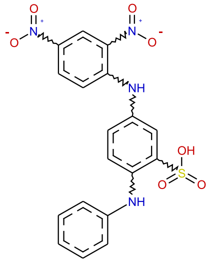
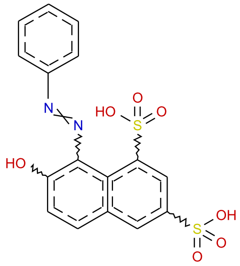
Mahalanobis Distance: 37.944

Mahalanobis Distance p-value: 1.02e-038

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	GLIPIZIDE	C.I. ACID ORANGE 3	C.I. ACID ORANGE 10
Structure			
Actual Endpoint (-log C)	3.94991	3.20573	3.435
Predicted Endpoint (-log C)	3.95594	3.55956	3.28295
Distance	0.630	0.677	0.685
Reference	NDA-17583	NTP REPORT # 335	NTP REPORT # 211

Model Applicability

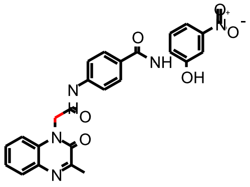
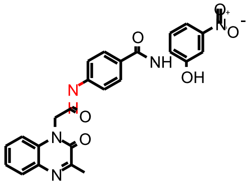
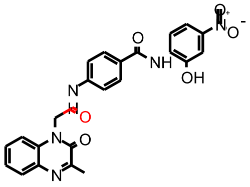
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_2 feature: 5: [*][O-]
3. Unknown FCFP_2 feature: -828984032: [*][N+](=[*])[c](:c:[*]):c:[*]
4. Unknown FCFP_2 feature: -1338588315: [*]:[c](:[*])[N+](=O)[O-]
5. Unknown FCFP_2 feature: 1872392852: [*][N+](=O)[*]
6. Unknown FCFP_2 feature: 260476081: [*][N+](=[*])[O-]
7. Unknown ECFP_6 feature: 1043790491: [*][N+](=[*])[*]
8. Unknown ECFP_6 feature: 781519895: [*][O-]
9. Unknown ECFP_6 feature: 1997021792: [*]:[cH]:[cH]:[cH]:[*]
10. Unknown ECFP_6 feature: -597295171: [*][c](:[*]):[c](:[cH]:[*])N=[*]
11. Unknown ECFP_6 feature: -1236953626: [*]N([*])[c](:[cH]:[*]):[c]([*]):[*]
12. Unknown ECFP_6 feature: 2085698692: [*]C(=N[c](:[*]):[*])[*]
13. Unknown ECFP_6 feature: 1483289300: [*]\N=C(\C)/C(=[*])[*]
14. Unknown ECFP_6 feature: 1945129186: [*]N([*])C(=O)C(=[*])[*]
15. Unknown ECFP_6 feature: -661097313: [*]CN(C(=[*])[*])[c](:[*]):[*]
16. Unknown ECFP_6 feature: -37698365: [*]N([*])CC(=[*])[*]
17. Unknown ECFP_6 feature: 1731843802: [*]CC(=O)N[*]
18. Unknown ECFP_6 feature: -177077903: [*]N[c](:[cH]:[*]):[cH]:[*]
19. Unknown ECFP_6 feature: -175146122: [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]
20. Unknown ECFP_6 feature: 1430169877: [*]NC(=O)[c](:[*]):[*]
21. Unknown ECFP_6 feature: 1335108269: [*]N[c](:[cH]:[*]):[c]([*]):[*]

22. Unknown ECFP_6 feature: -179073144: [*][N+](=[*])[c](:[cH]:[*]):[cH]:[*]
23. Unknown ECFP_6 feature: 2019062761: [*]:[c](:[*])O
24. Unknown ECFP_6 feature: -215026467: [*]:[c](:[*])[N+](=O)[O-]
25. Unknown ECFP_6 feature: 2104376220: [*][N+](=O)[*]
26. Unknown ECFP_6 feature: -659271057: [*][N+](=[*])[O-]

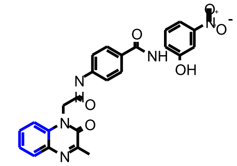
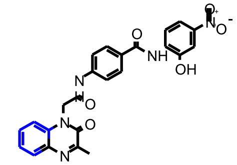
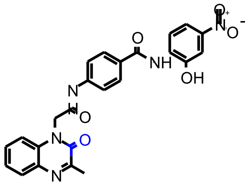
Feature Contribution

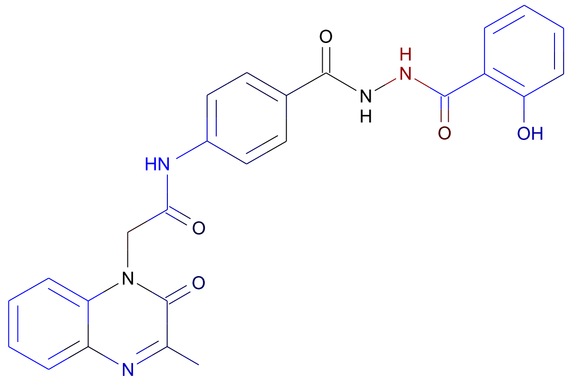
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	1559650422	 [*]C[*]	0.129
FCFP_6	3	 [*]N[*]	0.092
ECFP_6	2099970318	 [*]C(=O)[*]	0.077

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score

FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.134
ECFP_6	1564392544	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.133
ECFP_6	2106656448	 [*]C(=O)[*]	-0.110



$C_{25}H_{21}N_5O_5$

Molecular Weight: 471.46474

ALogP: 1.475

Rotatable Bonds: 6

Acceptors: 6

Donors: 4

Model Prediction

Prediction: 0.271

Unit: g/kg_body_weight

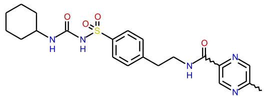
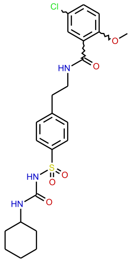
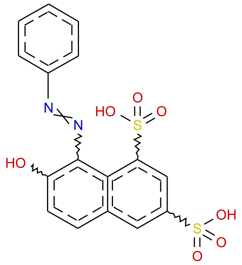
Mahalanobis Distance: 34.979

Mahalanobis Distance p-value: 1.35e-033

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	GLIPIZIDE	GLYBURIDE	C.I. ACID ORANGE 10
Structure			
Actual Endpoint (-log C)	3.94991	4.21661	3.435
Predicted Endpoint (-log C)	3.95594	4.21035	3.28295
Distance	0.631	0.736	0.788
Reference	NDA-17583	UPJ-26452	NTP REPORT # 211

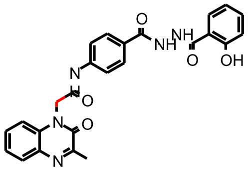
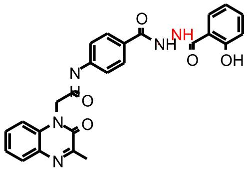
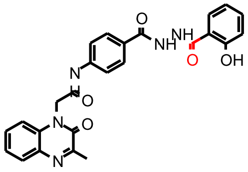
Model Applicability

Unknown features are fingerprint features in the query molecule, but not found in the training set.

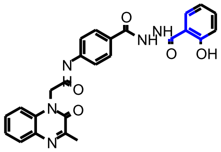
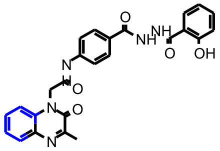
1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_6 feature: 1997021792: [*]:[cH]:[cH]:[cH]:[*]
3. Unknown ECFP_6 feature: -597295171: [*][c](:[*]):[c](:[cH]:[*])N=[*]
4. Unknown ECFP_6 feature: -1236953626: [*]N([*])[c](:[cH]:[*]):[c]([*]):[*]
5. Unknown ECFP_6 feature: 2085698692: [*]C(=N[c](:[*]):[*])[*]
6. Unknown ECFP_6 feature: 1483289300: [*]\N=C(\C)/C(=[*])[*]
7. Unknown ECFP_6 feature: 1945129186: [*]N([*])C(=O)C(=[*])[*]
8. Unknown ECFP_6 feature: -661097313: [*]CN(C(=[*])[*])[c](:[*]):[*]
9. Unknown ECFP_6 feature: 1635339976: [*]NNC(=[*])[*]
10. Unknown ECFP_6 feature: 1430169877: [*]NC(=O)[c](:[*]):[*]
11. Unknown ECFP_6 feature: 2019062761: [*]:[c](:[*])O
12. Unknown ECFP_6 feature: -37698365: [*]N([*])CC(=[*])[*]
13. Unknown ECFP_6 feature: 1731843802: [*]CC(=O)N[*]
14. Unknown ECFP_6 feature: -177077903: [*]N[c](:[cH]:[*]):[cH]:[*]
15. Unknown ECFP_6 feature: -175146122: [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]

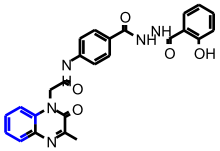
Feature Contribution

Top features for positive contribution			
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**

ECFP_6	1559650422	 [*]C[*]	0.129
FCFP_6	3	 [*]N[*]	0.092
ECFP_6	2099970318	 [*]C(=O)[*]	0.077

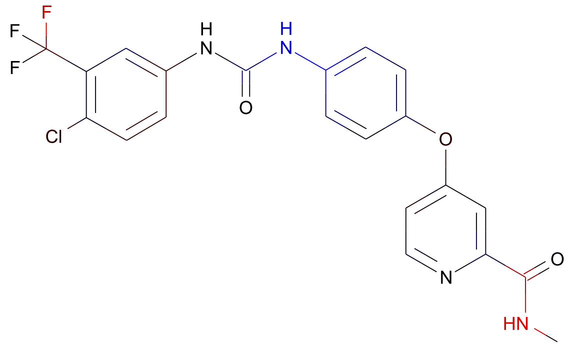
Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	1337040050	 [*]C(=[*])[c](:[cH]:[*]):[c]([*]):[*]	-0.158
FCFP_6	991735244	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.134

ECFP_6	1564392544	 [*]:[c]1:[*]:[cH]:[cH]:[cH]:[cH]:1	-0.133

Sorafenib

TOPKAT_Chronic_LOAEL



$C_{21}H_{16}ClF_3N_4O_3$

Molecular Weight: 464.82494

ALogP: 4.175

Rotatable Bonds: 6

Acceptors: 4

Donors: 3

Model Prediction

Prediction: 0.005

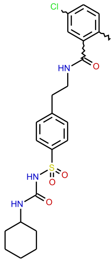
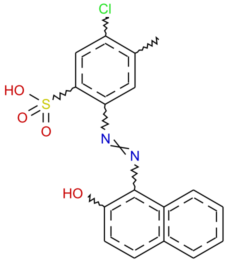
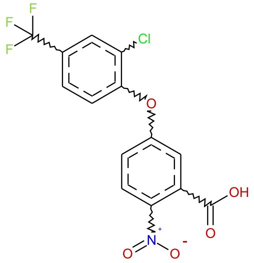
Unit: g/kg_body_weight

Mahalanobis Distance: 29.988

Mahalanobis Distance p-value: 1.21e-024

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	GLYBURIDE	D & C RED 9	SODIUM ACIFLUORFEN
Structure			
Actual Endpoint (-log C)	4.21661	3.87715	4.16036
Predicted Endpoint (-log C)	4.21035	3.6546	4.65915
Distance	0.636	0.722	0.736
Reference	UPJ-26452	NTP REPORT # 225	EPA COVER SHEET 0192;891101;(1)

Model Applicability

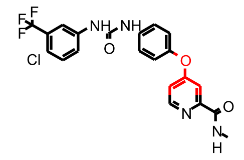
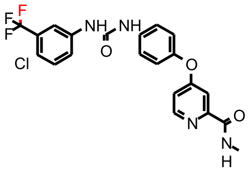
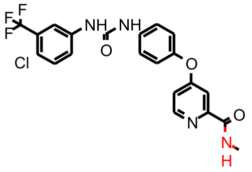
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_6 feature: -1046436026: [*]F
3. Unknown ECFP_6 feature: 1413420509: [*]C(=[*])[c](:[cH]:[*]):n:[*]
4. Unknown ECFP_6 feature: 1996163143: [*]:[cH]:[cH]:n:[*]
5. Unknown ECFP_6 feature: -677309799: [*][c](:[*]):n:[cH]:[*]
6. Unknown ECFP_6 feature: 1430169877: [*]NC(=O)[c](:[*]):[*]
7. Unknown ECFP_6 feature: 1338334141: [*]C(=[*])NC
8. Unknown ECFP_6 feature: 1305253718: [*]:[c](:[*])O[c](:[*]):[*]
9. Unknown ECFP_6 feature: -177077903: [*]N[c](:[cH]:[*]):[cH]:[*]
10. Unknown ECFP_6 feature: -649580166: [*]NC(=O)N[*]
11. Unknown ECFP_6 feature: 1336678434: [*][c](:[*]):[c](:[cH]:[*])C([*])([*])[*]
12. Unknown ECFP_6 feature: -1952889961: [*]:[c](:[*])C(F)(F)F
13. Unknown ECFP_6 feature: 226796801: [*]C([*])([*])F
14. Unknown ECFP_6 feature: 99947387: [*]:[c](:[*])Cl
15. Unknown ECFP_6 feature: 864287155: [*]NC

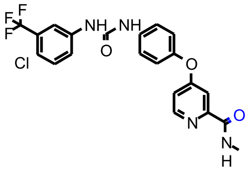
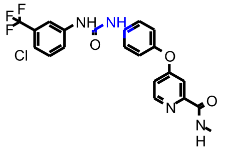
Feature Contribution

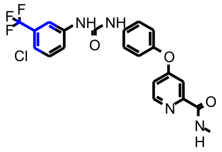
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score

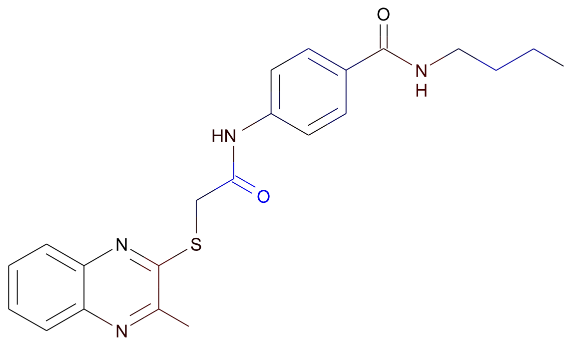
ECFP_6	-176455838	 [*]O[c](:[cH]:[*]):[cH]:[*]	0.106
FCFP_6	32	 [*]Cl	0.101
FCFP_6	3	 [*]N[*]	0.092

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_6	1	 [*]=O	-0.102
ECFP_6	-1236483485	 [*]C(=[*])N[c](:[*]):[*]	-0.075

FCFP_6	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.071

TOPKAT_Rat_Maximum_Tolerated_Dose_Feed



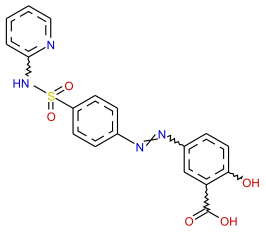
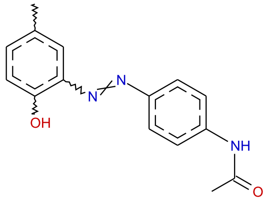
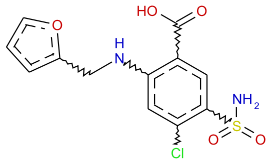
$C_{22}H_{24}N_4O_2S$
Molecular Weight: 408.51656
ALogP: 3.83
Rotatable Bonds: 8
Acceptors: 5
Donors: 2

Model Prediction

Prediction: 0.132
Unit: g/kg_body_weight
Mahalanobis Distance: 11.557
Mahalanobis Distance p-value: 8.5e-008

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	SALICYLAZOSULFAPYRIDINE	DISPERSE YELLOW 3	FUROSEMIDE
Structure			
Actual Endpoint (-log C)	3.375	2.77703	4.04236
Predicted Endpoint (-log C)	2.80292	2.80195	2.8614
Distance	0.750	0.754	0.769
Reference	NCI/NTP TR-457	NCI/NTP TR-222	NCI/NTP TR-356

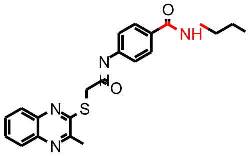
Model Applicability

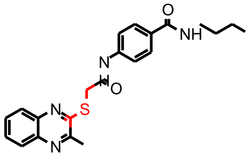
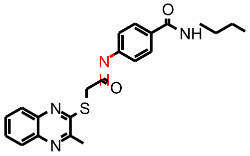
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. OPS PC8 out of range. Value: 4.0355. Training min, max, SD, explained variance: -3.8548, 3.9137, 1.331, 0.0400.
2. OPS PC9 out of range. Value: 3.781. Training min, max, SD, explained variance: -2.8548, 3.3954, 1.263, 0.0360.
3. OPS PC14 out of range. Value: 3.5834. Training min, max, SD, explained variance: -2.0656, 3.3808, 1.011, 0.0231.

