

201329011A

厚生労働科学研究費補助金

化学物質リスク研究事業

QSARによる化学物質の有害性予測の
迅速化・高度化に関する研究

平成25年度 総括・分担研究報告書

(H25-化学-一般-001)

研究代表者 栗原 正明

平成26(2014)年 3月

目 次

I. 総括研究報告

QSAR による化学物質の有害性予測の迅速化・高度化に関する研究 栗原 正明	1
---	---

II. 分担研究報告書

1. 構造類似学習母集団を用いたQSARモデルの作成 栗原 正明	5
2. 化学物質の構造の精密化・高度化 出水 庸介	35

厚生労働科学研究費補助金（化学物質リスク研究事業）
総括研究報告書

QSAR による化学物質の有害性予測の迅速化・高度化に関する研究

研究代表者：栗原正明 国立医薬品食品衛生研究所有機化学部 部長

研究要旨

本年度は化学物質の構造と毒性（Ames法）の定量的構造活性相関（QSAR）について、新しい方法論の開発に重点を置き、研究事業を遂行した。特に、化学物質の構造類似性に基づいたプレカテゴライゼーション法によるQSARモデルの構築を行った。

研究分担者

出水 庸介 国立医薬品食品衛生研究所
有機化学部 第二室長

A. 研究目的

化学物質のリスクから国民生活の安全を確保する観点から、化学物質の毒性に関する情報の取得が喫緊の課題である。しかし、動物を用いる安全性試験は莫大な時間と費用がかかるため、毒性が未知のすべての化学物質について動物試験により毒性を評価することは不可能である。また、動物愛護の観点からも動物試験を減らすことは重要である。そこで、構造活性相関、特にコンピュータを利用した定量的構造活性相関（QSAR）による毒性予測が化学物質管理の観点から非常に重要である。

本研究では、透明性を確保した新しい毒性予測システムを開発することを目的とする。

B. 研究方法

構造類似学習母集団を用いた QSAR モデル

の作成

【データの整理】

ISSTOX Chemical Toxicity Database（総数 6865）から以下のものを除外した。

（1）Ames 試験結果が不明確な化合物 1023 件

（2）Model Builder 未対応化合物（混合物や金属系化合物）735 件

（3）同一化合物に対して、異なる Ames 試験結果が掲載されている化合物 22 件
その結果、5067 件を有効母集団とした。

（巻末の参考資料）

2012 年度に作成した 3 種類の予測モデル（LX-C=O モデル、LA-NH₂ モデル、LA-BZN モデル）で誤分類された化合物（62 件）から、任意に 10 件（陽性：5 件、陰性 5 件）を選定した。

【類似学習母集団の選定】

検証化合物に対して類似度 95%、90%及び 85%の類似学習母集団を有効母集団（5067 件）から抽出した。追加検証母集団に対しては類似度 90%の類似学習母集団のみ抽出した。

【類似学習母集団からのモデル作成】

上記の方法で抽出した各検証化合物の類似学習母集団から以下の通りモデルを作成した。

(1) 類似学習母集団の確認

各検証化合物の類似学習母集団について、母集団の件数および陽性化合物と陰性化合物の比率を確認した。母集団の件数が45件以上かつ陽性化合物と陰性化合物の比率が0.5以上の母集団について、モデル式を作成した。それ以外の母集団については、モデル作成困難と判断した。

(2) パラメータ発生

上記(1)の条件を満たした類似学習母集団について、ADMEWORKS/ModelBuilderを用いて、物理化学的パラメータ(555個)と部分構造関連パラメータ(213個)、合計768個のパラメータを発生した。

(3) サンプル抽出

類似学習母集団について、パラメータの欠損値が多く、解析に向かない化合物を確認した結果、学習母集団から除外する必要な化合物はなかった。

(4) 特徴抽出

(2)で発生した768個のパラメータから欠損値、ゼロ値、高い相関のパラメータを除外し、線形重回帰にて有効なパラメータ(|t-statistics|>2.0かつp-value<0.05)に絞った。最終的にパラメータを47個に絞った。

