

生態毒性QSARモデルの 開発状況について

(独)国立環境研究所

環境リスク研究センター 白石 寛明



開発の背景

化学物質審査規制法での活用を目的とした(Q)SARの開発

- 新規化学物質審査における活用
- 既存化学物質安全性点検における活用

(Q)SARのエンドポイント

- 生分解性・濃縮性 経済産業省
- 人毒性(変異原性) 厚生労働省
- 生態毒性(魚類、甲殻類、藻類急性毒性)
環境省

既存モデルの概要

対象モデル

名称	開発元	生物種	特徴
ECOSAR	US EPA	魚類、甲殻類、藻類	<ul style="list-style-type: none"> ・部分構造等によるクラス分類 ・主にLogPとの単相関による予測
TOPKAT	Health Design Inc. & Oxford Molecular Group, Inc	魚類、甲殻類	<ul style="list-style-type: none"> ・部分構造等による最適モデルの選択 ・構造記述子、物性等を用いた重相関による予測
TIMES	Univ. of Burgas	魚類、甲殻类等17種の生物種	<ul style="list-style-type: none"> ・クラス分類 ・LogP、3D記述子等を用いた予測
MCASE	Multicase Inc.	魚類	<ul style="list-style-type: none"> ・部分構造等によるクラス分類 ・LogP他物性等を用いた予測

既存モデルの検証

検証に用いたデータセット

- 環境省生態影響試験結果(平成7~15年度)※
 - 魚類(メダカ) 230物質
 - 甲殻類(オオミジンコ) 294物質
 - 藻類 232物質
- EPA Fathead Minnowデータベース
 - 魚類(ファットヘッドミノー) 580物質

※限度試験結果を除く

検証方法

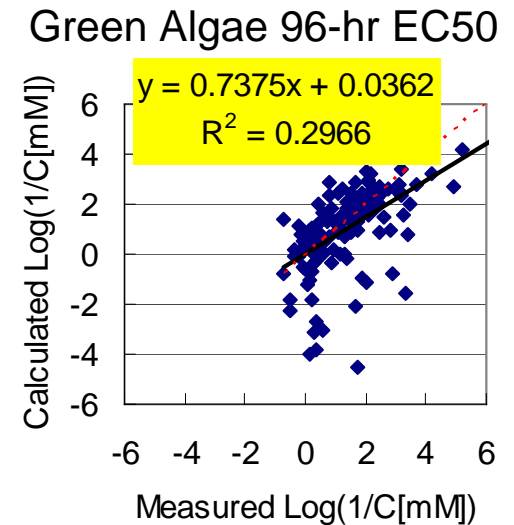
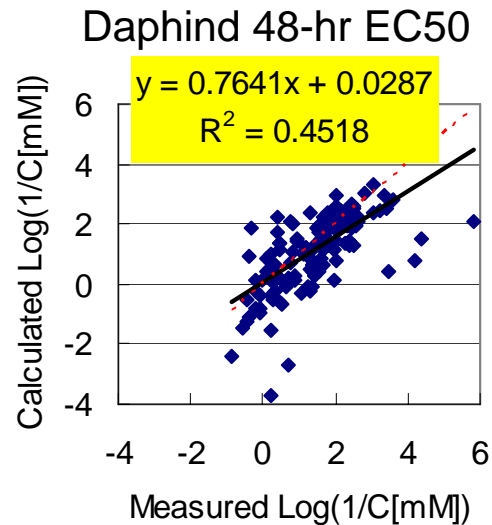
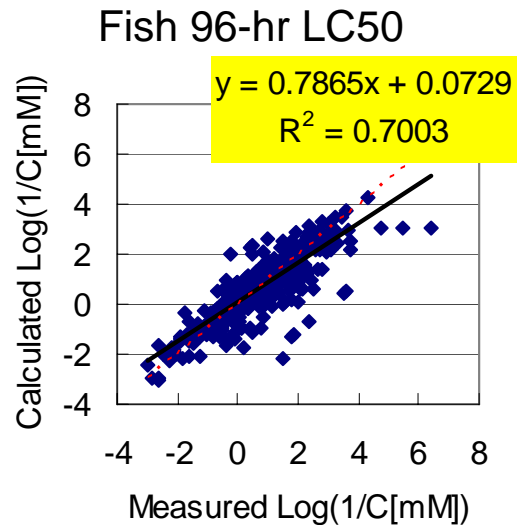
- 解析データセットの実測値とモデルによる予測値との相関解析