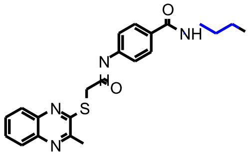
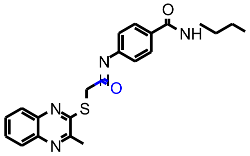
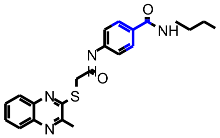
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	-885550502	 [*]CNC(=[*])[*]	0.115

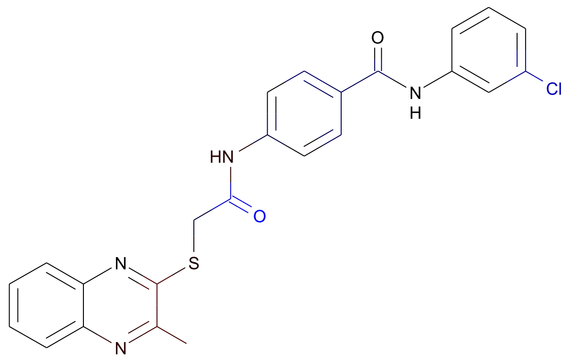
FCFP_2	1036089772	 [*]CS[c](:[*]):[*]	0.075
FCFP_2	3	 [*]N[*]	0.074

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	-1272798659	 [*]CCC[*]	-0.111
FCFP_2	1872154524	 [*]C(=O)[*]	-0.105
FCFP_2	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.083

15b

TOPKAT_Rat_Maximum_Tolerated_Dose_Feed



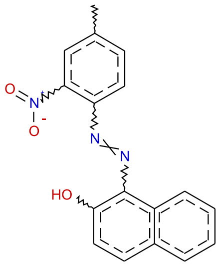
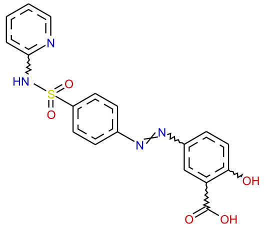
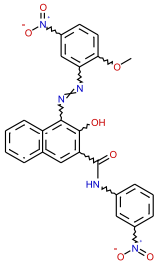
$C_{24}H_{19}ClN_4O_2S$
Molecular Weight: 462.95126
ALogP: 4.743
Rotatable Bonds: 6
Acceptors: 5
Donors: 2

Model Prediction

Prediction: 0.123
Unit: g/kg_body_weight
Mahalanobis Distance: 11.191
Mahalanobis Distance p-value: 4e-007

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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

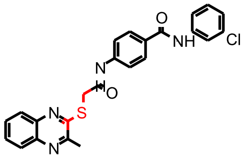
Name	C.I.PIGMENT RED 3	SALICYLAZOSULFAPYRIDINE	C.I.PIGMENT RED 23
Structure			
Actual Endpoint (-log C)	2.65635	3.375	2.30052
Predicted Endpoint (-log C)	2.97957	2.80292	3.55333
Distance	0.791	0.809	0.844
Reference	NCI/NTP TR-407	NCI/NTP TR-457	NCI/NTP TR-411

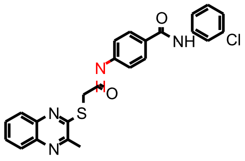
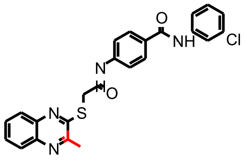
Model Applicability

Unknown features are fingerprint features in the query molecule, but not found in the training set.

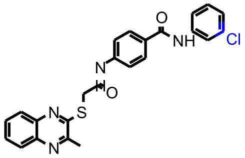
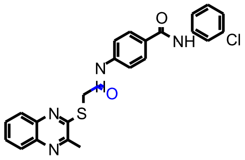
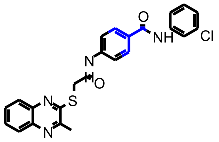
1. All properties and OPS components are within expected ranges.

Feature Contribution

Top features for positive contribution			
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	1036089772	 [*]CS[c](:[*]):[*]	0.075

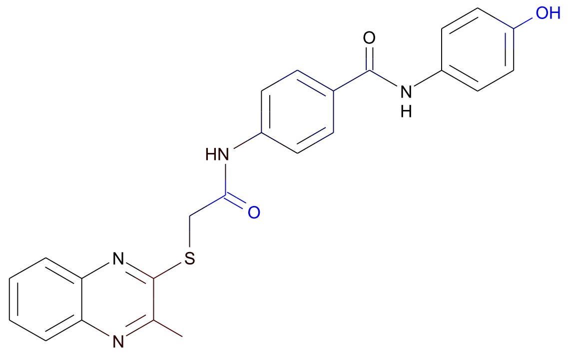
FCFP_2	3	 [*]N[*]	0.074
FCFP_2	136120670	 [*]:[c](:[*])C	0.064

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	71476542	 [*]:[c](:[*])Cl	-0.134
FCFP_2	1872154524	 [*]C(=O)[*]	-0.105
FCFP_2	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.083

15c

TOPKAT_Rat_Maximum_Tolerated_Dose_Feed



$C_{24}H_{20}N_4O_3S$

Molecular Weight: 444.5056

ALogP: 3.836

Rotatable Bonds: 6

Acceptors: 6

Donors: 3

Model Prediction

Prediction: 0.333

Unit: g/kg_body_weight

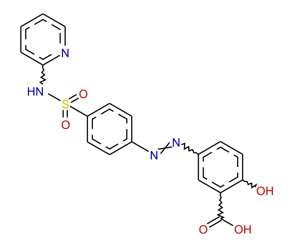
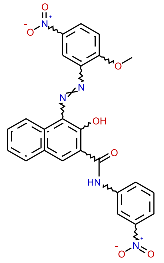
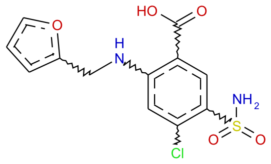
Mahalanobis Distance: 10.846

Mahalanobis Distance p-value: 1.65e-006

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	SALICYLAZOSULFAPYRIDINE	C.I.PIGMENT RED 23	FUROSEMIDE
Structure			
Actual Endpoint (-log C)	3.375	2.30052	4.04236
Predicted Endpoint (-log C)	2.80292	3.55333	2.8614
Distance	0.582	0.670	0.818
Reference	NCI/NTP TR-457	NCI/NTP TR-411	NCI/NTP TR-356

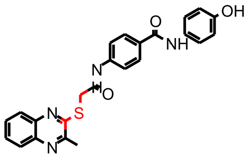
Model Applicability

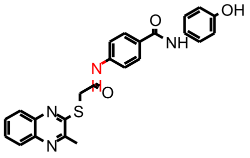
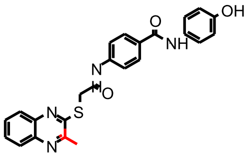
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

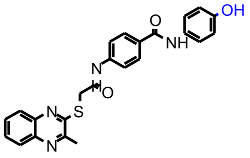
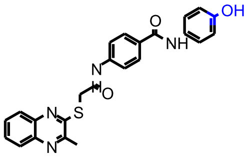
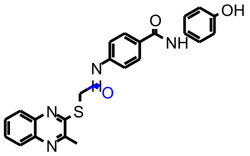
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	1036089772	 [*]CS[c](:[*]):[*]	0.075

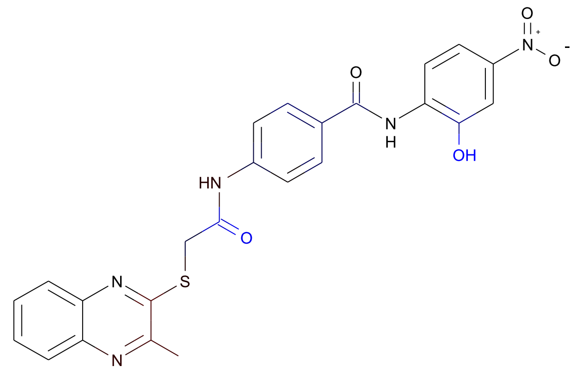
FCFP_2	3	 [*]N[*]	0.074
FCFP_2	136120670	 [*]:[c](:[*])C	0.064

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	7	 [*]O	-0.214
FCFP_2	-549108873	 [*]:[c](:[*])O	-0.127
FCFP_2	1872154524	 [*]C(=O)[*]	-0.105

15d

TOPKAT_Rat_Maximum_Tolerated_Dose_Feed



$C_{24}H_{19}N_5O_5S$

Molecular Weight: 489.50316

ALogP: 3.731

Rotatable Bonds: 7

Acceptors: 8

Donors: 3

Model Prediction

Prediction: 0.244

Unit: g/kg_body_weight

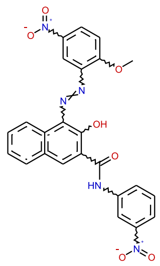
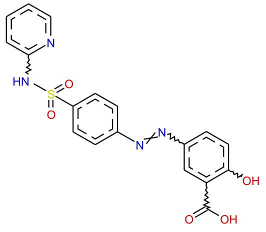
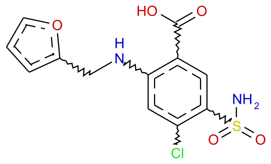
Mahalanobis Distance: 11.389

Mahalanobis Distance p-value: 1.74e-007

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	**C.I.PIGMENT RED 23**	**SALICYLAZOSULFAPYRIDINE**	**FUROSEMIDE**
Structure			
Actual Endpoint (-log C)	2.30052	3.375	4.04236
Predicted Endpoint (-log C)	3.55333	2.80292	2.8614
Distance	0.513	0.620	0.962
Reference	NCI/NTP TR-411	NCI/NTP TR-457	NCI/NTP TR-356

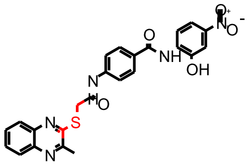
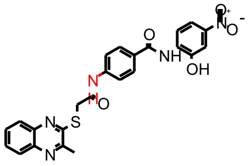
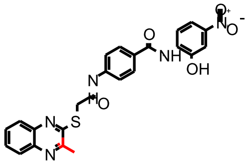
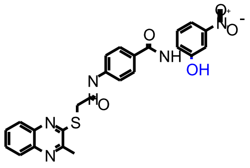
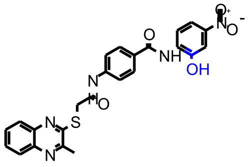
Model Applicability

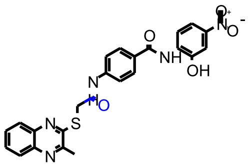
Unknown features are fingerprint features in the query molecule, but not found in the training set.

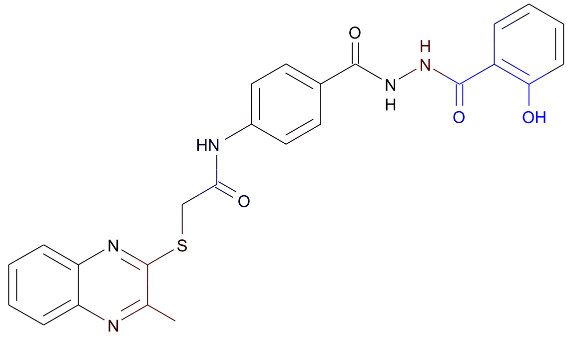
1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_2 feature: 8: [*][N+](=[*])[*]
3. Unknown FCFP_2 feature: 5: [*][O-]
4. Unknown FCFP_2 feature: -828984032: [*][N+](=[*])[c](:c:[*]):c:[*]
5. Unknown FCFP_2 feature: -1338588315: [*]:[c](:[*])[N+](=O)[O-]
6. Unknown FCFP_2 feature: 1872392852: [*][N+](=O)[*]
7. Unknown FCFP_2 feature: 260476081: [*][N+](=[*])[O-]

Feature Contribution

Top features for positive contribution			
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**

FCFP_2	1036089772	 [*]CS[c](:[*]):[*]	0.075
FCFP_2	3	 [*]N[*]	0.074
FCFP_2	136120670	 [*]:[c](:[*])C	0.064
Top Features for negative contribution			
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	7	 [*]O	-0.214
FCFP_2	-549108873	 [*]:[c](:[*])O	-0.127

FCFP_2	1872154524	 [*]C(=O)[*]	-0.105



$C_{25}H_{21}N_5O_4S$

Molecular Weight: 487.53034

ALogP: 3.093

Rotatable Bonds: 7

Acceptors: 7

Donors: 4

Model Prediction

Prediction: 0.364

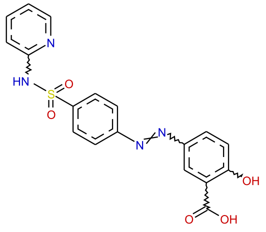
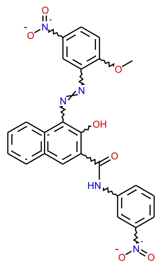
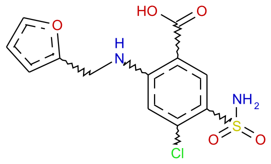
Unit: g/kg_body_weight

Mahalanobis Distance: 11.452

Mahalanobis Distance p-value: 1.33e-007

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	SALICYLAZOSULFAPYRIDINE	C.I.PIGMENT RED 23	FUROSEMIDE
Structure			
Actual Endpoint (-log C)	3.375	2.30052	4.04236
Predicted Endpoint (-log C)	2.80292	3.55333	2.8614
Distance	0.599	0.657	0.880
Reference	NCI/NTP TR-457	NCI/NTP TR-411	NCI/NTP TR-356

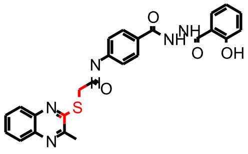
Model Applicability

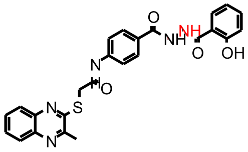
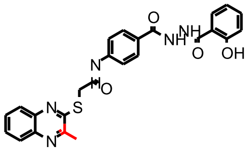
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_2 feature: -885461129: [*]NNC(=[*])[*]

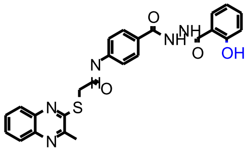
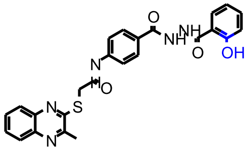
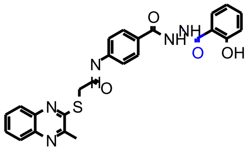
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	1036089772	 [*]CS[c](:[*]):[*]	0.075

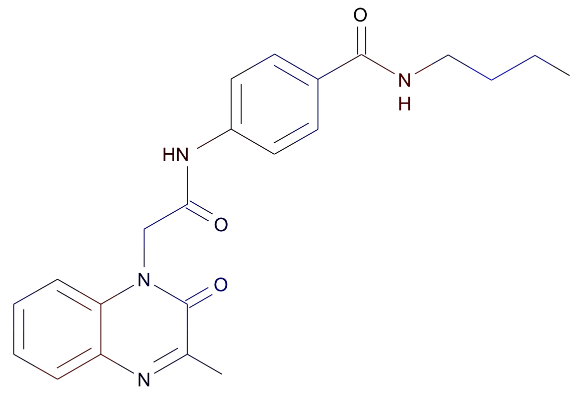
FCFP_2	3	 [*]N[*]	0.074
FCFP_2	136120670	 [*]:[c](:[*])C	0.064

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	7	 [*]O	-0.214
FCFP_2	-549108873	 [*]:[c](:[*])O	-0.127
FCFP_2	1872154524	 [*]C(=O)[*]	-0.105

17a

TOPKAT_Rat_Maximum_Tolerated_Dose_Feed



$C_{22}H_{24}N_4O_3$

Molecular Weight: 392.45096

ALogP: 2.212

Rotatable Bonds: 7

Acceptors: 4

Donors: 2

Model Prediction

Prediction: 0.097

Unit: g/kg_body_weight

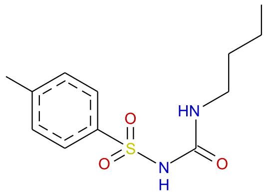
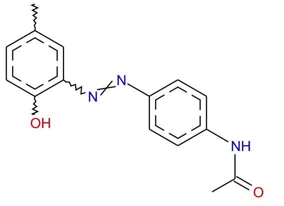
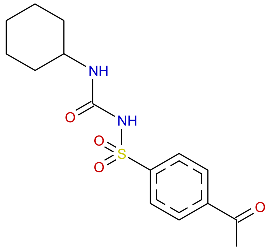
Mahalanobis Distance: 11.056

Mahalanobis Distance p-value: 7e-007

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	**TOLBUTAMIDE**	**DISPERSE YELLOW 3**	**ACETOHEXAMIDE**
Structure			
Actual Endpoint (-log C)	2.3985	2.77703	2.55683
Predicted Endpoint (-log C)	3.32272	2.80195	3.62413
Distance	0.596	0.607	0.609
Reference	NCI/NTP TR-031	NCI/NTP TR-222	NCI/NTP TR-050

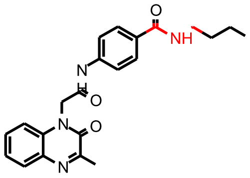
Model Applicability

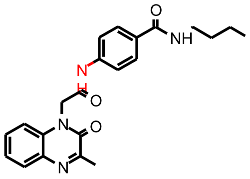
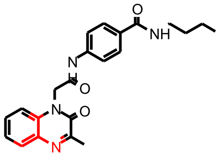
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. OPS PC9 out of range. Value: 3.6764. Training min, max, SD, explained variance: -2.8548, 3.3954, 1.263, 0.0360.