(5) モデル作成

類似学習母集団を用いて、SGP(Stochastic Gradient Perceptron)の線形学習手法によりモデルを作成し、該当する選定化合物を予測した。

化学物質の構造の精密化・高度化

本年度は昨年度までの研究事業で行った

部分構造によるプレカテゴリーゼーションを用いた毒性予測において、全ての予測法で誤分類された62件の化合物を、予測したい化合物(ターゲット分子)とした。このターゲット分子の構造の精密化、高度化に対応するため、エネルギー安定構造である三次元構造を求めた。

C. 研究結果

検証化合物に対して、95%、90%、85%の類似学習母集団にて作成した予測モデルを用いて毒性予測を行った。

追加検証化合物に対して、90%の類似学習母集団にて作成した予測モデルをもちいて毒性予測を行った。

2012年度に行った全ての方法で誤分類された化合物から選定された化合物のうち半分程度は予測することができた。

本プロジェクトで用いた、誤分類された62件の化合物について最安定構造(最安定コンフォマー)を計算した。

D. 考察

今回作成した予測モデルにて誤分類された検証化合物では学習母集団の中で最も類似している化合物の実測値が検証化合物の実測値と異なっている。これは、構造活性相関という手法に基づき作成している予測モデルのため、検証化合物の予測結果が最も類似している学習母集団の実測値の影響を大きく受けたため、誤分類されたと考えられる。一方、検証化合物に対して最も類似している母集団とその実測値が一致しているにも関わらず、誤分類された検証化合物が混在していないかどうかを確認する必要があると考えられる。

既存の毒性データベースにおける化合物の構造は不備が多い。2次元構造においてすら、特に構造異性体に関しては不明なも

のが多い。

E. 結論

類似性の高い構造類似集団を学習母集団として QSAR 法によりターゲット化合物の毒性を予測することは有効な方法論である。

本プロジェクトで用いた、誤分類された 62 件の化合物について最安定構造最安定構造を計算した。今後、分子の 3 次元構造を用いる構造活性相関の解析に使用できるようになった。

F. 健康危機情報

特に無し。

G. 研究発表

1. 論文発表

(論文発表-参考：コンピュータシミュレーションを含むもの)

栗原正明

コンピュータシミュレーションによる違法ドラッグの活性予測
YAKUGAKU ZASSHI, 133, 13-16 (2013)

Demizu, Y., Nagoya, S., Shirakawa, M., Kawamura, M., Yamagata, N., Sato, Y., Doi, M., Kurihara, M.; Development of stapled short helical peptides capable of inhibiting vitamin D receptor (VDR)-coactivator interactions; *Bioorg. Med. Chem. Lett.*, 23, 4292-4296(2013)

N. Sakakibara, T. Hamasaki, M. Baba, Y. Demizu, M. Kurihara, K. Irie, M. Iwai, E. Asada, Y. Kato, T. Maruyama; Synthesis and evaluation of novel 3-(3,5-dimethylbenzyl)-uracil analogs as potential anti-HIV-1 agents *Bioorg. Med. Chem.*, 21, 5900-5906(2013)

I. Kato, M. Oba, M. Kurihara, Y. Takano, M. Tanaka; Synthesis of Cyclic α,α -Disubstituted Amino Acid Bearing a Pendent Chiral Center; *Peptide Science* 2012,

129-130 (2013)

A. Imanishi, M. Oba, Y. Demizu, M. Kurihara, M. Doi, H. Takazaki, H. Suemune, M. Tanaka; Synthesis of Chiral Five-membered Ring Amino Acids with an Azido Group, and Their Peptides; *Peptide Science* 2012, 131-132 (2013)

T. Sugiyama, Y. Imamura, Y. Demizu, M. Kurihara, M. Takano, A. Kittaka; Synthesis of b-Chiral Peptide Nucleic Acids Bearing Lysine Side Chains; *Peptide Science* 2012, 385-386 (2013)

N. Yamazaki, Y. Demizu, Y. Sato, M. Doi, M. Kurihara; Helical Foldamer Containing a Combination of Cyclopentane-1,2-diamine and 2,2-Dimethylmalonic Acid; *J. Org. Chem.*, 2013, 78, pp 9991-9994

M. Oba, N. Ishikawa, Y. Demizu, M. Kurihara, H. Suemune, M. Tanaka; Helical Oligomer with Changeable Chiral Acetal Moiety; *Eur. J. Org. Chem.* 2013, 7679-7682