検証の指標

- 傾き(~ 1)、切片(~ 0)、決定係数(R^2 値)
- 2乗平均平方根誤差(RMSE: Root Mean Square Error)

$$\text{RMSE} = [\sum (\text{Predicted Value} - \text{Observed Value})^2 / n]^{0.5}$$

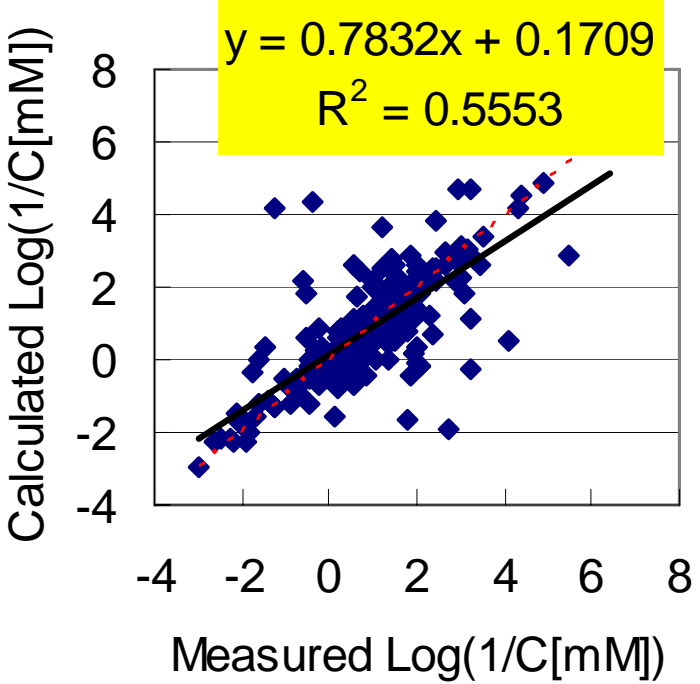
ECOSARの評価



Class	Fish 96-hrLC50					Daphnid 48-hr EC50					Algae 96-hr EC50				
	物質数	傾き	決定係数	平均誤差	RMS E	物質数	傾き	決定係数	平均誤差	RMS E	物質数	傾き	決定係数	平均誤差	RMS E
Acrylates	5	0.65	0.90	-0.08	0.19	4	-	-	-	-	5	1.83	0.99	0.02	0.26
Aldehydes	31	0.23	0.21	-0.23	0.74	1	-	-	-	-	4	-	-	-	-
Aliphatic Amines	8	0.47	0.77	0.09	0.66	9	1.38	0.59	0.58	0.80	14	0.60	0.32	-0.14	1.23
Aromatic Amines	29	0.91	0.87	-0.13	0.38	16	0.13	0.10	0.26	0.57	-	-	-	-	-
Esters	29	0.48	0.80	-0.26	1.00	13	0.70	0.48	-0.49	1.13	7	0.67	0.67	1.31	1.42
Methacrylates	6	0.40	0.33	-0.29	0.69	-	-	-	-	-	-	-	-	-	-
Neutral Organics	195	0.89	0.72	-0.12	0.72	65	1.13	0.68	-0.493	0.86	73	0.82	0.35	-0.68	1.58
Phenols	48	0.70	0.78	0.09	0.53	17	0.42	0.37	-0.443	0.86	19	0.79	0.24	-0.14	1.05
Phenols (dinitro)	5	0.52	0.66	-0.10	0.43	-	-	-	-	-	-	-	-	-	-
Propargyl Alc-hindered	9	0.20	0.16	-0.39	1.20	-	-	-	-	-	-	-	-	-	-
Vinyl/Allyl Halides	6	0.15	0.05	0.15	1.19	2	-	-	-	-	4	-	-	-	-
All	401	0.79	0.70	-0.10	0.74	146	0.76	0.45	-0.32	0.99	137	0.74	0.30	-0.33	1.43