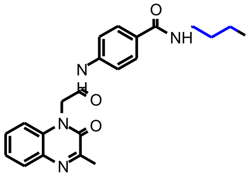
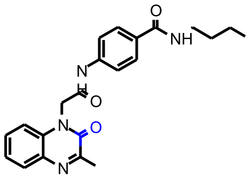
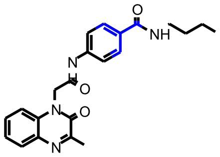
Feature Contribution

Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	-885550502	 [*]CNC(=[*])[*]	0.115

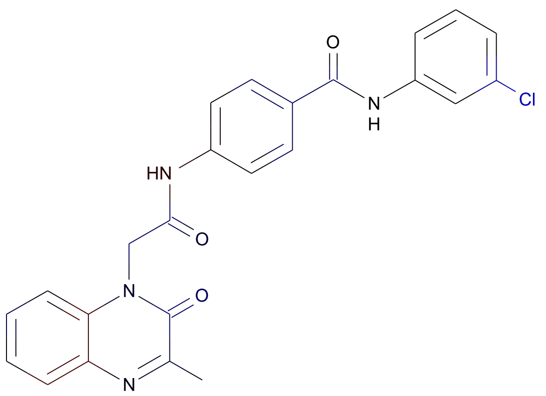
FCFP_2	3	 [*]N[*]	0.074
FCFP_2	332760439	 [*][c](:[*]):[c](:[cH]:[*])N=[*]	0.061

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	-1272798659	 [*]CCC[*]	-0.111
FCFP_2	1872154524	 [*]C(=O)[*]	-0.105
FCFP_2	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.083

17b

TOPKAT_Rat_Maximum_Tolerated_Dose_Feed



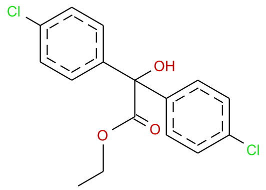
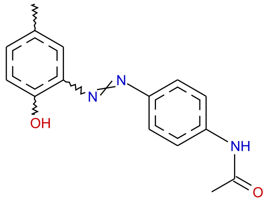
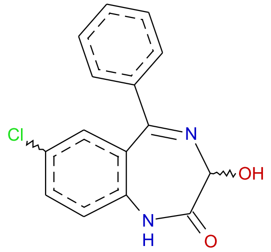
$C_{24}H_{19}ClN_4O_3$
Molecular Weight: 446.88566
ALogP: 3.125
Rotatable Bonds: 5
Acceptors: 4
Donors: 2

Model Prediction

Prediction: 0.096
Unit: g/kg_body_weight
Mahalanobis Distance: 10.657
Mahalanobis Distance p-value: 3.54e-006

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	CHLORBENZILATE	DISPERSE YELLOW 3	OXAZEPAM
Structure			
Actual Endpoint (-log C)	3.38252	2.77703	3.05262
Predicted Endpoint (-log C)	3.27894	2.80195	3.13073
Distance	0.650	0.661	0.680
Reference	NCI/NTP TR-75	NCI/NTP TR-222	NCI/NTP TR-468

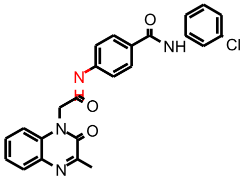
Model Applicability

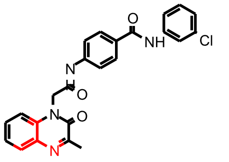
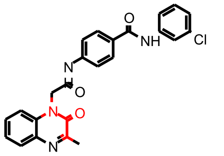
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

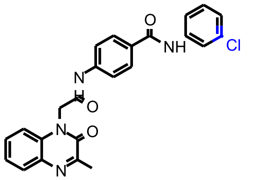
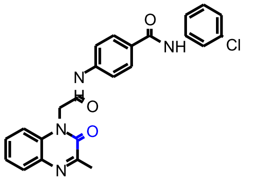
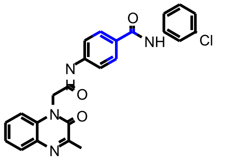
Feature Contribution

Top features for positive contribution

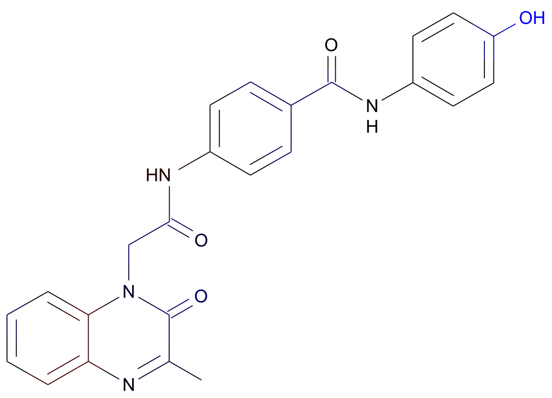
Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	3	 [*]N[*]	0.074

FCFP_2	332760439	 [*][c](:[*]):[c](:[cH]:[*])N=[*]	0.061
FCFP_2	565998553	 [*]N([*])C(=O)C(=[*])[*]	0.008

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	71476542	 [*]:[c](:[*])Cl	-0.134
FCFP_2	1872154524	 [*]C(=O)[*]	-0.105
FCFP_2	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.083

TOPKAT_Rat_Maximum_Tolerated_Dose_Feed



$C_{24}H_{20}N_4O_4$

Molecular Weight: 428.44

ALogP: 2.218

Rotatable Bonds: 5

Acceptors: 5

Donors: 3

Model Prediction

Prediction: 0.259

Unit: g/kg_body_weight

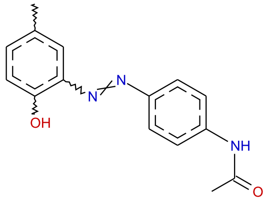
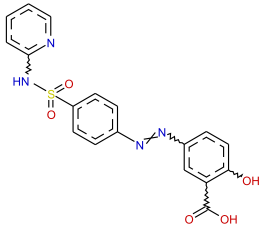
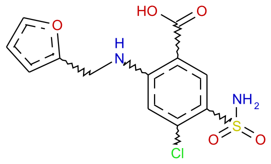
Mahalanobis Distance: 10.670

Mahalanobis Distance p-value: 3.35e-006

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	DISPERSE YELLOW 3	SALICYLAZOSULFAPYRIDINE	FUROSEMIDE
Structure			
Actual Endpoint (-log C)	2.77703	3.375	4.04236
Predicted Endpoint (-log C)	2.80195	2.80292	2.8614
Distance	0.639	0.657	0.683
Reference	NCI/NTP TR-222	NCI/NTP TR-457	NCI/NTP TR-356

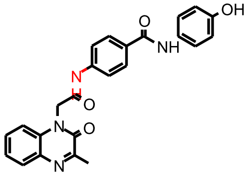
Model Applicability

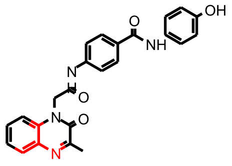
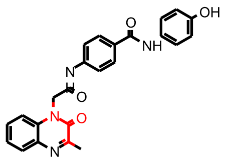
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

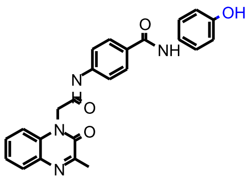
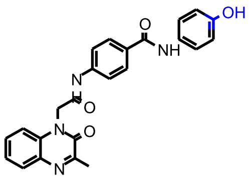
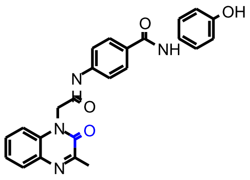
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	3	 [*]N[*]	0.074

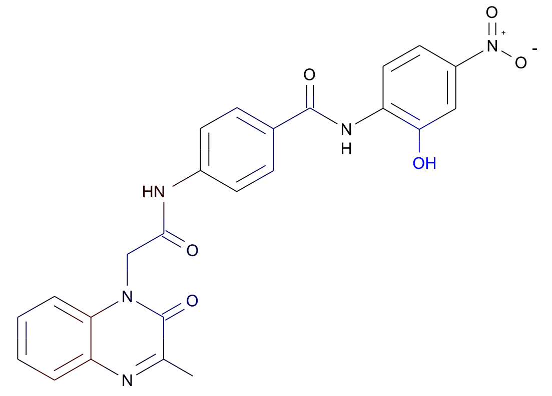
FCFP_2	332760439	 [*][c](:[*]):[c](:[cH]:[*])N=[*]	0.061
FCFP_2	565998553	 [*]N([*])C(=O)C(=[*])[*]	0.008

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	7	 [*]O	-0.214
FCFP_2	-549108873	 [*]:[c](:[*])O	-0.127
FCFP_2	1872154524	 [*]C(=O)[*]	-0.105

17d

TOPKAT_Rat_Maximum_Tolerated_Dose_Feed



$C_{24}H_{19}N_5O_6$

Molecular Weight: 473.43756

ALogP: 2.113

Rotatable Bonds: 6

Acceptors: 7

Donors: 3

Model Prediction

Prediction: 0.191

Unit: g/kg_body_weight

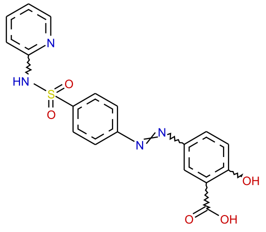
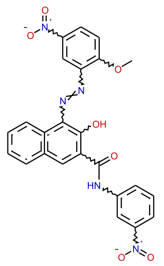
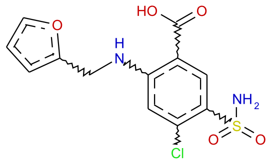
Mahalanobis Distance: 11.201

Mahalanobis Distance p-value: 3.84e-007

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	**SALICYLAZOSULFAPYRIDINE**	**C.I.PIGMENT RED 23**	**FUROSEMIDE**
Structure			
Actual Endpoint (-log C)	3.375	2.30052	4.04236
Predicted Endpoint (-log C)	2.80292	3.55333	2.8614
Distance	0.525	0.720	0.755
Reference	NCI/NTP TR-457	NCI/NTP TR-411	NCI/NTP TR-356

Model Applicability

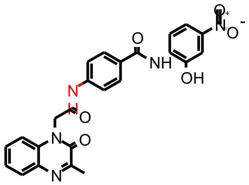
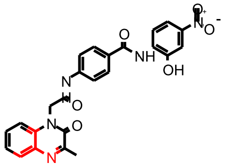
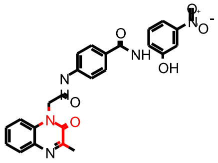
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_2 feature: 8: [*][N+](=[*])[*]
3. Unknown FCFP_2 feature: 5: [*][O-]
4. Unknown FCFP_2 feature: -828984032: [*][N+](=[*])[c](:c:[*]):c:[*]
5. Unknown FCFP_2 feature: -1338588315: [*]:[c](:[*])[N+](=O)[O-]
6. Unknown FCFP_2 feature: 1872392852: [*][N+](=O)[*]
7. Unknown FCFP_2 feature: 260476081: [*][N+](=[*])[O-]

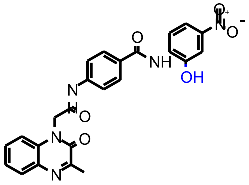
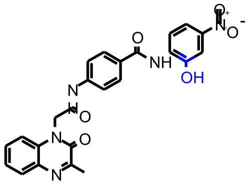
Feature Contribution

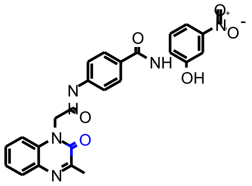
Top features for positive contribution

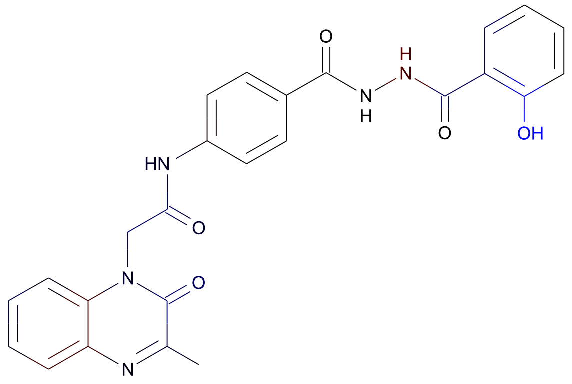
Fingerprint	Bit/Smiles	Feature Structure	Score

FCFP_2	3	 [*]N[*]	0.074
FCFP_2	332760439	 [*][c](:[*]):[c](:[cH]:[*])N=[*]	0.061
FCFP_2	565998553	 [*]N([*])C(=O)C(=[*])[*]	0.008

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	7	 [*]O	-0.214
FCFP_2	-549108873	 [*]:[c](:[*])O	-0.127

FCFP_2	1872154524	 [*]C(=O)[*]	-0.105



$C_{25}H_{21}N_5O_5$

Molecular Weight: 471.46474

ALogP: 1.475

Rotatable Bonds: 6

Acceptors: 6

Donors: 4

Model Prediction

Prediction: 0.284

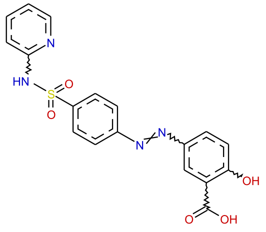
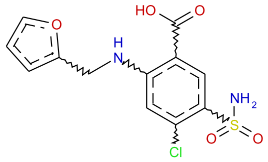
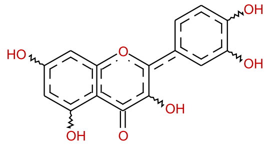
Unit: g/kg_body_weight

Mahalanobis Distance: 11.529

Mahalanobis Distance p-value: 9.57e-008

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	SALICYLAZOSULFAPYRIDINE	FUROSEMIDE	QUERCETIN
Structure			
Actual Endpoint (-log C)	3.375	4.04236	2.2016
Predicted Endpoint (-log C)	2.80292	2.8614	2.27782
Distance	0.613	0.720	0.798
Reference	NCI/NTP TR-457	NCI/NTP TR-356	NCI/NTP TR-409

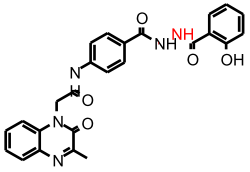
Model Applicability

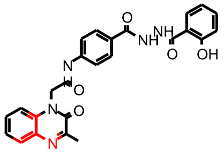
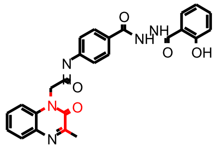
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_2 feature: -885461129: [*]NNC(=[*])[*]

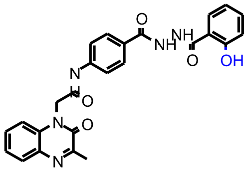
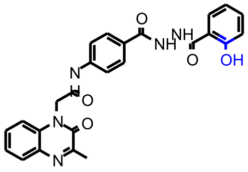
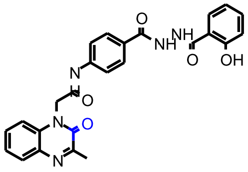
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	3	 [*]N[*]	0.074

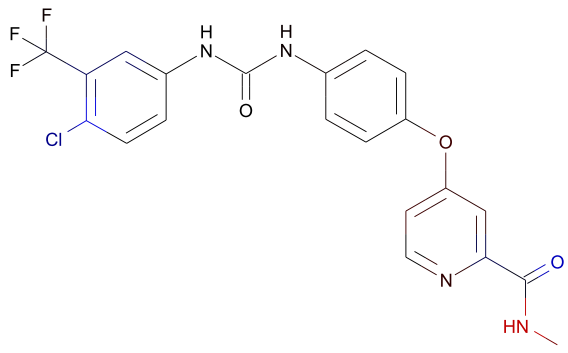
FCFP_2	332760439	 [*][c](:[*]):[c](:[cH]:[*])N=[*]	0.061
FCFP_2	565998553	 [*]N([*])C(=O)C(=[*])[*]	0.008

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	7	 [*]O	-0.214
FCFP_2	-549108873	 [*]:[c](:[*])O	-0.127
FCFP_2	1872154524	 [*]C(=O)[*]	-0.105

Sorafenib

TOPKAT_Rat_Maximum_Tolerated_Dose_Feed



$C_{21}H_{16}ClF_3N_4O_3$

Molecular Weight: 464.82494

ALogP: 4.175

Rotatable Bonds: 6

Acceptors: 4

Donors: 3

Model Prediction

Prediction: 0.089

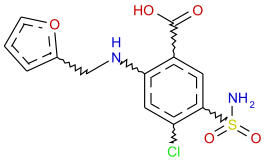
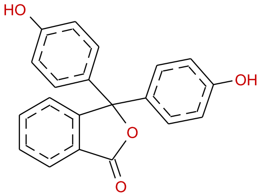
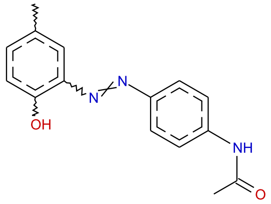
Unit: g/kg_body_weight

Mahalanobis Distance: 12.441

Mahalanobis Distance p-value: 1.76e-009

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	FUROSEMIDE	PHENOLPHTHALEIN	DISPERSE YELLOW 3
Structure			
Actual Endpoint (-log C)	4.04236	2.20184	2.77703
Predicted Endpoint (-log C)	2.8614	2.8857	2.80195
Distance	0.741	0.780	0.799
Reference	NCI/NTP TR-356	NCI/NTP TR-465	NCI/NTP TR-222

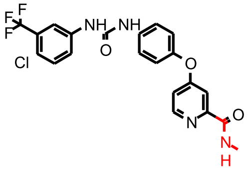
Model Applicability

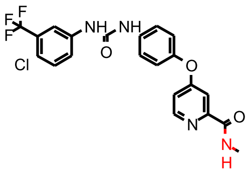
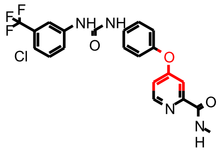
Unknown features are fingerprint features in the query molecule, but not found in the training set.