Y. Demizu, H. Yamashita, N. Yamazaki, Y. Sato, M. Doi, M. Tanaka, M. Kurihara; Oligopeptides with equal amounts of l- and d-amino acids may prefer a helix screw sense; *J. Org. Chem.*, 2013, 78, 12106-12113

2. 学会発表

栗原正明, 出水庸介, 佐藤由紀子; QSAR による化学物質の有害性予測法の開発 日本薬学会第 133 年会(2013/03/28-30, 横浜)

栗原正明, 佐藤由紀子, 出水庸介; コンピュータシミュレーションによる違法ドラッグの活性予測 日本薬学会第 133 年会(2013/03/28-30, 横浜)

H. 知的財産権の出願・登録状況

なし

厚生労働科学研究費補助金（化学物質リスク研究事業）
分担研究報告書

分担研究課題：構造類似学習母集団を用いた QSAR モデルの作成

研究分担者：栗原正明 国立医薬品食品衛生研究所有機化学部 部長

研究要旨

構造類似学習母集団を用いた QSAR モデルの作成を行った。有効母集団 5067 件の中から、予測したい化合物（ターゲット化合物）と類似性の高い（例：90%）化合物を抽出した。それを構造類似学習母集団として QSAR 法によりターゲット化合物の毒性を予測した。

A. 研究目的

本研究では、透明性を確保した新しい毒性予測システムを開発することを目的とする。

B. 研究方法

データの整理

ISSTOX Chemical Toxicity Database（総数 6865）から以下のものを除外した。

（1）Ames 試験結果が不明確な化合物 1023 件

（2）Model Builder 未対応化合物（混合物や金属系化合物）735 件

（3）同一化合物に対して、異なる Ames 試験結果が掲載されている化合物 22 件

その結果、5067 件を有効母集団とした。

（巻末の参考資料）

2012 年度に作成した 3 種類の予測モデル

（LX-C=O モデル，LA-NH₂ モデル，LA-BZN モデル）で誤分類された化合物

（62 件：表 1）から、任意に 10 件（陽性：5 件，陰性 5 件）を選定した。

検証化合物 10 件を表 2 に示した。

追加検証化合物 10 件を表 3 に示した。

類似学習母集団の選定

図 1 の通り，表 2 の検証化合物に対して類似度 95%，90% 及び 85% の類似学習母集団を有効母集団（5067 件）から抽出した。

表 3 の追加検証母集団に対しては類似度 90% の類似学習母集団のみ抽出した。

類似学習母集団からのモデル作成

上記の方法で抽出した各検証化合物の類似学習母集団から以下の通りモデルを作成した。

（1）類似学習母集団の確認

各検証化合物の類似学習母集団について，母集団の件数および陽性化合物と陰性化合物の比率を確認した。母集団の件数が 45 件以上かつ陽性化合物と陰性化合物の比率が 0.5 以上の母集団について，モデル式を作成した。それ以外の母集団については，

モデル作成困難と判断した。

(2) パラメータ発生

上記(1)の条件を満たした類似学習母集団について、ADMEWORKS/ModelBuilderを用いて、物理化学的パラメータ(555個)と部分構造関連パラメータ(213個)、合計768個のパラメータを発生した。

(3) サンプル抽出

類似学習母集団について、パラメータの欠損値が多く、解析に向かない化合物を確認した結果、学習母集団から除外する必要な化合物はなかった。

(4) 特徴抽出

(2)で発生した768個のパラメータから欠損値、ゼロ値、高い相関のパラメータを除外し、線形重回帰にて有効なパラメータ(|t-statistics|>2.0かつp-value<0.05)に絞った。最終的にパラメータを47個に絞った。

(5) モデル作成

類似学習母集団を用いて、SGP(Stochastic Gradient Perceptron)の線形学習手法によりモデルを作成し、該当する選定化合物を予測した。

C. 研究結果

表2の検証化合物に対して、95%、90%、85%の類似学習母集団にて作成した予測モデルの結果は表4~表6の通り。

表3の追加検証化合物に対して、90%の類似学習母集団にて作成した予測モデルの結果は表7の通りだった。

全ての方法で誤分類された化合物の半分程度は予測することができた。

D. 考察

表4、表5、表6、表7より作成した予測モデルにて誤分類された検証化合物では

学習母集団の中で最も類似している化合物の実測値が検証化合物の実測値と異なっている。追加検証化合物の結果である表8でも、同様の結果となった。

これは、構造活性相関という手法に基づき作成している予測モデルのため、検証化合物の予測結果が最も類似している学習母集団の実測値の影響を大きく受けたため、誤分類されたと考えられる。一方、検証化合物に対して最も類似している母集団とその実測値が一致しているにも関わらず、誤分類された検証化合物が混在していないかどうかを確認する必要があると考えられる。