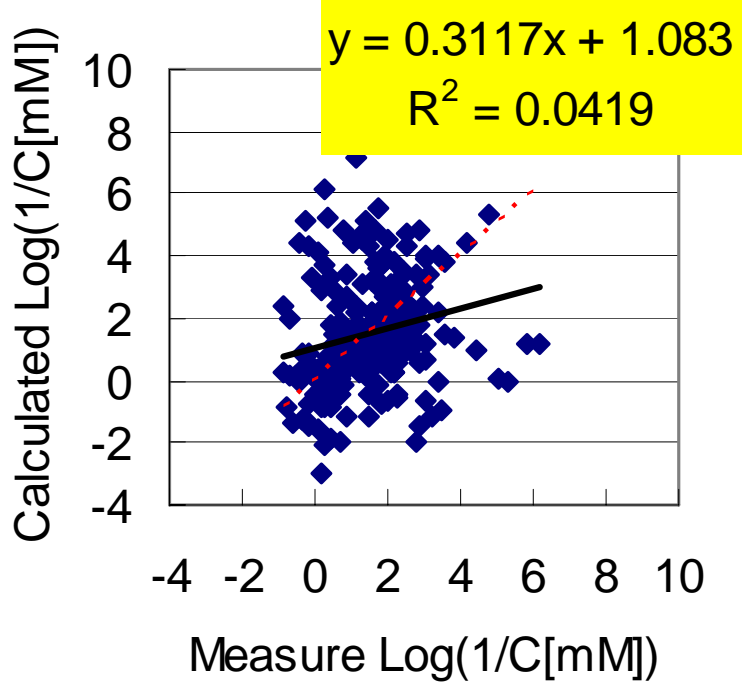
TOPKATの評価

Fathead Minnow LC50 v3.2

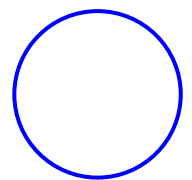
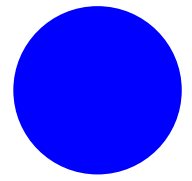
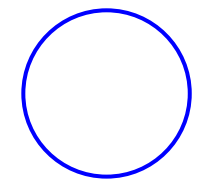