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

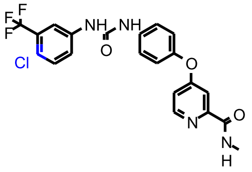
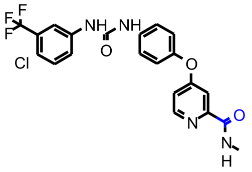
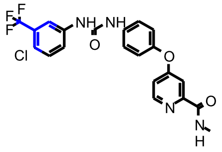
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	-885550502	 [*]CNC(=[*])[*]	0.115

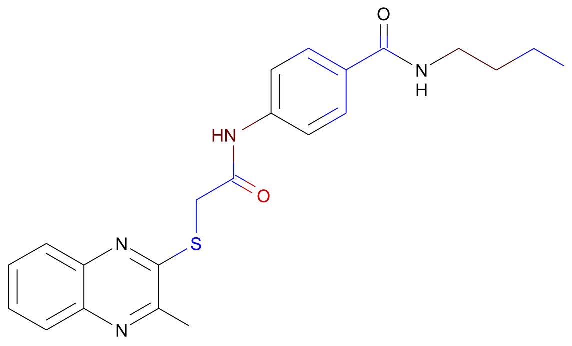
FCFP_2	3	 [*]N[*]	0.074
FCFP_2	332760439	 [*][c](:[*]):[c](:[cH]:[*])N=[*]	0.061

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	71476542	 [*]:[c](:[*])Cl	-0.134
FCFP_2	1872154524	 [*]C(=O)[*]	-0.105
FCFP_2	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.083

15a

TOPKAT_Rat_Maximum_Tolerated_Dose_Gavage



$C_{22}H_{24}N_4O_2S$

Molecular Weight: 408.51656

ALogP: 3.83

Rotatable Bonds: 8

Acceptors: 5

Donors: 2

Model Prediction

Prediction: 0.250

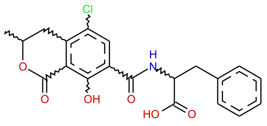
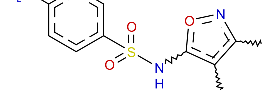
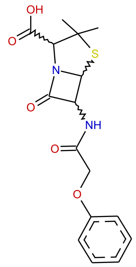
Unit: g/kg_body_weight

Mahalanobis Distance: 10.534

Mahalanobis Distance p-value: 1.04e-006

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	OCHRATOXIN	SULFISOOXAZOLE	PENICILLIN VK
Structure			
Actual Endpoint (-log C)	6.28396	2.82494	2.54455
Predicted Endpoint (-log C)	5.12358	3.0705	3.9702
Distance	0.778	0.888	0.997
Reference	NCI/NTP TR-358	NCI/NTP TR-138	NCI/NTP TR-336

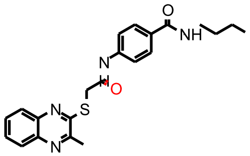
Model Applicability

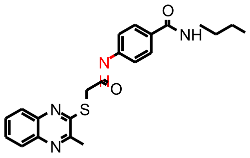
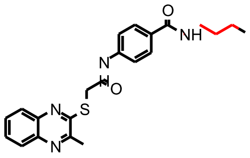
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. Num_AromaticRings out of range. Value: 3. Training min, max, mean, SD: 0, 2, 0.5625, 0.693.
2. Unknown FCFP_2 feature: -1410049896: [*]S[c](:n:[*]):[c]([*]):[*]

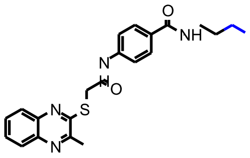
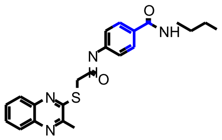
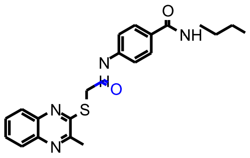
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	1	 [*]=O	0.511

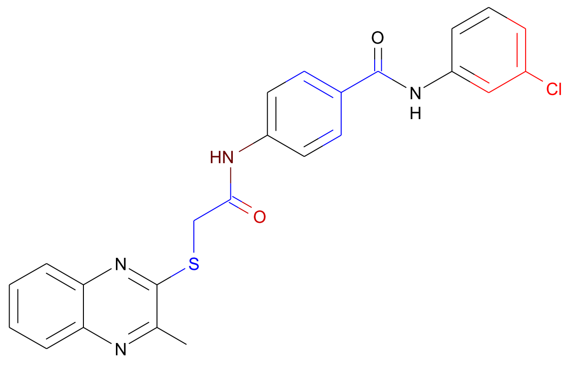
FCFP_2	3	 [*]N[*]	0.104
FCFP_2	-1272798659	 [*]CCC[*]	0.070

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	136597326	 [*]CC	-0.489
FCFP_2	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.406
FCFP_2	1872154524	 [*]C(=O)[*]	-0.307

15b

TOPKAT_Rat_Maximum_Tolerated_Dose_Gavage



$C_{24}H_{19}ClN_4O_2S$

Molecular Weight: 462.95126

ALogP: 4.743

Rotatable Bonds: 6

Acceptors: 5

Donors: 2

Model Prediction

Prediction: 0.003

Unit: g/kg_body_weight

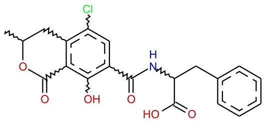
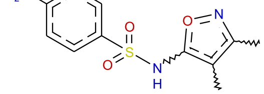
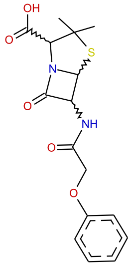
Mahalanobis Distance: 10.761

Mahalanobis Distance p-value: 5.01e-007

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	OCHRATOXIN	SULFISOOXAZOLE	PENICILLIN VK
Structure			
Actual Endpoint (-log C)	6.28396	2.82494	2.54455
Predicted Endpoint (-log C)	5.12358	3.0705	3.9702
Distance	0.951	1.080	1.290
Reference	NCI/NTP TR-358	NCI/NTP TR-138	NCI/NTP TR-336

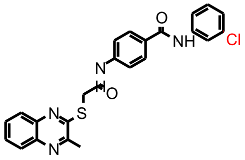
Model Applicability

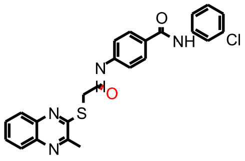
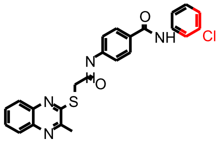
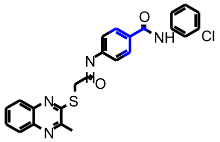
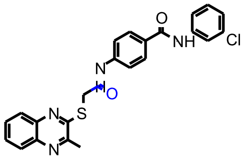
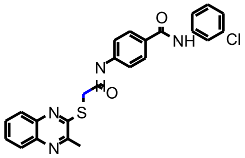
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. Molecular_Weight out of range. Value: 462.95. Training min, max, mean, SD: 68.074, 434.63, 171.13, 85.06.
2. Num_AromaticRings out of range. Value: 4. Training min, max, mean, SD: 0, 2, 0.5625, 0.693.
3. OPS PC5 out of range. Value: -3.5338. Training min, max, SD, explained variance: -3.4, 4.1587, 1.489, 0.0686.
4. OPS PC7 out of range. Value: -2.8709. Training min, max, SD, explained variance: -2.8003, 2.9332, 1.16, 0.0416.
5. Unknown FCFP_2 feature: -1410049896: [*]S[c](:n:[*]):[c]([*]):[*]

Feature Contribution

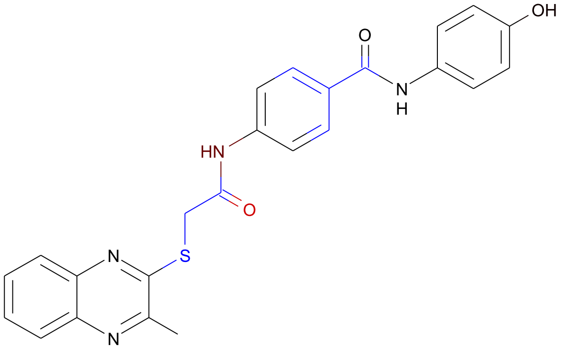
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	32	 [*]Cl	0.526

FCFP_2	1	 [*]=O	0.511
FCFP_2	367998008	 [*]:[cH]:[c](Cl):[cH]:[*]	0.413
Top Features for negative contribution			
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.406
FCFP_2	1872154524	 [*]C(=O)[*]	-0.307
FCFP_2	0	 [*]C[*]	-0.290

15c

TOPKAT_Rat_Maximum_Tolerated_Dose_Gavage



$C_{24}H_{20}N_4O_3S$

Molecular Weight: 444.5056

ALogP: 3.836

Rotatable Bonds: 6

Acceptors: 6

Donors: 3

Model Prediction

Prediction: 0.014

Unit: g/kg_body_weight

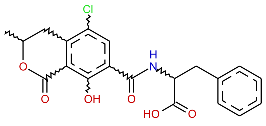
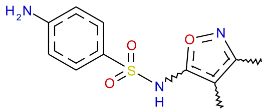
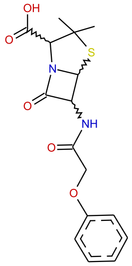
Mahalanobis Distance: 12.626

Mahalanobis Distance p-value: 1.14e-009

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	OCHRATOXIN	SULFISOOXAZOLE	PENICILLIN VK
Structure			
Actual Endpoint (-log C)	6.28396	2.82494	2.54455
Predicted Endpoint (-log C)	5.12358	3.0705	3.9702
Distance	0.873	1.073	1.239
Reference	NCI/NTP TR-358	NCI/NTP TR-138	NCI/NTP TR-336

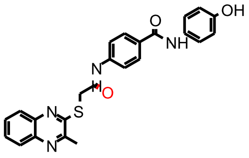
Model Applicability

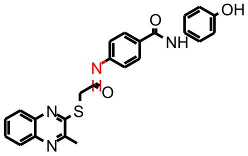
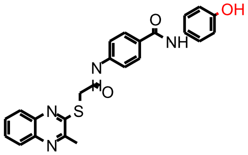
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. Molecular_Weight out of range. Value: 444.51. Training min, max, mean, SD: 68.074, 434.63, 171.13, 85.06.
2. Num_AromaticRings out of range. Value: 4. Training min, max, mean, SD: 0, 2, 0.5625, 0.693.
3. OPS PC1 out of range. Value: 8.7573. Training min, max, SD, explained variance: -4.0008, 7.9165, 2.861, 0.2531.
4. Unknown FCFP_2 feature: -1410049896: [*]S[c](:n:[*]):[c]([*]):[*]

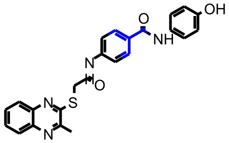
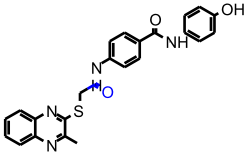
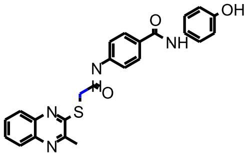
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	1	 [*]=O	0.511

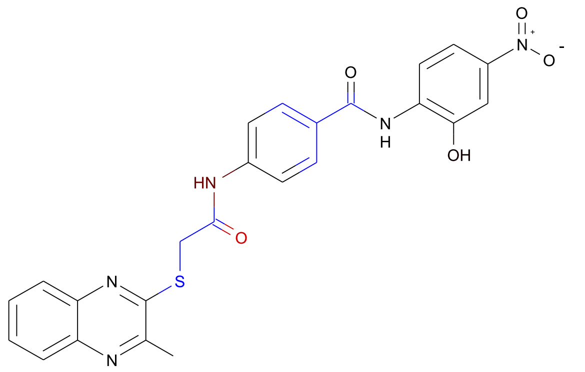
FCFP_2	3	 [*]N[*]	0.104
FCFP_2	7	 [*]O	0.014

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.406
FCFP_2	1872154524	 [*]C(=O)[*]	-0.307
FCFP_2	0	 [*]C[*]	-0.290

15d

TOPKAT_Rat_Maximum_Tolerated_Dose_Gavage



$C_{24}H_{19}N_5O_5S$

Molecular Weight: 489.50316

ALogP: 3.731

Rotatable Bonds: 7

Acceptors: 8

Donors: 3

Model Prediction

Prediction: 0.004

Unit: g/kg_body_weight

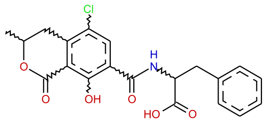
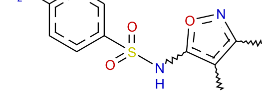
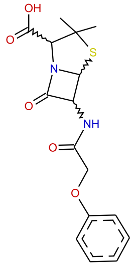
Mahalanobis Distance: 12.287

Mahalanobis Distance p-value: 3.45e-009

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	**OCHRATOXIN**	**SULFISOOXAZOLE**	**PENICILLIN VK**
Structure			
Actual Endpoint (-log C)	6.28396	2.82494	2.54455
Predicted Endpoint (-log C)	5.12358	3.0705	3.9702
Distance	1.098	1.356	1.399
Reference	NCI/NTP TR-358	NCI/NTP TR-138	NCI/NTP TR-336

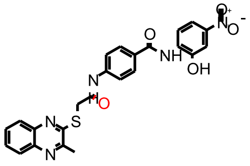
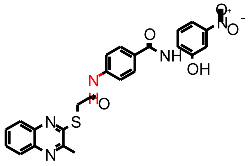
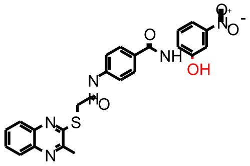
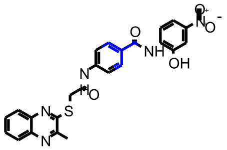
Model Applicability

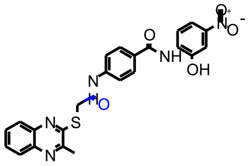
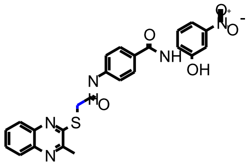
Unknown features are fingerprint features in the query molecule, but not found in the training set.

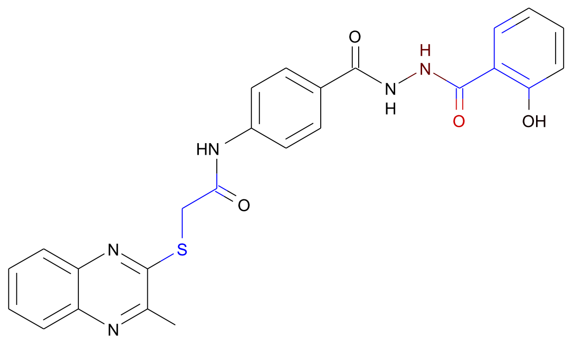
1. Molecular_Weight out of range. Value: 489.5. Training min, max, mean, SD: 68.074, 434.63, 171.13, 85.06.
2. Num_H_Acceptors out of range. Value: 8. Training min, max, mean, SD: 0, 6, 1.6146, 1.644.
3. Num_AromaticRings out of range. Value: 4. Training min, max, mean, SD: 0, 2, 0.5625, 0.693.
4. Molecular_PolarSASA out of range. Value: 265.99. Training min, max, mean, SD: 0, 223.97, 50.816, 55.15.
5. Molecular_PolarSurfaceArea out of range. Value: 175.33. Training min, max, mean, SD: 0, 138.03, 28.978, 32.1.
6. OPS PC1 out of range. Value: 10.108. Training min, max, SD, explained variance: -4.0008, 7.9165, 2.861, 0.2531.
7. Unknown FCFP_2 feature: 8: [*][N+](=[*])[*]
8. Unknown FCFP_2 feature: 5: [*][O-]
9. Unknown FCFP_2 feature: -1410049896: [*]S[c](:n:[*]):[c]([*]):[*]
10. Unknown FCFP_2 feature: -828984032: [*][N+](=[*])[c](:c:[*]):c:[*]
11. Unknown FCFP_2 feature: -1338588315: [*]:[c](:[*])[N+](=O)[O-]
12. Unknown FCFP_2 feature: 1872392852: [*][N+](=O)[*]
13. Unknown FCFP_2 feature: 260476081: [*][N+](=[*])[O-]

Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	1	 [*]=O	0.511
FCFP_2	3	 [*]N[*]	0.104
FCFP_2	7	 [*]O	0.014
Top Features for negative contribution			
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.406

FCFP_2	1872154524	 [*]C(=O)[*]	-0.307
FCFP_2	0	 [*]C[*]	-0.290



$C_{25}H_{21}N_5O_4S$

Molecular Weight: 487.53034

ALogP: 3.093

Rotatable Bonds: 7

Acceptors: 7

Donors: 4

Model Prediction

Prediction: 0.012

Unit: g/kg_body_weight

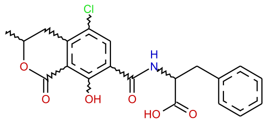
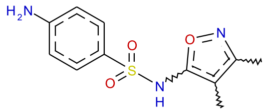
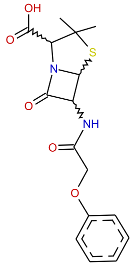
Mahalanobis Distance: 12.599

Mahalanobis Distance p-value: 1.24e-009

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	OCHRATOXIN	SULFISOOXAZOLE	PENICILLIN VK
Structure			
Actual Endpoint (-log C)	6.28396	2.82494	2.54455
Predicted Endpoint (-log C)	5.12358	3.0705	3.9702
Distance	1.016	1.276	1.366
Reference	NCI/NTP TR-358	NCI/NTP TR-138	NCI/NTP TR-336

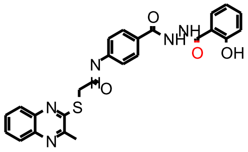
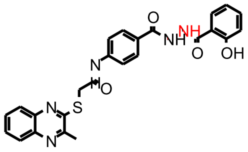
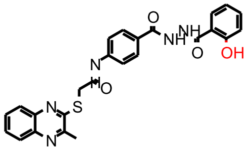
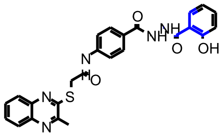
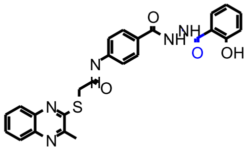
Model Applicability

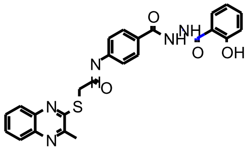
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. Molecular_Weight out of range. Value: 487.53. Training min, max, mean, SD: 68.074, 434.63, 171.13, 85.06.
2. Num_H_Donors out of range. Value: 4. Training min, max, mean, SD: 0, 3, 0.4375, 0.8311.
3. Num_H_Acceptors out of range. Value: 7. Training min, max, mean, SD: 0, 6, 1.6146, 1.644.
4. Num_AromaticRings out of range. Value: 4. Training min, max, mean, SD: 0, 2, 0.5625, 0.693.
5. Molecular_PolarSASA out of range. Value: 234.94. Training min, max, mean, SD: 0, 223.97, 50.816, 55.15.
6. Molecular_PolarSurfaceArea out of range. Value: 158.61. Training min, max, mean, SD: 0, 138.03, 28.978, 32.1.
7. OPS PC1 out of range. Value: 9.9011. Training min, max, SD, explained variance: -4.0008, 7.9165, 2.861, 0.2531.
8. Unknown FCFP_2 feature: -885461129: [*]NNC(=[*])[*]
9. Unknown FCFP_2 feature: -1410049896: [*]S[c](:n:[*]):[c]([*]):[*]