E. 結論

類似性の高い構造類似集団を学習母集団としてQSAR法によりターゲット化合物の毒性を予測することは有効な方法論である。

F. 健康危機情報

特になし

G. 研究発表

1. 論文発表

(論文発表-参考:コンピュータシミュレーションを含むもの)

栗原正明

コンピュータシミュレーションによる違法ドラッグの活性予測

YAKUGAKU ZASSHI, 133, 13-16 (2013)

Demizu, Y., Nagoya, S., Shirakawa, M., Kawamura, M., Yamagata, N., Sato, Y., Doi, M., Kurihara, M.; Development of stapled short helical peptides capable of inhibiting vitamin D receptor (VDR)-coactivator interactions; *Bioorg. Med. Chem. Lett.*, 23, 4292-4296(2013)

N. Sakakibara, T. Hamasaki, M. Baba, Y.

Demizu, M. Kurihara, K. Irie, M. Iwai, E. Asada, Y. Kato, T. Maruyama; Synthesis and evaluation of novel 3-(3,5-dimethylbenzyl)-uracil analogs as potential anti-HIV-1 agents *Bioorg. Med. Chem.*, 21, 5900-5906(2013)

I. Kato, M. Oba, M. Kurihara, Y. Takano, M. Tanaka; Synthesis of Cyclic α,α -Disubstituted Amino Acid Bearing a Pendent Chiral Center; *Peptide Science 2012*, 129-130 (2013)

A. Imanishi, M. Oba, Y. Demizu, M. Kurihara, M. Doi, H. Takazaki, H. Suemune, M. Tanaka; Synthesis of Chiral Five-membered Ring Amino Acids with an Azido Group, and Their Peptides; *Peptide Science 2012*, 131-132 (2013)

T. Sugiyama, Y. Imamura, Y. Demizu, M. Kurihara, M. Takano, A. Kittaka; Synthesis of b-Chiral Peptide Nucleic Acids Bearing Lysine Side Chains; *Peptide Science 2012*, 385-386 (2013)

N. Yamazaki, Y. Demizu, Y. Sato, M. Doi, M. Kurihara; Helical Foldamer Containing a Combination of Cyclopentane-1,2-diamine and 2,2-Dimethylmalonic Acid; *J. Org. Chem.*, 2013, 78, pp 9991-9994

M. Oba, N. Ishikawa, Y. Demizu, M. Kurihara, H. Suemune, M. Tanaka; Helical Oligomer with Changeable Chiral Acetal Moiety; *Eur. J. Org. Chem.* 2013, 7679-7682

Y. Demizu, H. Yamashita, N. Yamazaki, Y. Sato, M. Doi, M. Tanaka, M. Kurihara; Oligopeptides with equal amounts of l- and d-amino acids may prefer a helix screw sense; *J. Org. Chem.*, 2013, 78, 12106-12113

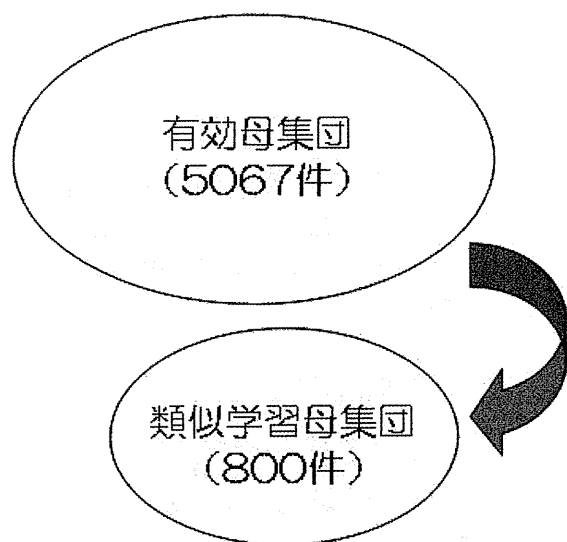
2. 学会発表

栗原正明, 出水庸介, 佐藤由紀子; QSARによる化学物質の有害性予測法の開発
日本薬学会第133年会(2013/03/28-30, 横浜)