RMSE = 0.93

Daphnia magna EC50 v3.1

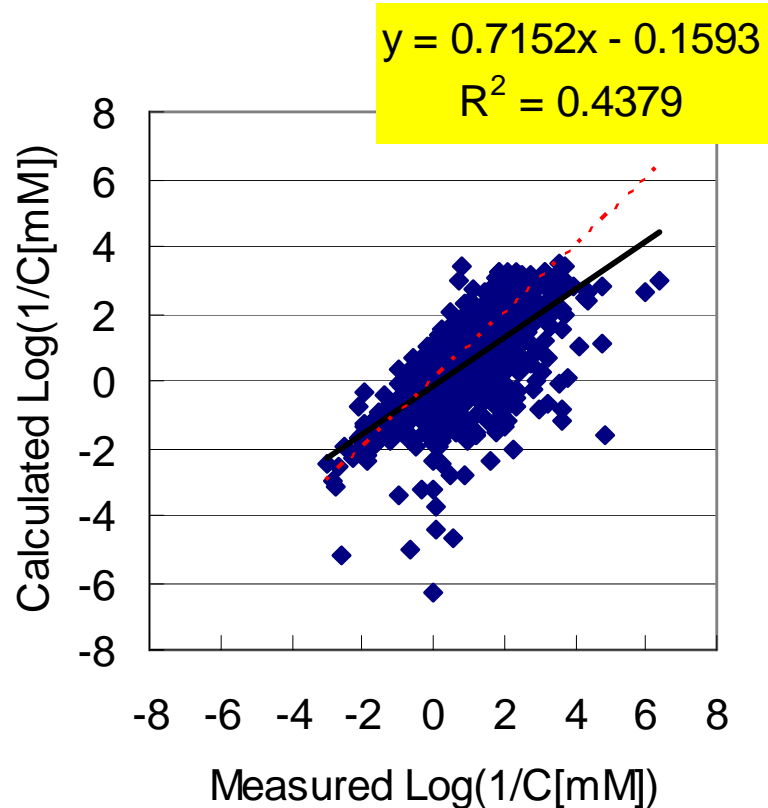


RMSE = 1.88



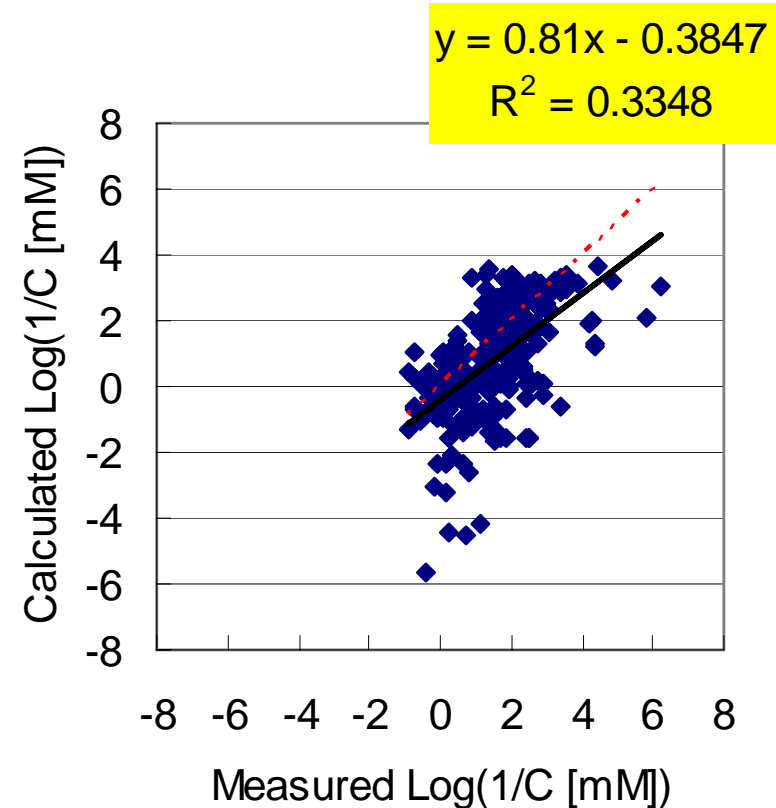
TIMESの評価

Fathead Minnow 96-hr LC50
(Interspecies Model)

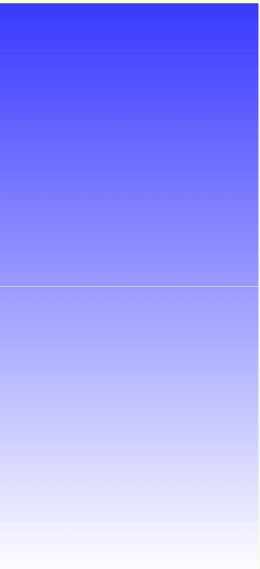
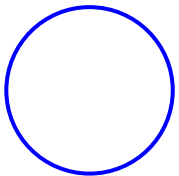
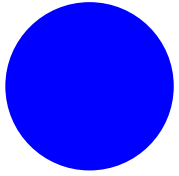
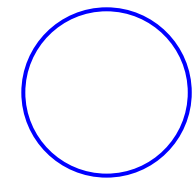


RMSE = 1.21

Daphnia magna 48-hr EC50
(Interspecies Model)