Feature Contribution

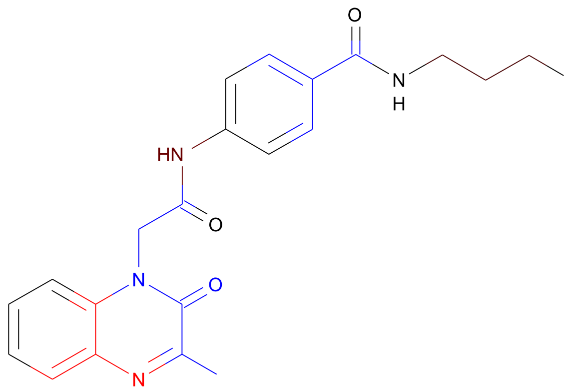
Top features for positive contribution			
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**

FCFP_2	1	 [*]=O	0.511
FCFP_2	3	 [*]N[*]	0.104
FCFP_2	7	 [*]O	0.014
Top Features for negative contribution			
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.406
FCFP_2	1872154524	 [*]C(=O)[*]	-0.307

FCFP_2	0	 [*]C[*]	-0.290

17a

TOPKAT_Rat_Maximum_Tolerated_Dose_Gavage



$C_{22}H_{24}N_4O_3$

Molecular Weight: 392.45096

ALogP: 2.212

Rotatable Bonds: 7

Acceptors: 4

Donors: 2

Model Prediction

Prediction: 0.611

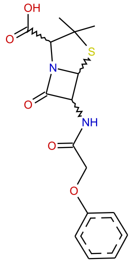
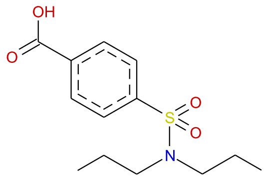
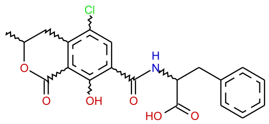
Unit: g/kg_body_weight

Mahalanobis Distance: 9.374

Mahalanobis Distance p-value: 3.95e-005

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	PENICILLIN VK	PROBENECID	OCHRATOXIN
Structure			
Actual Endpoint (-log C)	2.54455	2.85333	6.28396
Predicted Endpoint (-log C)	3.9702	2.4258	5.12358
Distance	0.760	0.763	0.768
Reference	NCI/NTP TR-336	NCI/NTP TR-395	NCI/NTP TR-358

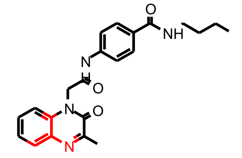
Model Applicability

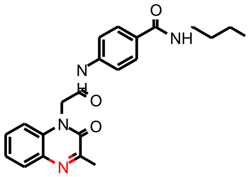
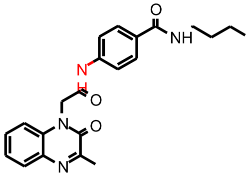
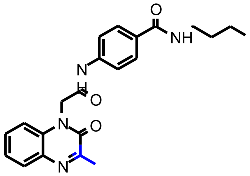
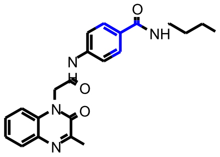
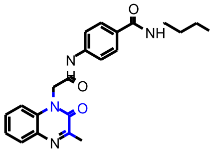
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. OPS PC6 out of range. Value: -2.8056. Training min, max, SD, explained variance: -2.4321, 2.9885, 1.256, 0.0488.
2. Unknown FCFP_2 feature: 580453787: [*]C(=N[c](:[*]):[*])[*]

Feature Contribution

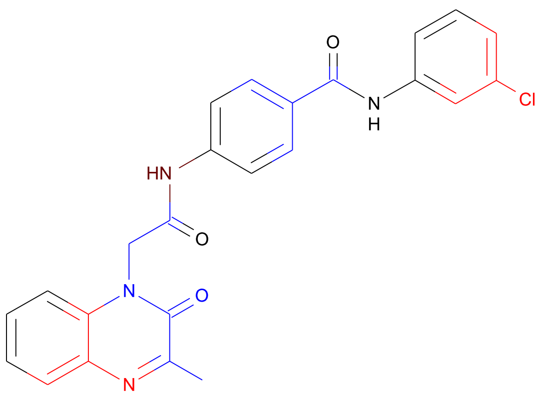
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	332760439	 [*][c](:[*]):[c](:[cH]:[*])N=[*]	0.672

FCFP_2	1	 [*]=O	0.511
FCFP_2	3	 [*]N[*]	0.104
Top Features for negative contribution			
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	136597326	 [*]CC	-0.489
FCFP_2	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.406
FCFP_2	565998553	 [*]N([*])C(=O)C(=[*])[*]	-0.348

17b

TOPKAT_Rat_Maximum_Tolerated_Dose_Gavage



$C_{24}H_{19}ClN_4O_3$

Molecular Weight: 446.88566

ALogP: 3.125

Rotatable Bonds: 5

Acceptors: 4

Donors: 2

Model Prediction

Prediction: 0.022

Unit: g/kg_body_weight

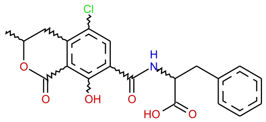
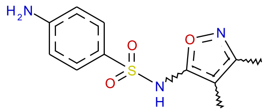
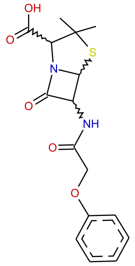
Mahalanobis Distance: 9.375

Mahalanobis Distance p-value: 3.94e-005

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	OCHRATOXIN	SULFISOOXAZOLE	PENICILLIN VK
Structure			
Actual Endpoint (-log C)	6.28396	2.82494	2.54455
Predicted Endpoint (-log C)	5.12358	3.0705	3.9702
Distance	0.777	0.899	1.005
Reference	NCI/NTP TR-358	NCI/NTP TR-138	NCI/NTP TR-336

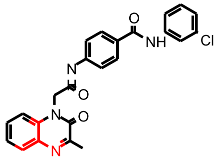
Model Applicability

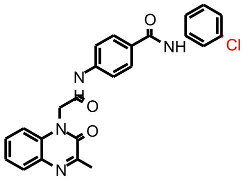
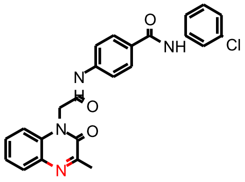
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. Molecular_Weight out of range. Value: 446.89. Training min, max, mean, SD: 68.074, 434.63, 171.13, 85.06.
2. Num_AromaticRings out of range. Value: 3. Training min, max, mean, SD: 0, 2, 0.5625, 0.693.
3. OPS PC10 out of range. Value: 3.2414. Training min, max, SD, explained variance: -3.9696, 2.3971, 0.982, 0.0298.
4. Unknown FCFP_2 feature: 580453787: [*]C(=N[c](:[*]):[*])[*]

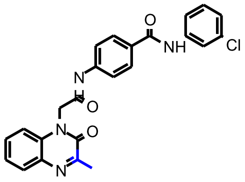
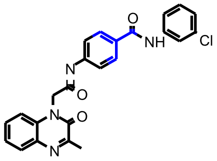
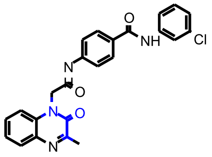
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	332760439	 [*][c](:[*]):[c](:[cH]:[*])N=[*]	0.672

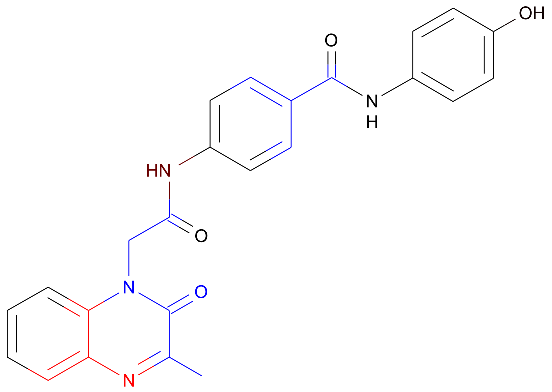
FCFP_2	32	 [*]Cl	0.526
FCFP_2	1	 [*]=O	0.511

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	136597326	 [*]CC	-0.489
FCFP_2	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.406
FCFP_2	565998553	 [*]N([*])C(=O)C(=[*])[*]	-0.348

17c

TOPKAT_Rat_Maximum_Tolerated_Dose_Gavage



$C_{24}H_{20}N_4O_4$

Molecular Weight: 428.44

ALogP: 2.218

Rotatable Bonds: 5

Acceptors: 5

Donors: 3

Model Prediction

Prediction: 0.104

Unit: g/kg_body_weight

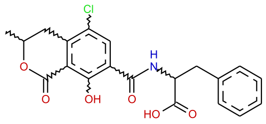
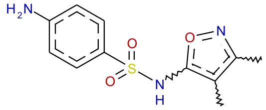
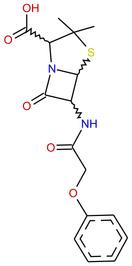
Mahalanobis Distance: 9.317

Mahalanobis Distance p-value: 4.69e-005

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	**OCHRATOXIN**	**SULFISOOXAZOLE**	**PENICILLIN VK**
Structure			
Actual Endpoint (-log C)	6.28396	2.82494	2.54455
Predicted Endpoint (-log C)	5.12358	3.0705	3.9702
Distance	0.627	0.864	0.916
Reference	NCI/NTP TR-358	NCI/NTP TR-138	NCI/NTP TR-336

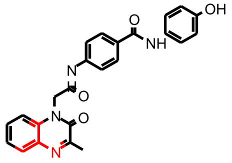
Model Applicability

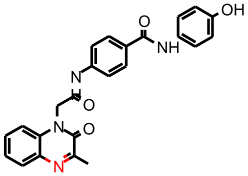
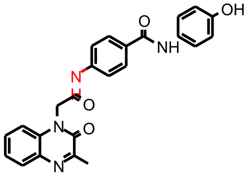
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. Num_AromaticRings out of range. Value: 3. Training min, max, mean, SD: 0, 2, 0.5625, 0.693.
2. OPS PC1 out of range. Value: 8.6203. Training min, max, SD, explained variance: -4.0008, 7.9165, 2.861, 0.2531.
3. OPS PC10 out of range. Value: 2.4631. Training min, max, SD, explained variance: -3.9696, 2.3971, 0.982, 0.0298.
4. Unknown FCFP_2 feature: 580453787: [*]C(=N[c](:[*]):[*])[*]

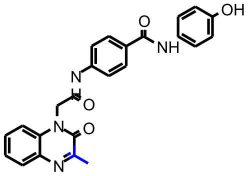
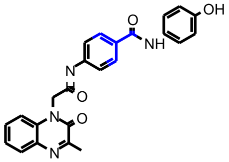
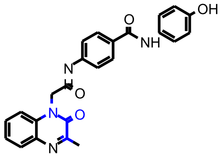
Feature Contribution

Top features for positive contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	332760439	 [*][c](:[*]):[c](:[cH]:[*])N=[*]	0.672

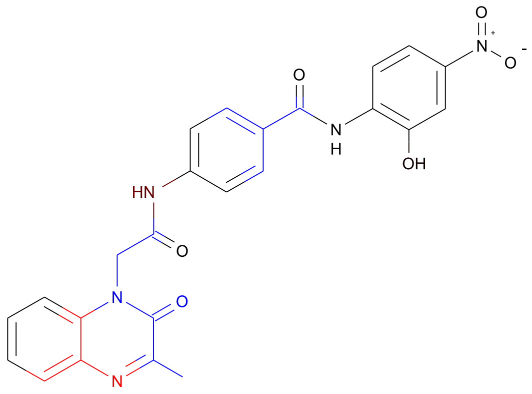
FCFP_2	1	 [*]=O	0.511
FCFP_2	3	 [*]N[*]	0.104

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	136597326	 [*]CC	-0.489
FCFP_2	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.406
FCFP_2	565998553	 [*]N([*])C(=O)C(=[*])[*]	-0.348

17d

TOPKAT_Rat_Maximum_Tolerated_Dose_Gavage



$C_{24}H_{19}N_5O_6$

Molecular Weight: 473.43756

ALogP: 2.113

Rotatable Bonds: 6

Acceptors: 7

Donors: 3

Model Prediction

Prediction: 0.032

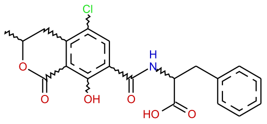
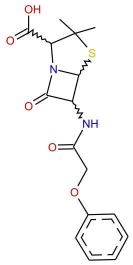
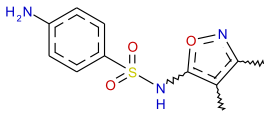
Unit: g/kg_body_weight

Mahalanobis Distance: 9.408

Mahalanobis Distance p-value: 3.56e-005

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	OCHRATOXIN	PENICILLIN VK	SULFISOOXAZOLE
Structure			
Actual Endpoint (-log C)	6.28396	2.54455	2.82494
Predicted Endpoint (-log C)	5.12358	3.9702	3.0705
Distance	0.795	1.036	1.122
Reference	NCI/NTP TR-358	NCI/NTP TR-336	NCI/NTP TR-138

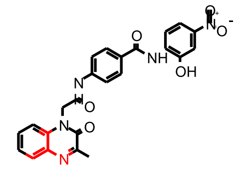
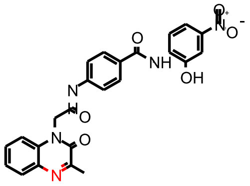
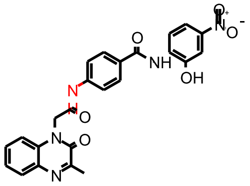
Model Applicability

Unknown features are fingerprint features in the query molecule, but not found in the training set.

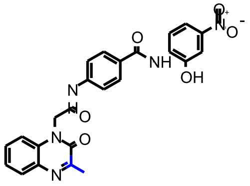
1. Molecular_Weight out of range. Value: 473.44. Training min, max, mean, SD: 68.074, 434.63, 171.13, 85.06.
2. Num_H_Acceptors out of range. Value: 7. Training min, max, mean, SD: 0, 6, 1.6146, 1.644.
3. Num_AromaticRings out of range. Value: 3. Training min, max, mean, SD: 0, 2, 0.5625, 0.693.
4. Molecular_PolarSASA out of range. Value: 240.14. Training min, max, mean, SD: 0, 223.97, 50.816, 55.15.
5. Molecular_PolarSurfaceArea out of range. Value: 156.91. Training min, max, mean, SD: 0, 138.03, 28.978, 32.1.
6. OPS PC1 out of range. Value: 9.9705. Training min, max, SD, explained variance: -4.0008, 7.9165, 2.861, 0.2531.
7. Unknown FCFP_2 feature: 8: [*][N+](=[*])[*]
8. Unknown FCFP_2 feature: 5: [*][O-]
9. Unknown FCFP_2 feature: 580453787: [*]C(=N[c](:[*]):[*])[*]
10. Unknown FCFP_2 feature: -828984032: [*][N+](=[*])[c](:c:[*]):c:[*]
11. Unknown FCFP_2 feature: -1338588315: [*]:[c](:[*])[N+](=O)[O-]
12. Unknown FCFP_2 feature: 1872392852: [*][N+](=O)[*]
13. Unknown FCFP_2 feature: 260476081: [*][N+](=[*])[O-]

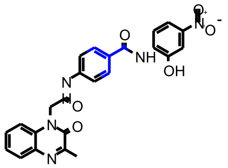
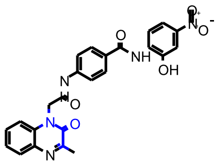
Feature Contribution

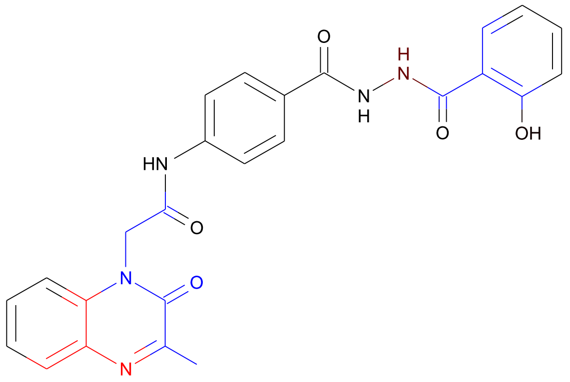
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	332760439	 [*][c](:[*]):[c](:[cH]:[*])N=[*]	0.672
FCFP_2	1	 [*]=O	0.511
FCFP_2	3	 [*]N[*]	0.104

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
FCFP_2	136597326	 [*]CC	-0.489

FCFP_2	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.406
FCFP_2	565998553	 [*]N([*])C(=O)C(=[*])[*]	-0.348



$C_{25}H_{21}N_5O_5$
Molecular Weight: 471.46474
ALogP: 1.475
Rotatable Bonds: 6
Acceptors: 6
Donors: 4

Model Prediction

Prediction: 0.093

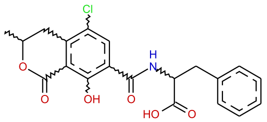
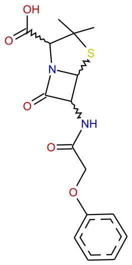
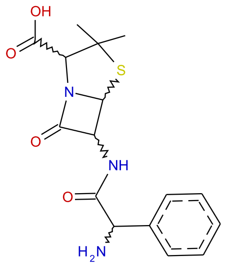
Unit: g/kg_body_weight