栗原正明, 佐藤由紀子, 出水庸介; コンピュータシミュレーションによる違法ドラッグの活性予測
日本薬学会第133年会(2013/03/28-30, 横浜)

H. 知的財産権の出願・登録状況

なし



ADMEWORKS/ModelBuilderの類似検索機能を用いて、指定した類似度で類似構造のものを抽出した。
(例: 選定化合物m0078との類似度が90%以上)

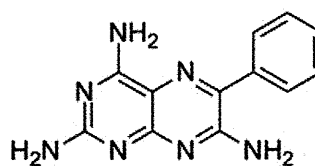
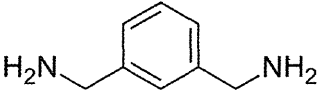
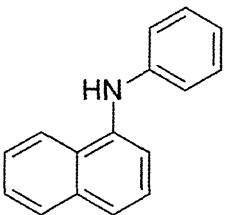
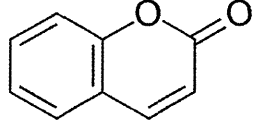
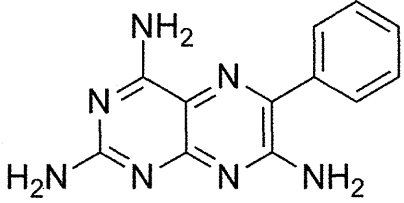
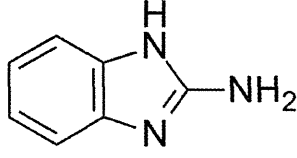
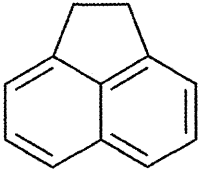
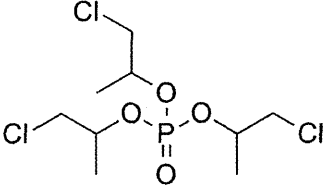
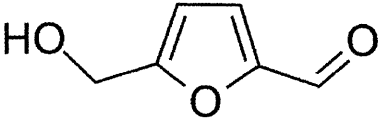
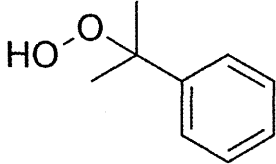
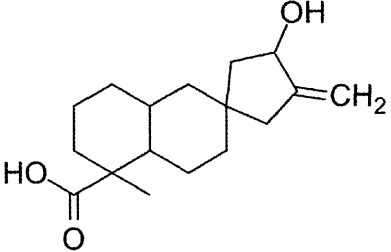
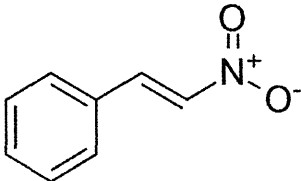


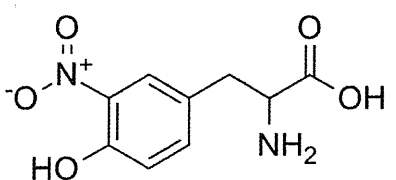
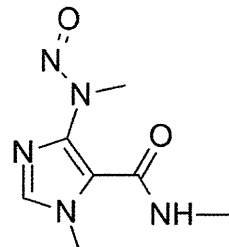
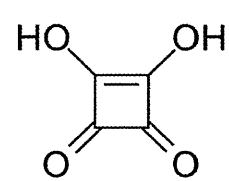
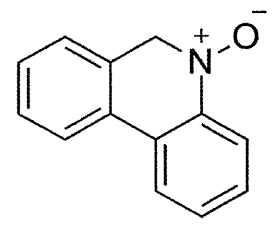
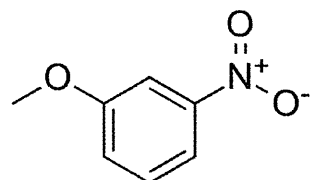
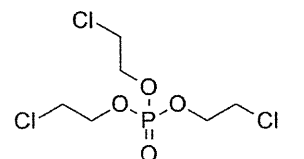
図 1