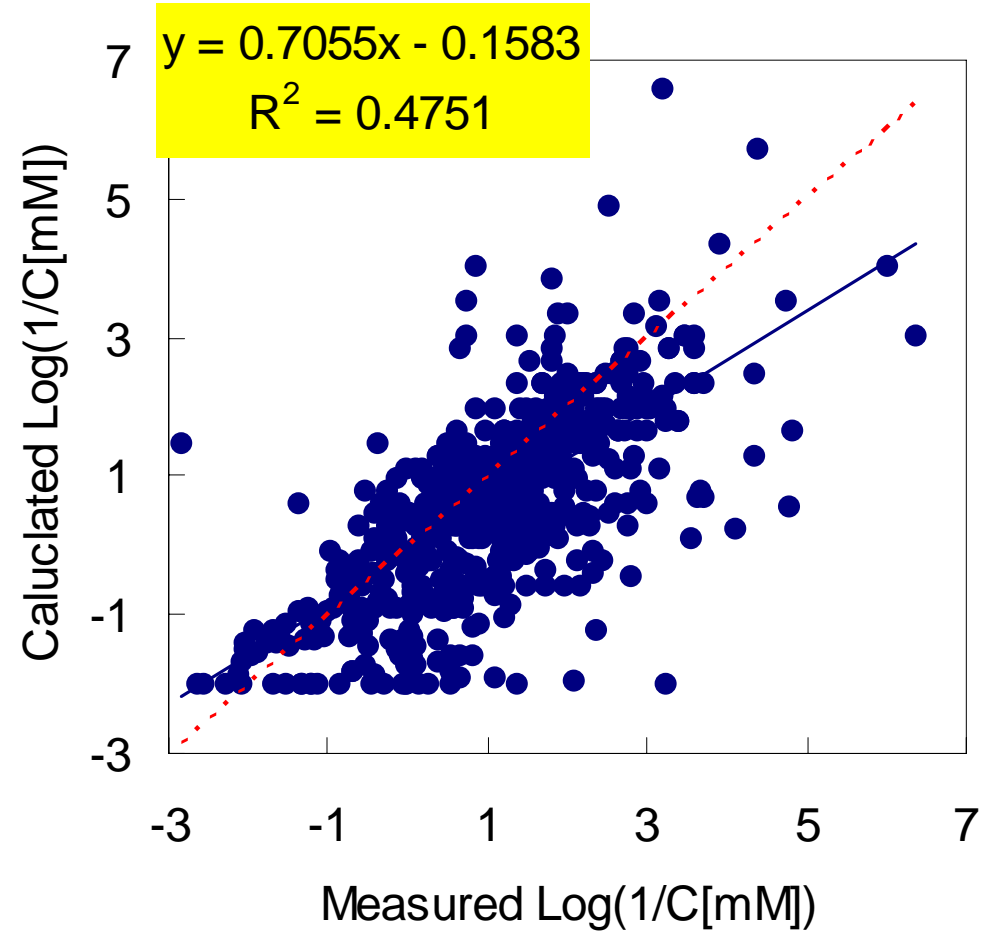


RMSE = 1.48

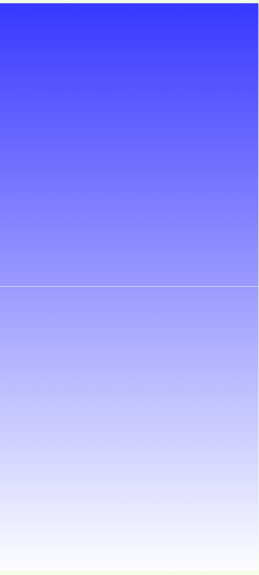
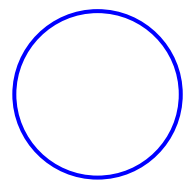
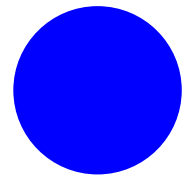
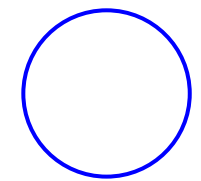


MCASEの評価

Fathead Minnow 96-hr LC50



RMSE = 1.12



既存QSARシステムの評価のまとめ

モデル名称	対象生物	傾き	決定係数	2乗平均平方根誤差 (RMSE)
ECOSAR	魚類	0.79	0.70	0.74
	甲殻類	0.76	0.45	0.99
	藻類	0.74	0.30	1.43
TOPKAT	魚類	0.78	0.55	0.93
	甲殻類	0.31	0.04	1.88
TIMES	魚類	0.72	0.44	1.21
	甲殻類	0.81	0.33	1.48
MCASE	魚類	0.71	0.48	1.12

生態毒性予測システムの開発

エンドポイント及びデータセット

エンドポイント

- 魚類、甲殻類、藻類の急性毒性
 - 魚類急性毒性試験 96hr LC50
 - 甲殻類急性遊泳阻害試験 48hr EC50
 - 藻類生長阻害試験 72hr EC50