Mahalanobis Distance: 10.392

Mahalanobis Distance p-value: 1.64e-006

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	OCHRATOXIN	PENICILLIN VK	AMPICILLIN TRIHYDRATE
Structure			
Actual Endpoint (-log C)	6.28396	2.54455	2.36724
Predicted Endpoint (-log C)	5.12358	3.9702	2.27651
Distance	0.767	1.047	1.082
Reference	NCI/NTP TR-358	NCI/NTP TR-336	NCI/NTP TR-318

Model Applicability

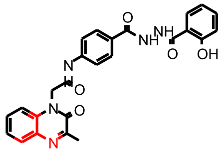
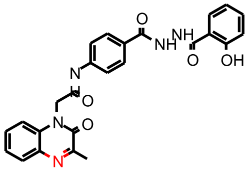
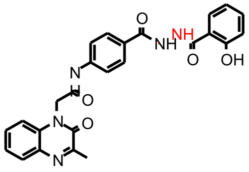
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. Molecular_Weight out of range. Value: 471.46. Training min, max, mean, SD: 68.074, 434.63, 171.13, 85.06.
2. Num_H_Donors out of range. Value: 4. Training min, max, mean, SD: 0, 3, 0.4375, 0.8311.
3. Num_AromaticRings out of range. Value: 3. Training min, max, mean, SD: 0, 2, 0.5625, 0.693.
4. Molecular_PolarSurfaceArea out of range. Value: 140.19. Training min, max, mean, SD: 0, 138.03, 28.978, 32.1.
5. OPS PC1 out of range. Value: 9.764. Training min, max, SD, explained variance: -4.0008, 7.9165, 2.861, 0.2531.
6. Unknown FCFP_2 feature: 580453787: [*]C(=N[c](:[*]):[*])[*]
7. Unknown FCFP_2 feature: -885461129: [*]NNC(=[*])[*]

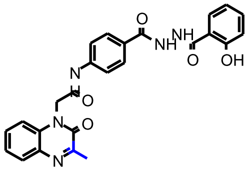
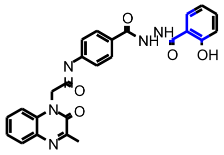
Feature Contribution

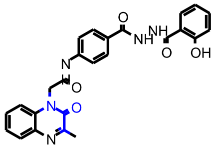
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score

FCFP_2	332760439	 [*][c](:[*]):[c](:[cH]:[*])N=[*]	0.672
FCFP_2	1	 [*]=O	0.511
FCFP_2	3	 [*]N[*]	0.104

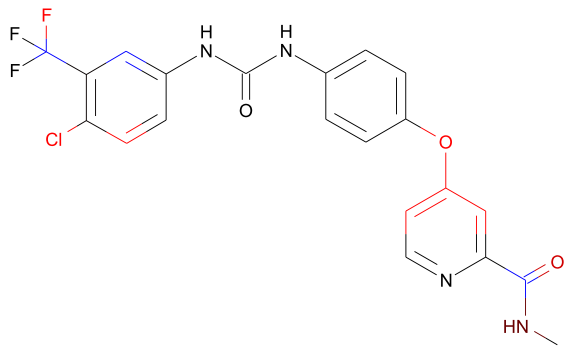
Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	136597326	 [*]CC	-0.489
FCFP_2	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.406

FCFP_2	565998553	 [*]N([*])C(=O)C(=[*])[*]	-0.348

Sorafenib

TOPKAT_Rat_Maximum_Tolerated_Dose_Gavage



$C_{21}H_{16}ClF_3N_4O_3$

Molecular Weight: 464.82494

ALogP: 4.175

Rotatable Bonds: 6

Acceptors: 4

Donors: 3

Model Prediction

Prediction: 0.001

Unit: g/kg_body_weight

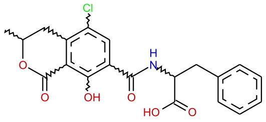
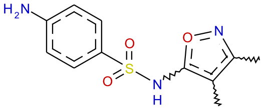
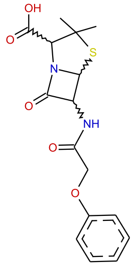
Mahalanobis Distance: 12.193

Mahalanobis Distance p-value: 4.69e-009

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	OCHRATOXIN	SULFISOOXAZOLE	PENICILLIN VK
Structure			
Actual Endpoint (-log C)	6.28396	2.82494	2.54455
Predicted Endpoint (-log C)	5.12358	3.0705	3.9702
Distance	0.758	0.997	1.159
Reference	NCI/NTP TR-358	NCI/NTP TR-138	NCI/NTP TR-336

Model Applicability

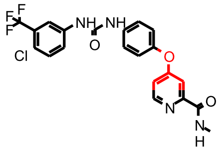
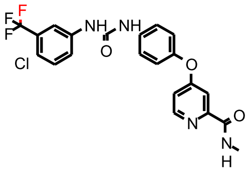
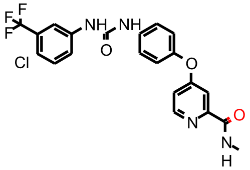
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. Molecular_Weight out of range. Value: 464.82. Training min, max, mean, SD: 68.074, 434.63, 171.13, 85.06.
2. Num_AromaticRings out of range. Value: 3. Training min, max, mean, SD: 0, 2, 0.5625, 0.693.
3. OPS PC5 out of range. Value: -3.5737. Training min, max, SD, explained variance: -3.4, 4.1587, 1.489, 0.0686.
4. OPS PC7 out of range. Value: -3.8342. Training min, max, SD, explained variance: -2.8003, 2.9332, 1.16, 0.0416.
5. Unknown FCFP_2 feature: 1499521844: [*]NC(=O)N[*]
6. Unknown FCFP_2 feature: -1029533685: [*]:[c](:[*])C(F)(F)F
7. Unknown FCFP_2 feature: 136686699: [*]NC

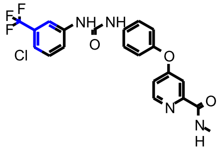
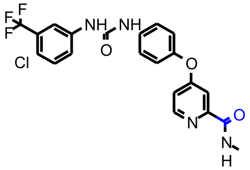
Feature Contribution

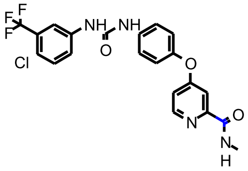
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score

FCFP_2	332760439	 [*][c](:[*]):[c](:[cH]:[*])N=[*]	0.672
FCFP_2	32	 [*]Cl	0.526
FCFP_2	1	 [*]=O	0.511

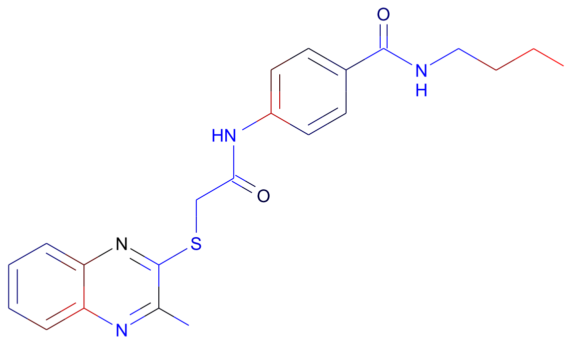
Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
FCFP_2	203677720	 [*]C(=[*])[c](:[cH]:[*]):[cH]:[*]	-0.406
FCFP_2	1872154524	 [*]C(=O)[*]	-0.307

FCFP_2	0	 [*]C[*]	-0.290

15a

TOPKAT_Rat_Oral_LD50



$C_{22}H_{24}N_4O_2S$

Molecular Weight: 408.51656

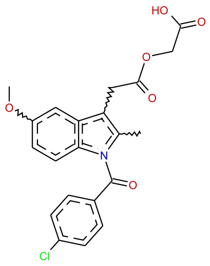
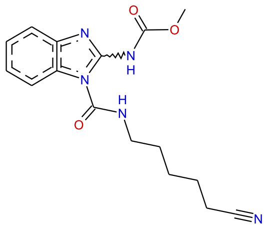
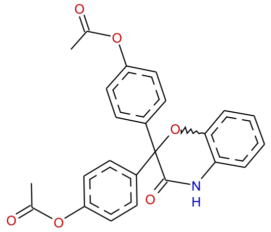
ALogP: 3.83

Rotatable Bonds: 8

Acceptors: 5

Donors: 2

Structural Similar Compounds

Name	ACEMETACIN	CARBAMIC ACID; [1-[(5-CYANOPENTYL)CARBAMOYL]BENZIMIDAZOL-2-YL]-; METHYL ESTER	bis-OXATIN ACETATE
Structure			
Actual Endpoint (-log C)	4.235	2.12	1.717
Predicted Endpoint (-log C)	3.39415	1.78415	2.40947
Distance	0.614	0.643	0.647
Reference	ARZNAD 30;1398;80	85ARAE 4;118;76/77	NIIRDN 6;609;82

Model Prediction

Prediction: 12.496

Unit: g/kg_body_weight

Mahalanobis Distance: 21.154

Mahalanobis Distance p-value: 6.56e-013

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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.

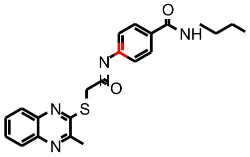
Model Applicability

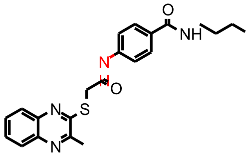
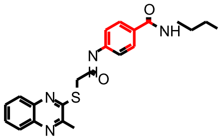
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_6 feature: 16: [*][c](:[*]):[*]
3. Unknown FCFP_6 feature: 1618154665: [*][c](:[*]):[cH]:[cH]:[*]
4. Unknown FCFP_6 feature: 1747237384: [*][c](:[*]):n:[c](:[*]):[*]
5. Unknown FCFP_6 feature: -1410049896: [*]S[c](:n:[*]):[c]([*]):[*]

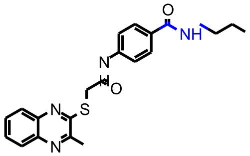
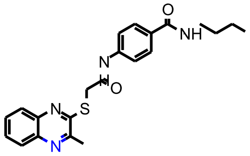
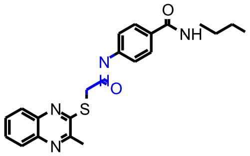
Feature Contribution

Top features for positive contribution

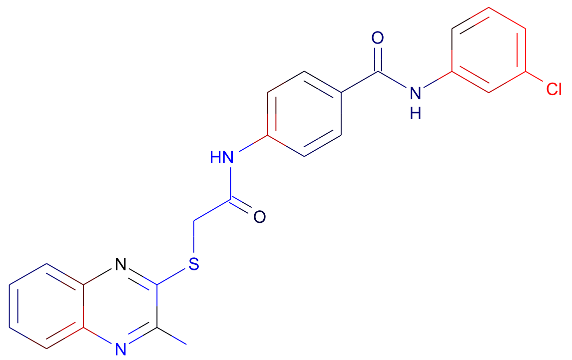
Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	642810091	 [*][c](:[*]):[*]	0.281

ECFP_6	-1897341097	 [*]N[*]	0.216
ECFP_6	1444581947	 [*]C(=[*])[c]1:[cH]:[*]:[c]([*]):[cH]:[cH]:1	0.163

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
ECFP_6	497523368	 [*]CNC(=[*])[*]	-0.301
ECFP_6	655739385	 [*]:n:[*]	-0.239
FCFP_6	566058135	 [*]CC(=O)N[*]	-0.216

15b



$C_{24}H_{19}ClN_4O_2S$

Molecular Weight: 462.95126

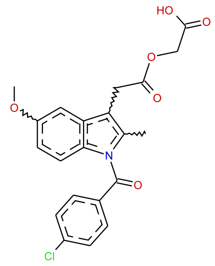
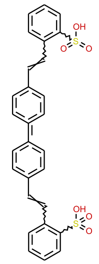
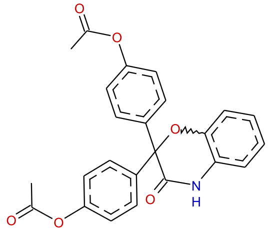
ALogP: 4.743

Rotatable Bonds: 6

Acceptors: 5

Donors: 2

Structural Similar Compounds

Name	**ACEMETACIN**	**BENZENESULFONIC ACID; 2;2'-(4;4'-BIPHENYLYLENEDIVINYLENE)DI-; DISODIUM SALT (Na STRIPPED)**	**bis-OXATIN ACETATE**
Structure			
Actual Endpoint (-log C)	4.235	1.968	1.717
Predicted Endpoint (-log C)	3.39415	1.72109	2.40947
Distance	0.698	0.716	0.742
Reference	ARZNAD 30;1398;80	MVCRB3 2;193;73	NIIRDN 6;609;82

Model Prediction

Prediction: 4.969

Unit: g/kg_body_weight

Mahalanobis Distance: 20.014

Mahalanobis Distance p-value: 5.71e-009

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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

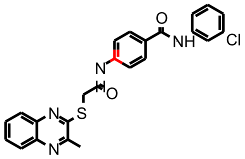
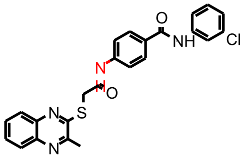
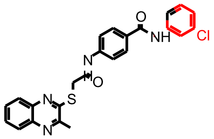
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_6 feature: 16: [*][c](:[*]):[*]
3. Unknown FCFP_6 feature: 1618154665: [*][c](:[*]):[cH]:[cH]:[*]
4. Unknown FCFP_6 feature: 1747237384: [*][c](:[*]):n:[c](:[*]):[*]
5. Unknown FCFP_6 feature: -1410049896: [*]S[c](:n:[*]):[c]([*]):[*]
6. Unknown FCFP_6 feature: 71476542: [*]:[c](:[*])Cl

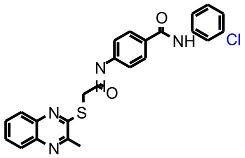
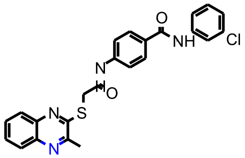
Feature Contribution

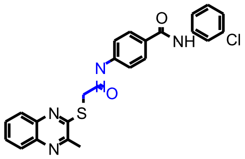
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score

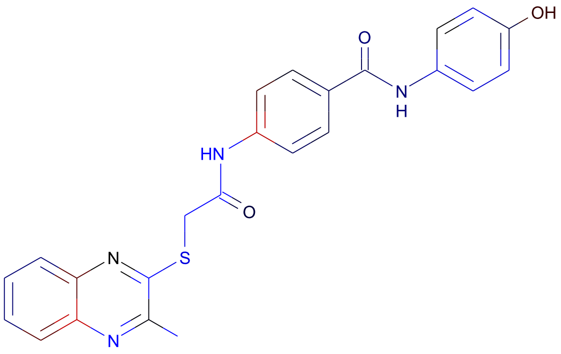
ECFP_6	642810091	 [*][c](:[*]):[*]	0.281
ECFP_6	-1897341097	 [*]N[*]	0.216
ECFP_6	577592657	 [*][c]1:[*]:[cH]:[cH] :[c](Cl):[cH]:1	0.194

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
ECFP_6	-817402818	 [*]Cl	-0.263
ECFP_6	655739385	 [*]:n:[*]	-0.239

FCFP_6	566058135	 [*]CC(=O)N[*]	-0.216

15c



$C_{24}H_{20}N_4O_3S$

Molecular Weight: 444.5056

ALogP: 3.836

Rotatable Bonds: 6

Acceptors: 6

Donors: 3

Model Prediction

Prediction: 5.932

Unit: g/kg_body_weight

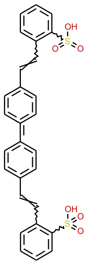
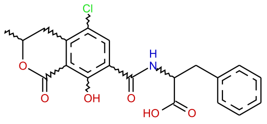
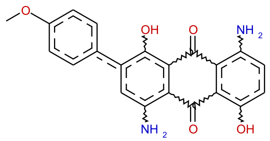
Mahalanobis Distance: 20.758

Mahalanobis Distance p-value: 1.85e-011

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	BENZENESULFONIC ACID; 2;2'-(4;4'-BIPHENYLYLENEDIVINYLENE)DI-; DISODIUM SALT (Na STRIPPED)	OCHRATOXIN A	ANTHRAQUINONE; 1;5-DIAMINO-4;8-DIHYDROXY-3-(p-METHOXYPHENYL)-
Structure			
Actual Endpoint (-log C)	1.968	4.305	1.771
Predicted Endpoint (-log C)	1.72109	3.03558	2.1122
Distance	0.748	0.825	0.829
Reference	MVCRB3 2;193;73	FCTXAV 6;479;68	28ZPAK -;245;72

Model Applicability

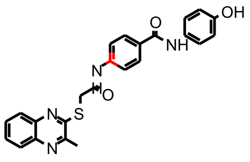
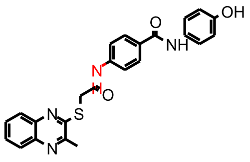
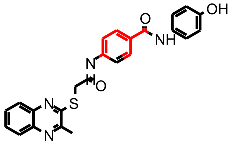
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_6 feature: 16: [*][c](:[*]):[*]
3. Unknown FCFP_6 feature: 1618154665: [*][c](:[*]):[cH]:[cH]:[*]
4. Unknown FCFP_6 feature: 1747237384: [*][c](:[*]):n:[c](:[*]):[*]
5. Unknown FCFP_6 feature: -1410049896: [*]S[c](:n:[*]):[c]([*]):[*]
6. Unknown FCFP_6 feature: 74595001: [*]:[cH]:[c](O):[cH]:[*]
7. Unknown FCFP_6 feature: -549108873: [*]:[c](:[*])O

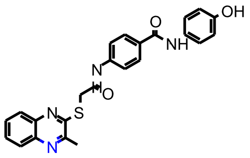
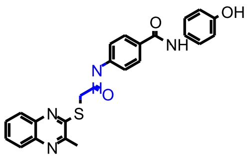
Feature Contribution