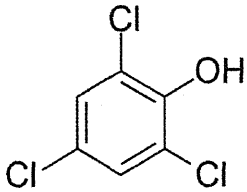
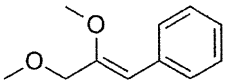
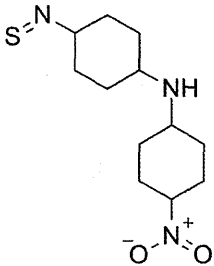
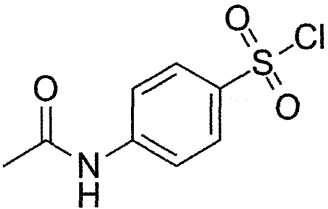
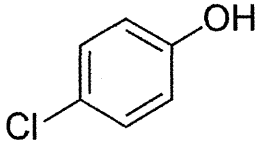
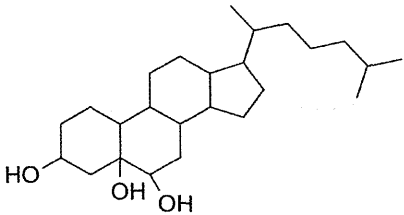
表 1

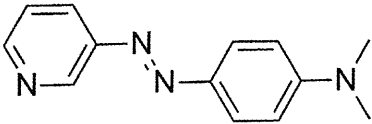
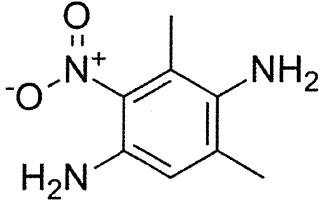
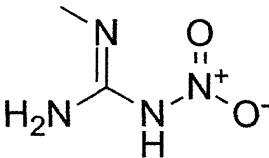
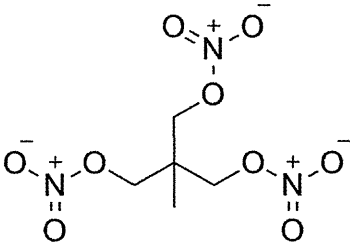
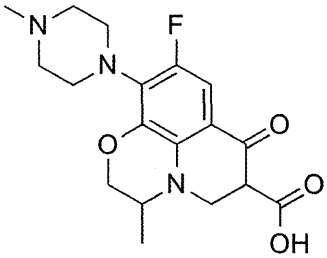
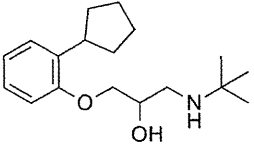
今までの予測モデルで必ず誤分類される化合物

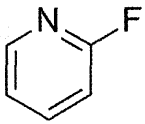
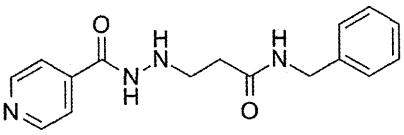
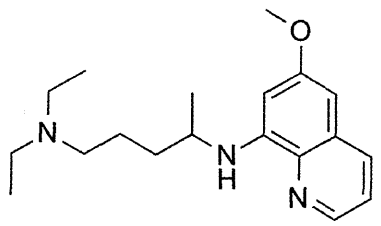
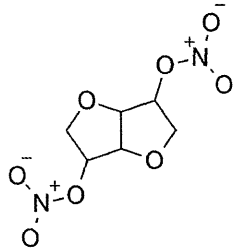
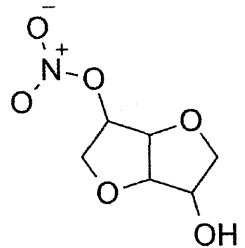
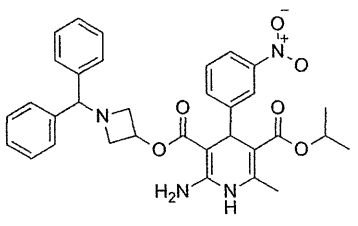
	Structure	ID	Name	AMES
1		m0034	m-xylylenediamine	-
2		m0035	N-phenyl-1-naphthylamine	-
3		m0065	coumarin	+
4		m0078	triamterene	-
5		m0081	2-aminobenzimidazole	-

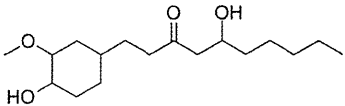
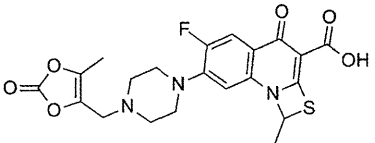
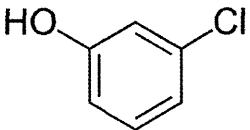
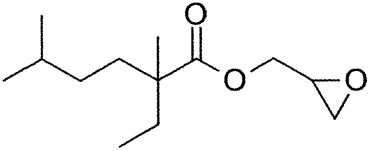
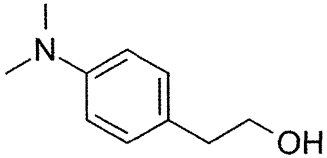
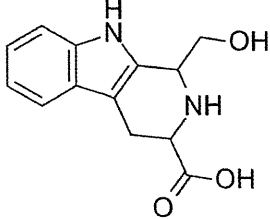
6		m0099	acenaphthene	
7		m0102	tris(1-chloro-2-propyl)phosphate	
8		m0121	5-hydroxymethylfurfural	
9		m0123	cumenehydroperoxide	+
10		m0156	steviol	+
11		m0164	beta-nitrostyrene	

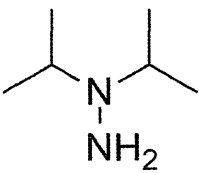
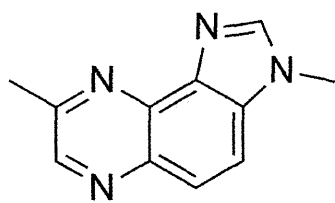
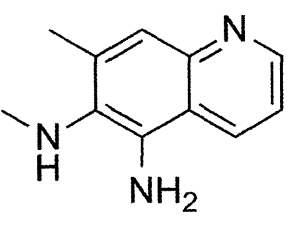
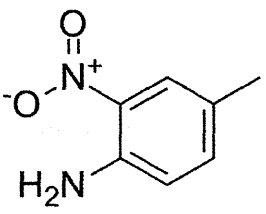
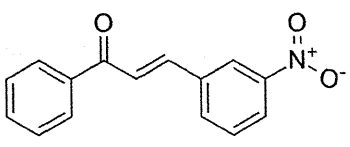
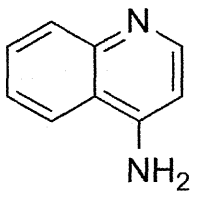
12		m0175	3-nitro-L-tyrosine	
13		m0193	mononitrosocaffeidine	
14		m0196	3,4-dihydroxy-3-cyclobuten- e- 1,2-dione	
15		m0201	2-nitrobiphenyl	
16		m0238	m-nitroanisole	
17	 This name appears to be ambiguous	m0242	tris(chloroethyl)phosphate	

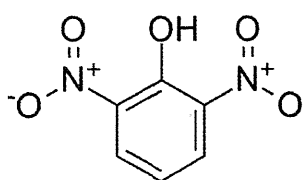
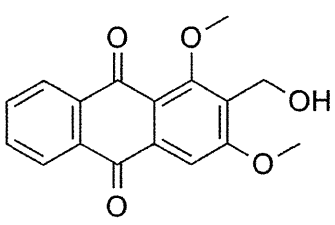
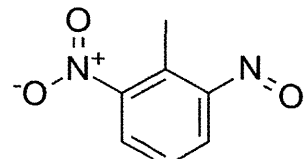
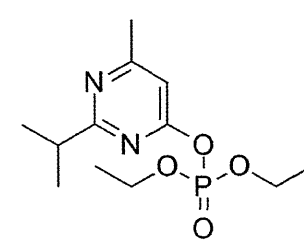
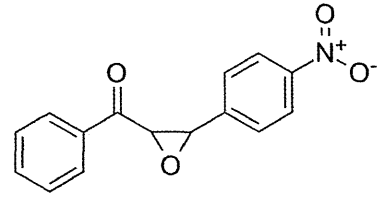
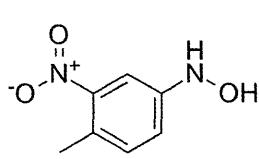
18		m0248	2,4,6-trichlorophenol	+
19	 <p>This name appears to be ambiguous</p>	m0263	z-2,3-dimethoxypropenylbenzene	+
20		m0268	amoscanate	-
21		m0283	4-(acetylamino)benzenesulfonyl chloride	-
22		m0302	4-chlorophenol	+
23		m0354	cholestane-3beta,5alpha,6beta-triol	+