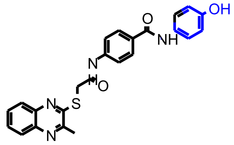
Top features for positive contribution

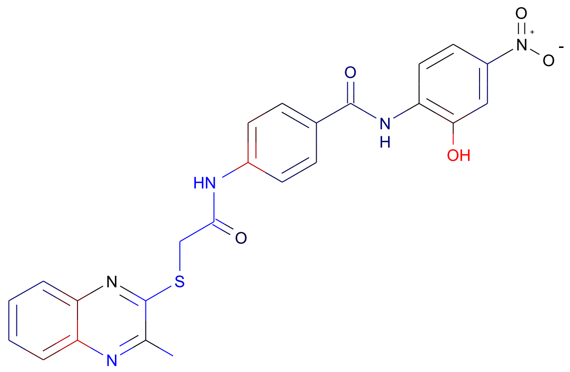
Fingerprint	Bit/Smiles	Feature Structure	Score

ECFP_6	642810091	 [*][c](:[*]):[*]	0.281
ECFP_6	-1897341097	 [*]N[*]	0.216
ECFP_6	1444581947	 [*]C(=[*])[c]1:[cH]:[*]:[c]([*]):[cH]:[cH]:1	0.163

Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
ECFP_6	655739385	 [*]:n:[*]	-0.239
FCFP_6	566058135	 [*]CC(=O)N[*]	-0.216

FCFP_6	946589555	 [*][c]1:[*]:[cH]:[c](O):[cH]:[cH]:1	-0.204



$C_{24}H_{19}N_5O_5S$

Molecular Weight: 489.50316

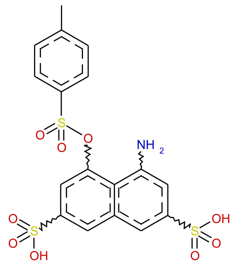
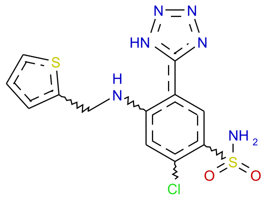
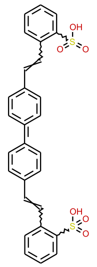
ALogP: 3.731

Rotatable Bonds: 7

Acceptors: 8

Donors: 3

Structural Similar Compounds

Name	2;7-NAPHTHALENE DISULFONIC ACID; 4-AMINO-5-HYDROXY-; p-TOLUENE SULFONATE ESTER	AZOSEMIDE	BENZENESULFONIC ACID; 2;2'-(4;4'-BIPHENYLYLENEDIVINYLENE)DI-; DISODIUM SALT (Na STRIPPED)
Structure			
Actual Endpoint (-log C)	1.615	2.163	1.968
Predicted Endpoint (-log C)	1.79606	2.21052	1.72109
Distance	0.840	0.847	0.865
Reference	85JCAE -;1063;86	IYKEDH 18;666;87	MVCRB3 2;193;73

Model Prediction

Prediction: 6.811

Unit: g/kg_body_weight

Mahalanobis Distance: 23.298

Mahalanobis Distance p-value: 3.53e-022

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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

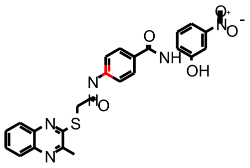
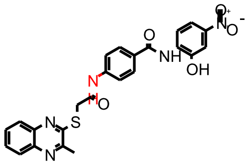
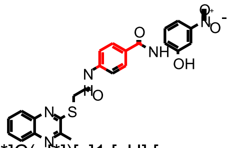
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_2 feature: 1043790491: [*][N+](=[*])[*]
3. Unknown ECFP_2 feature: 781519895: [*][O-]
4. Unknown ECFP_2 feature: -179073144: [*][N+](=[*])[c](:c:[*]):c:[*]
5. Unknown ECFP_2 feature: -215026467: [*]:[c](:[*])[N+](=O)[O-]
6. Unknown ECFP_2 feature: 2104376220: [*][N+](=O)[*]
7. Unknown ECFP_2 feature: -659271057: [*][N+](=[*])[O-]
8. Unknown FCFP_6 feature: 16: [*][c](:[*]):[*]
9. Unknown FCFP_6 feature: 8: [*][N+](=[*])[*]
10. Unknown FCFP_6 feature: 5: [*][O-]
11. Unknown FCFP_6 feature: 1618154665: [*][c](:[*]):[cH]:[cH]:[*]
12. Unknown FCFP_6 feature: 1747237384: [*][c](:[*]):n:[c](:[*]):[*]
13. Unknown FCFP_6 feature: -1410049896: [*]S[c](:n:[*]):[c]([*]):[*]
14. Unknown FCFP_6 feature: 74595001: [*]:[cH]:[c](O):[cH]:[*]
15. Unknown FCFP_6 feature: -828984032: [*][N+](=[*])[c](:[cH]:[*]):[cH]:[*]
16. Unknown FCFP_6 feature: -549108873: [*]:[c](:[*])O
17. Unknown FCFP_6 feature: -1338588315: [*]:[c](:[*])[N+](=O)[O-]
18. Unknown FCFP_6 feature: 1872392852: [*][N+](=O)[*]

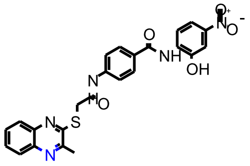
19. Unknown FCFP_6 feature: 260476081: [*][N+](=[*])[O-]

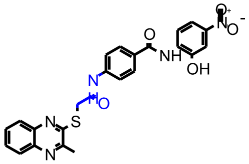
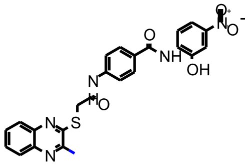
Feature Contribution

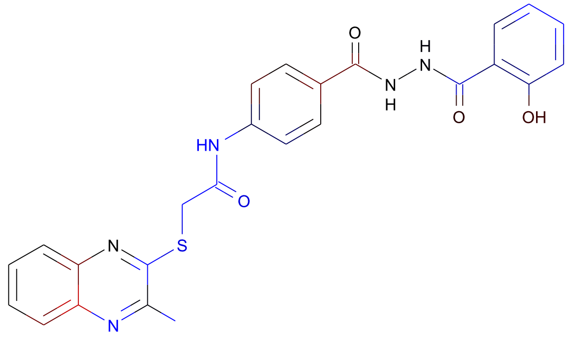
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	642810091	 [*][c](:[*]):[*]	0.281
ECFP_6	-1897341097	 [*]N[*]	0.216
ECFP_6	1444581947	 [*]C(=[*])[c]1:[cH]:[*]:[c]([*]):[cH]:[cH]:1	0.163

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	655739385	 [*]:n:[*]	-0.239

FCFP_6	566058135	 [*]CC(=O)N[*]	-0.216
ECFP_6	734603939	 [*]C	-0.201



$C_{25}H_{21}N_5O_4S$

Molecular Weight: 487.53034

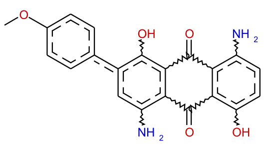
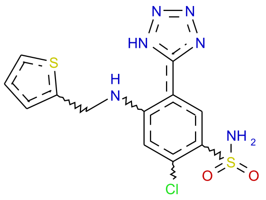
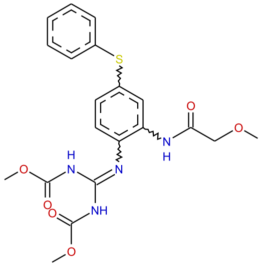
ALogP: 3.093

Rotatable Bonds: 7

Acceptors: 7

Donors: 4

Structural Similar Compounds

Name	**ANTHRAQUINONE; 1;5-DIAMINO-4;8-DIHYDROXY-3-(p-METHOXYPHENYL)-**	**AZOSEMIDE**	**FEBANTEL**
Structure			
Actual Endpoint (-log C)	1.771	2.163	1.624
Predicted Endpoint (-log C)	2.1122	2.21052	2.37098
Distance	0.800	0.818	0.890
Reference	28ZPAK -;245;72	IYKEDH 18;666;87	ARZNAD 28;2193;78

Model Prediction

Prediction: 4.703

Unit: g/kg_body_weight

Mahalanobis Distance: 23.110

Mahalanobis Distance p-value: 2.84e-021

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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

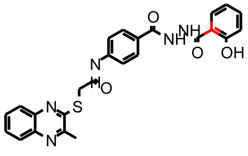
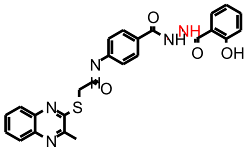
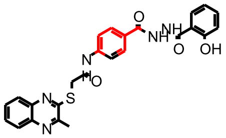
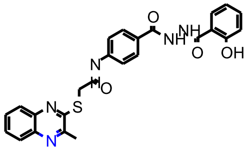
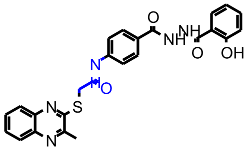
Unknown features are fingerprint features in the query molecule, but not found in the training set.

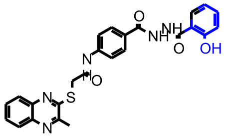
1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_6 feature: 16: [*][c](:[*]):[*]
3. Unknown FCFP_6 feature: -885461129: [*]NNC(=[*])[*]
4. Unknown FCFP_6 feature: 74595001: [*]:[cH]:[c](O):[cH]:[*]
5. Unknown FCFP_6 feature: 1618154665: [*][c](:[*]):[cH]:[cH]:[*]
6. Unknown FCFP_6 feature: -549108873: [*]:[c](:[*])O
7. Unknown FCFP_6 feature: 1747237384: [*][c](:[*]):n:[c](:[*]):[*]
8. Unknown FCFP_6 feature: -1410049896: [*]S[c](:n:[*]):[c]([*]):[*]

Feature Contribution

Top features for positive contribution

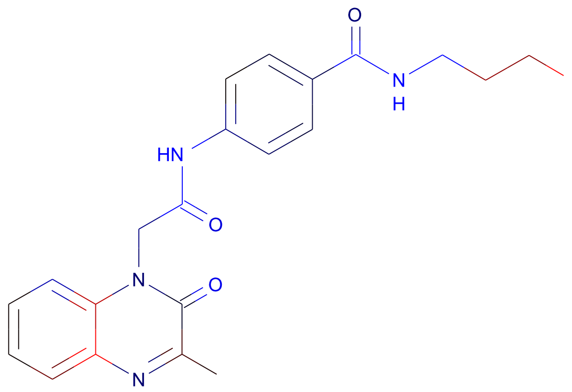
Fingerprint	Bit/Smiles	Feature Structure	Score

ECFP_6	642810091	 [*][c](:[*]):[*]	0.281
ECFP_6	-1897341097	 [*]N[*]	0.216
ECFP_6	1444581947	 [*]C(=[*])[c]1:[cH]:[*]:[c]([*]):[cH]:[cH]:1	0.163
Top Features for negative contribution			
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
ECFP_6	655739385	 [*]:n:[*]	-0.239
FCFP_6	566058135	 [*]CC(=O)N[*]	-0.216

FCFP_6	946589555	 [*][c]1:[*]:[cH]:[c](O):[cH]:[cH]:1	-0.204

17a

TOPKAT_Rat_Oral_LD50



$C_{22}H_{24}N_4O_3$

Molecular Weight: 392.45096

ALogP: 2.212

Rotatable Bonds: 7

Acceptors: 4

Donors: 2

Model Prediction

Prediction: 18.136

Unit: g/kg_body_weight

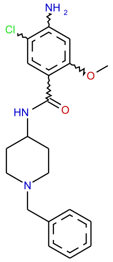
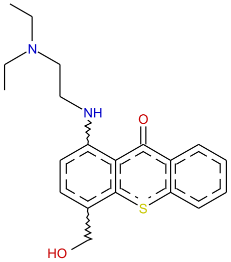
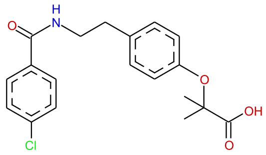
Mahalanobis Distance: 23.642

Mahalanobis Distance p-value: 7.05e-024

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	CLEBOPRIDE; MALATE SALT (MALATE STRIPPED)	HYCANTHONE	BEZAFIBRATE
Structure			
Actual Endpoint (-log C)	2.168	2.561	1.946
Predicted Endpoint (-log C)	2.89926	2.91901	2.54395
Distance	0.534	0.547	0.553
Reference	OYYAA2 25;803;83	EJBLAB 1;181;74	ARZNAD 30;2023;80

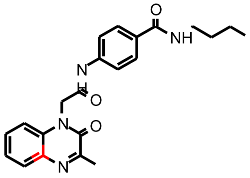
Model Applicability

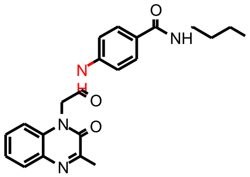
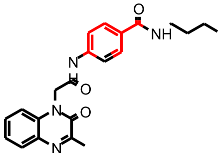
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_6 feature: 16: [*][c](:[*]):[*]
3. Unknown FCFP_6 feature: 1618154665: [*][c](:[*]):[cH]:[cH]:[*]
4. Unknown FCFP_6 feature: 580453787: [*]C(=N[c](:[*]):[*])[*]

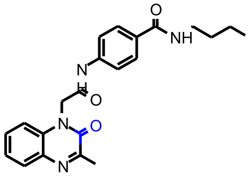
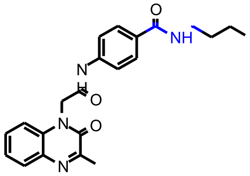
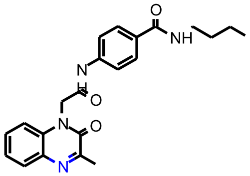
Feature Contribution

Top features for positive contribution

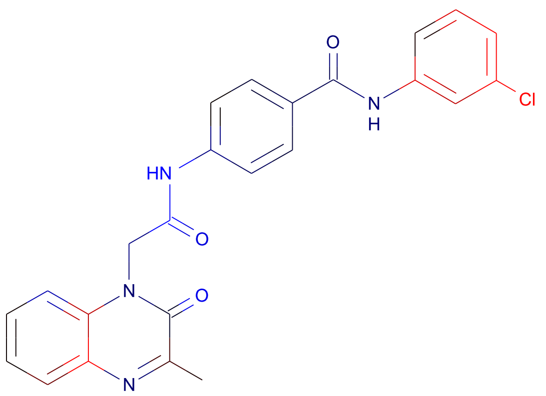
Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	642810091	 [*][c](:[*]):[*]	0.281

ECFP_6	-1897341097	 [*]N[*]	0.216
ECFP_6	1444581947	 [*]C(=[*])[c]1:[cH]:[*]:[c]([*]):[cH]:[cH]:1	0.163

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	2106656448	 [*]C(=O)[*]	-0.352
ECFP_6	497523368	 [*]CNC(=[*])[*]	-0.301
ECFP_6	655739385	 [*]:n:[*]	-0.239

17b



$C_{24}H_{19}ClN_4O_3$

Molecular Weight: 446.88566

ALogP: 3.125

Rotatable Bonds: 5

Acceptors: 4

Donors: 2

Model Prediction

Prediction: 7.450

Unit: g/kg_body_weight

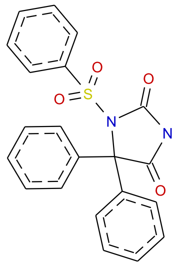
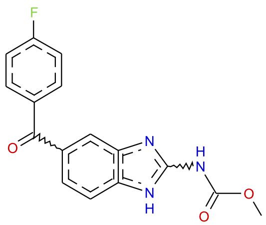
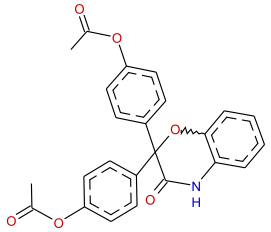
Mahalanobis Distance: 22.536

Mahalanobis Distance p-value: 1.29e-018

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	**1-BENZENESULFONYL-5;5-DIPHENYLHYDANTOIN**	**FLUBENDAZOLE**	**bis-OXATIN ACETATE**
Structure			
Actual Endpoint (-log C)	2.363	2.088	1.717
Predicted Endpoint (-log C)	2.34793	2.69288	2.40947
Distance	0.620	0.630	0.636
Reference	ARZNAD 20;1579;70	YRTMA6 9;11;78	NIIRDN 6;609;82

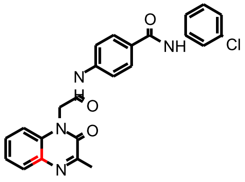
Model Applicability

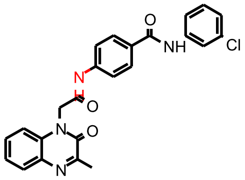
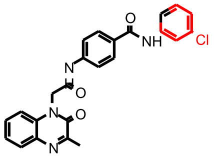
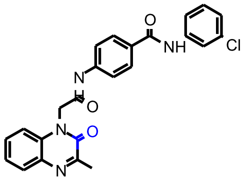
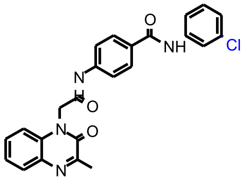
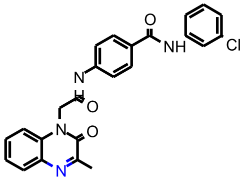
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_6 feature: 16: [*][c](:[*]):[*]
3. Unknown FCFP_6 feature: 1618154665: [*][c](:[*]):[cH]:[cH]:[*]
4. Unknown FCFP_6 feature: 580453787: [*]C(=N[c](:[*]):[*])[*]
5. Unknown FCFP_6 feature: 71476542: [*]:[c](:[*])Cl