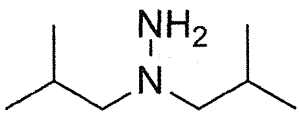
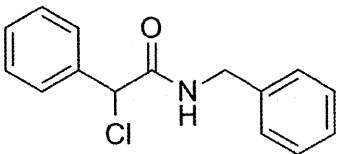
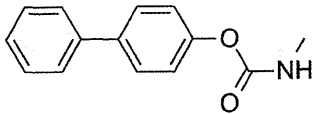
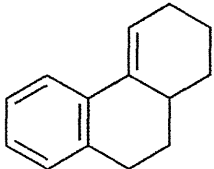
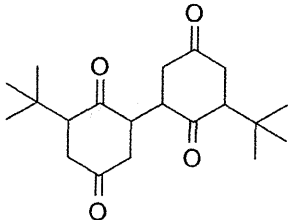
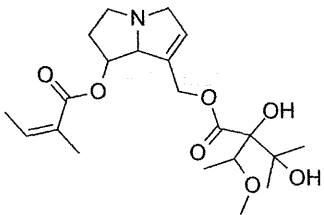
24		m0370	4'-(n,n-dimethylamino)-1'-phenylazo-3-pyridine	
25		m0380	4-amino-3-nitro-2,6-dimethylaniline	
26		m0390	n-methyl-n'-nitroguanidine	
27		m0404	trimethylolethanetrinitrate	
28		m0447	dr3354	
29	 <p data-bbox="223 1917 647 1944">Caution: Stereochemical terms discarded: -</p>	m0448	(-)-penbutolol	+

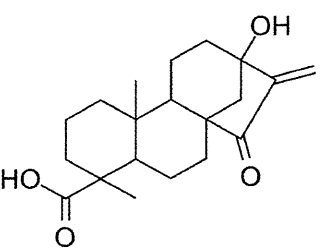
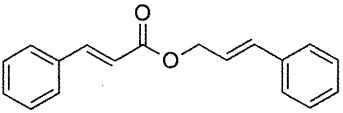
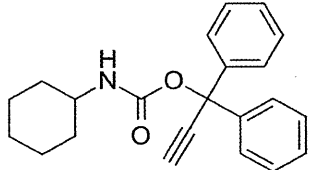
30		m0465	2-fluoropyridine	+
31		m0482	nialamide	+
32		m0487	pamaquine	+
33		m0489	isodinit	-
34		m0491	monisid	-
35		m0495	azelnidipine	-

36		m0509	gingerol	+
37		m0529	prulifloxacin	-
38		m0536	3-chlorophenol	+
39		m0552	neodecanoic acid, 2,3-epoxypropyl ester	+
40		m0577	4-(dimethylamino)benzene ethanol	+
41		m0586	tetrahydro-beta-carboline-3-carboxylic acid	-

42		m0596	n,n-diisopropylhydrazine	+
43		m0624	3,8-dimethyl-3h-imidazo [4,5-f]quinoxaline	-
44		m0625	6,7-dimethylquinoline- 5,6-diamine	-
45		m0634	4-methyl-2-nitroaniline	-
46		m0645	3-nitrochalcone	-
47		m0648	4-aminoquinoline	-

48		m0653	2,6-dinitrophenol	-
49		m0663	2-(hydroxymethyl)- 1,3-dimethoxyanthraquinone	-
50		m0671	2-nitroso-6-nitrotoluene	-
51		m0694	diazoxon	+
52		m0718	4-nitrochalcone oxide	-
53		m0720	4-hydroxylamino-2-nitrotoluen e	-

54		m0726	n,n-diisobutylhydrazine	+
55		m0812	n-benzyl-2-chloro-2-phenylacetamide	+
56		m0814	p-biphenyl-N-methylcarbamate	+
57		m0828	1,2,3,9,10,10a-hexahydrophenanthrene	+
58		m0873	3,3'-di-tert-butylbiphenyldiquinone-(2,5,2',5')	+
59		m0879	lasiocarpine	+

60		m0896	15-oxosteviol	+
61		m0930	cinnamylcinnamate	+
62		m0959	1,1-diphenyl-2-propynyl- N-cyclohexylcarbamate	+