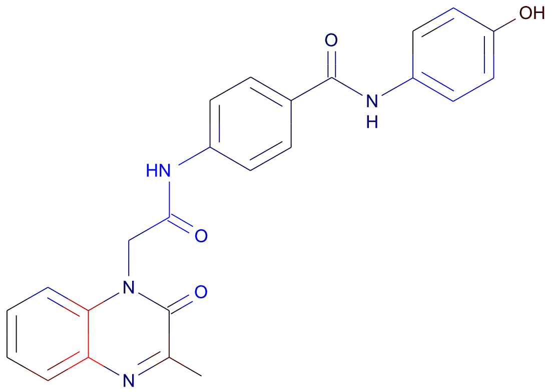
Feature Contribution

Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	642810091	 [*][c](:[*]):[*]	0.281

ECFP_6	-1897341097	 [*]N[*]	0.216
ECFP_6	577592657	 [*][c]1:[*]:[cH]:[cH]:[c](Cl):[cH]:1	0.194
Top Features for negative contribution			
Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
ECFP_6	2106656448	 [*]C(=O)[*]	-0.352
ECFP_6	-817402818	 [*]Cl	-0.263
ECFP_6	655739385	 [*]:n:[*]	-0.239

17c



$C_{24}H_{20}N_4O_4$

Molecular Weight: 428.44

ALogP: 2.218

Rotatable Bonds: 5

Acceptors: 5

Donors: 3

Model Prediction

Prediction: 7.868

Unit: g/kg_body_weight

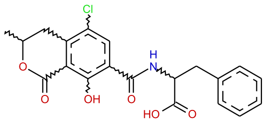
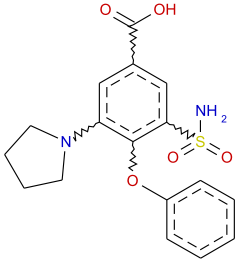
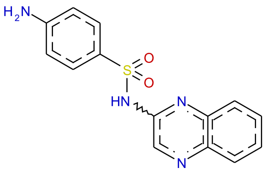
Mahalanobis Distance: 22.461

Mahalanobis Distance p-value: 2.77e-018

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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	OCHRATOXIN A	PIRETANIDE	SULFAQUINOXALINE
Structure			
Actual Endpoint (-log C)	4.305	1.811	2.341
Predicted Endpoint (-log C)	3.03558	1.83976	2.42674
Distance	0.639	0.714	0.725
Reference	FCTXAV 6;479;68	DRFUD4 2;393;77	MahWM# 16NOV82

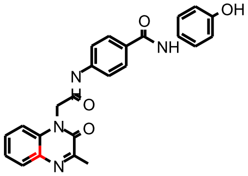
Model Applicability

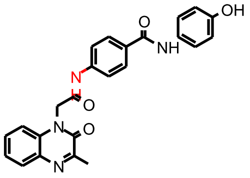
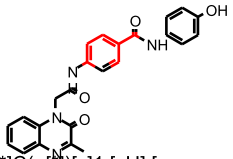
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_6 feature: 16: [*][c](:[*]):[*]
3. Unknown FCFP_6 feature: 1618154665: [*][c](:[*]):[cH]:[cH]:[*]
4. Unknown FCFP_6 feature: 580453787: [*]C(=N[c](:[*]):[*])[*]
5. Unknown FCFP_6 feature: 74595001: [*]:[cH]:[c](O):[cH]:[*]
6. Unknown FCFP_6 feature: -549108873: [*]:[c](:[*])O

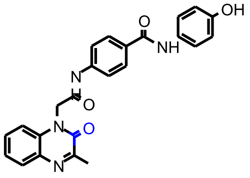
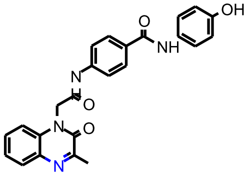
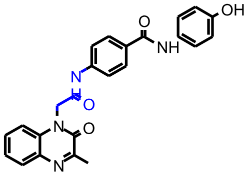
Feature Contribution

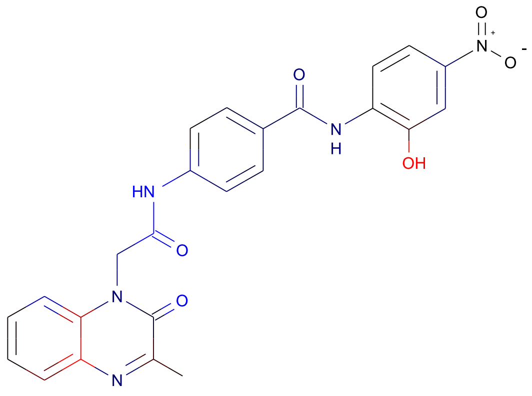
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	642810091	 [*][c](:[*]):[*]	0.281

ECFP_6	-1897341097	 [*]N[*]	0.216
ECFP_6	1444581947	 [*]C(=[*])[c]1:[cH]:[*]:[c]([*]):[cH]:[cH]:1	0.163

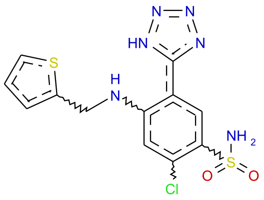
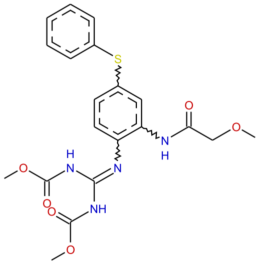
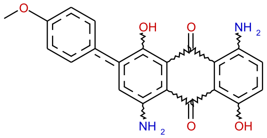
Top Features for negative contribution

Fingerprint	**Bit/Smiles**	**Feature Structure**	**Score**
ECFP_6	2106656448	 [*]C(=O)[*]	-0.352
ECFP_6	655739385	 [*]:n:[*]	-0.239
FCFP_6	566058135	 [*]CC(=O)N[*]	-0.216



$C_{24}H_{19}N_5O_6$
Molecular Weight: 473.43756
ALogP: 2.113
Rotatable Bonds: 6
Acceptors: 7
Donors: 3

Structural Similar Compounds

Name	AZOSEMIDE	FEBANTEL	ANTHRAQUINONE; 1;5-DIAMINO-4;8-DIHYDROXY-3-(p-METHOXYPHENYL)-
Structure			
Actual Endpoint (-log C)	2.163	1.624	1.771
Predicted Endpoint (-log C)	2.21052	2.37098	2.1122
Distance	0.650	0.730	0.732
Reference	IYKEDH 18;666;87	ARZNAD 28;2193;78	28ZPAK -;245;72

Model Prediction

Prediction: 5.470
Unit: g/kg_body_weight
Mahalanobis Distance: 26.984
Mahalanobis Distance p-value: 4.2e-043

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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.

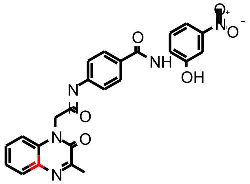
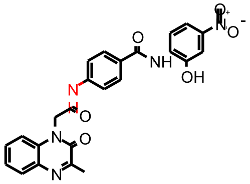
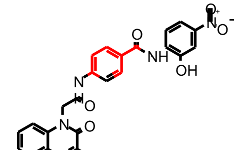
Model Applicability

Unknown features are fingerprint features in the query molecule, but not found in the training set.

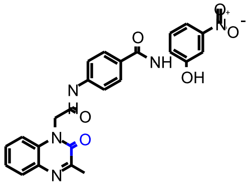
1. All properties and OPS components are within expected ranges.
2. Unknown ECFP_2 feature: 1043790491: [*][N+](=[*])[*]
3. Unknown ECFP_2 feature: 781519895: [*][O-]
4. Unknown ECFP_2 feature: -179073144: [*][N+](=[*])[c](:c:[*]):c:[*]
5. Unknown ECFP_2 feature: -215026467: [*]:[c](:[*])[N+](=O)[O-]
6. Unknown ECFP_2 feature: 2104376220: [*][N+](=O)[*]
7. Unknown ECFP_2 feature: -659271057: [*][N+](=[*])[O-]
8. Unknown FCFP_6 feature: 16: [*][c](:[*]):[*]
9. Unknown FCFP_6 feature: 8: [*][N+](=[*])[*]
10. Unknown FCFP_6 feature: 5: [*][O-]
11. Unknown FCFP_6 feature: 1618154665: [*][c](:[*]):[cH]:[cH]:[*]
12. Unknown FCFP_6 feature: 580453787: [*]C(=N[c](:[*]):[*])[*]
13. Unknown FCFP_6 feature: 74595001: [*]:[cH]:[c](O):[cH]:[*]
14. Unknown FCFP_6 feature: -828984032: [*][N+](=[*])[c](:[cH]:[*]):[cH]:[*]
15. Unknown FCFP_6 feature: -549108873: [*]:[c](:[*])O
16. Unknown FCFP_6 feature: -1338588315: [*]:[c](:[*])[N+](=O)[O-]
17. Unknown FCFP_6 feature: 1872392852: [*][N+](=O)[*]
18. Unknown FCFP_6 feature: 260476081: [*][N+](=[*])[O-]

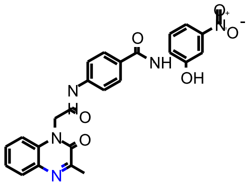
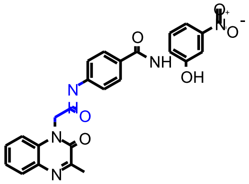
Feature Contribution

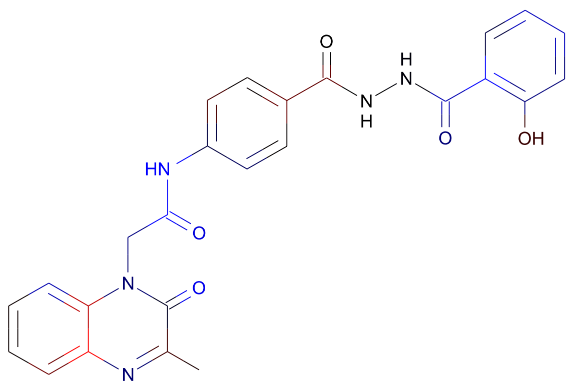
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	642810091	 [*][c](:[*]):[*]	0.281
ECFP_6	-1897341097	 [*]N[*]	0.216
ECFP_6	1444581947	 [*]C(=[*])[c]1:[cH]:[*]:[c]([*]):[cH]:[cH]:1	0.163

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	2106656448	 [*]C(=O)[*]	-0.352

ECFP_6	655739385	 [*]:n:[*]	-0.239
FCFP_6	566058135	 [*]CC(=O)N[*]	-0.216



$C_{25}H_{21}N_5O_5$

Molecular Weight: 471.46474

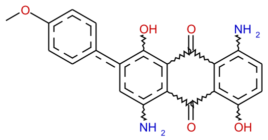
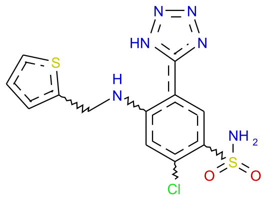
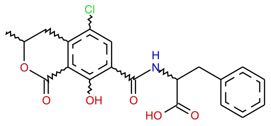
ALogP: 1.475

Rotatable Bonds: 6

Acceptors: 6

Donors: 4

Structural Similar Compounds

Name	**ANTHRAQUINONE; 1;5-DIAMINO-4;8-DIHYDROXY-3-(p-METHOXYPHENYL)-**	**AZOSEMIDE**	**OCHRATOXIN A**
Structure			
Actual Endpoint (-log C)	1.771	2.163	4.305
Predicted Endpoint (-log C)	2.1122	2.21052	3.03558
Distance	0.707	0.730	0.748
Reference	28ZPAK -;245;72	IYKEDH 18;666;87	FCTXAV 6;479;68

Model Prediction

Prediction: 5.956

Unit: g/kg_body_weight

Mahalanobis Distance: 24.187

Mahalanobis Distance p-value: 1.11e-026

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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

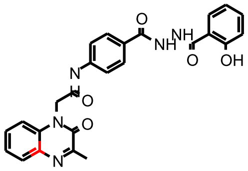
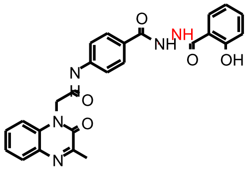
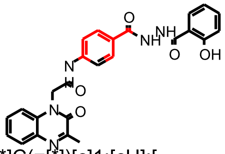
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_6 feature: 16: [*][c](:[*]):[*]
3. Unknown FCFP_6 feature: 1618154665: [*][c](:[*]):[cH]:[cH]:[*]
4. Unknown FCFP_6 feature: 580453787: [*]C(=N[c](:[*]):[*])[*]
5. Unknown FCFP_6 feature: -885461129: [*]NNC(=[*])[*]
6. Unknown FCFP_6 feature: 74595001: [*]:[cH]:[c](O):[cH]:[*]
7. Unknown FCFP_6 feature: -549108873: [*]:[c](:[*])O

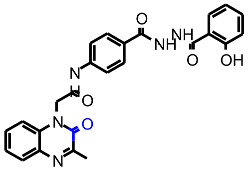
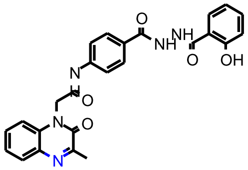
Feature Contribution

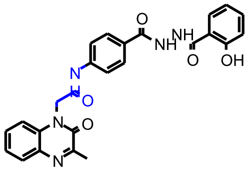
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score

ECFP_6	642810091	 [*][c](:[*]):[*]	0.281
ECFP_6	-1897341097	 [*]N[*]	0.216
ECFP_6	1444581947	 [*]C(=[*])[c]1:[cH]:[*]:[c]([*]):[cH]:[cH]:1	0.163

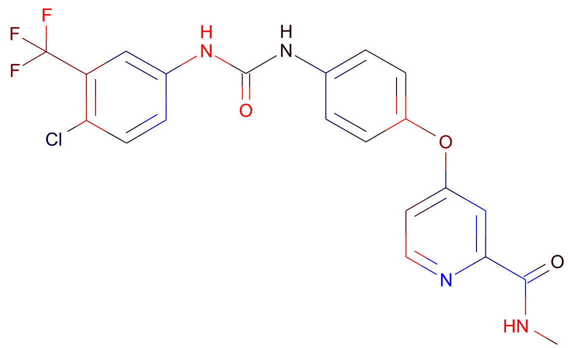
Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	2106656448	 [*]C(=O)[*]	-0.352
ECFP_6	655739385	 [*]:n:[*]	-0.239

FCFP_6	566058135	 [*]CC(=O)N[*]	-0.216

Sorafenib

TOPKAT_Rat_Oral_LD50



$C_{21}H_{16}ClF_3N_4O_3$

Molecular Weight: 464.82494

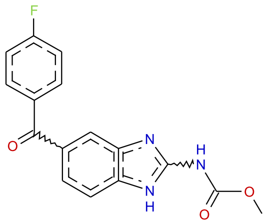
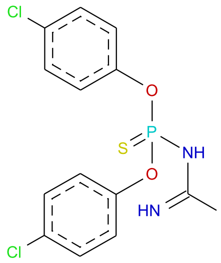
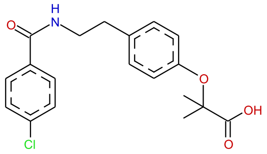
ALogP: 4.175

Rotatable Bonds: 6

Acceptors: 4

Donors: 3

Structural Similar Compounds

Name	FLUBENDAZOLE	PHOSPHORAMIDOTHIOIC ACID; ACETIMIDOYL-; O;O-bis-(p-CHLOROPHENYL)ESTER	BEZAFIBRATE
Structure			
Actual Endpoint (-log C)	2.088	5.006	1.946
Predicted Endpoint (-log C)	2.69288	3.23989	2.54395
Distance	0.697	0.703	0.721
Reference	YRTMA6 9;11;78	FMCHA2 -;C149;89	ARZNAD 30;2023;80

Model Prediction

Prediction: 0.823

Unit: g/kg_body_weight

Mahalanobis Distance: 21.029

Mahalanobis Distance p-value: 1.93e-012

Mahalanobis Distance: The Mahalanobis distance (MD) is a generalization of the Euclidean distance that accounts for correlations among the X properties. It is calculated as 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

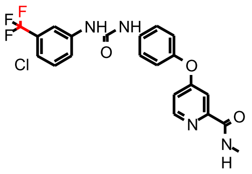
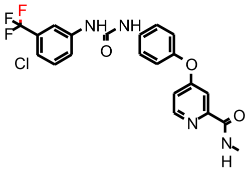
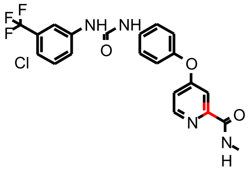
Unknown features are fingerprint features in the query molecule, but not found in the training set.

1. All properties and OPS components are within expected ranges.
2. Unknown FCFP_6 feature: 16: [*][c](:[*]):[*]
3. Unknown FCFP_6 feature: 1618154665: [*][c](:[*]):[cH]:[cH]:[*]
4. Unknown FCFP_6 feature: 1747237384: [*][c](:[*]):n:[c](:[*]):[*]
5. Unknown FCFP_6 feature: 71476542: [*]:[c](:[*])Cl
6. Unknown FCFP_6 feature: 136686699: [*]NC

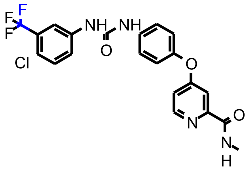
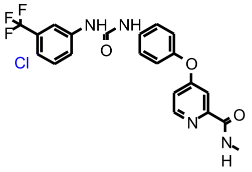
Feature Contribution

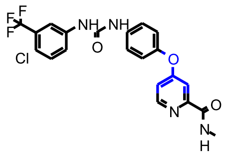
Top features for positive contribution

Fingerprint	Bit/Smiles	Feature Structure	Score

FCFP_6	71953198	 [*]C([*])([*])F	0.392
ECFP_6	-1046436026	 [*]F	0.349
ECFP_6	642810091	 [*][c](:[*]):[*]	0.281

Top Features for negative contribution

Fingerprint	Bit/Smiles	Feature Structure	Score
ECFP_6	226796801	 [*]C([*])([*])F	-0.320
ECFP_6	-817402818	 [*]Cl	-0.263

ECFP_6	-176455838	 [*]O[c](:[cH]:[*]):[cH]:[*]	-0.257