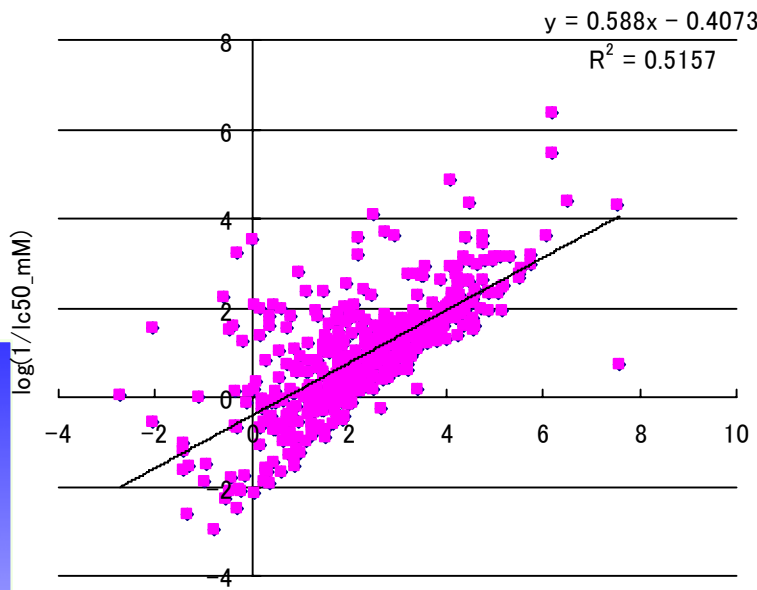
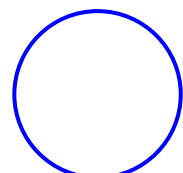
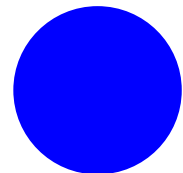
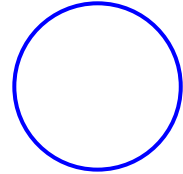
データセット

- 環境省生態影響試験結果(平成7年度~17年度)※
 - 魚類(メダカ) 275物質
 - 甲殻類(オオミジンコ) 347物質
 - 藻類 247物質
- EPA Fathead Minnowデータベース
 - 魚類(ファットヘッドミノー) 580物質

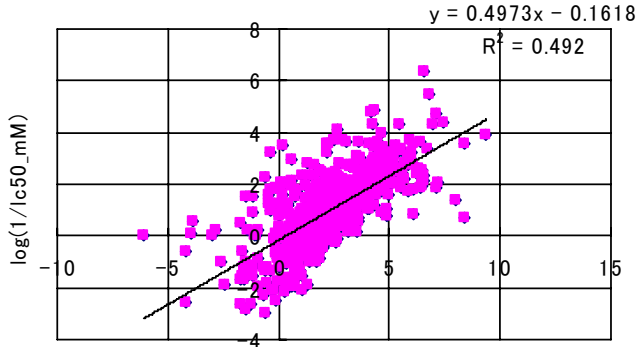
※ 限度試験を除く。



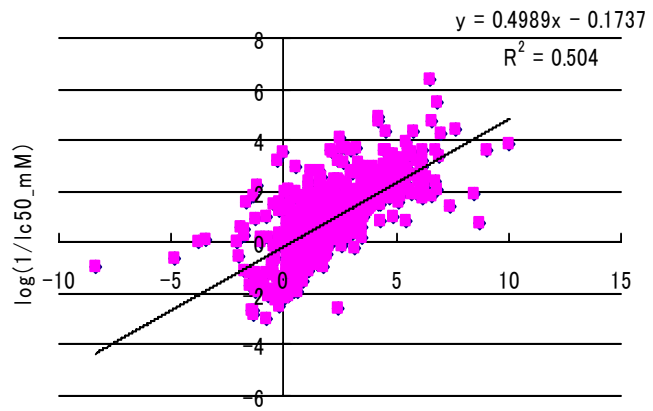
魚類96hLC50と分配係数の関係



実測値: LogP



予測値: KowWin



予測値: ClogP



ECOSAR の場合 (カテゴリーの重複)

Category1	Category2	Category3	Chemicals	
Aldehydes	AromaticAmines		1	unpredictable
Aldehydes	Phenols		2	unpredictable
Aliphatic Amines	Aromatic Amines		3	AliphaticAmines
Aliphatic Amines	Aromatic Amines	Esters	1	unpredictable
Aliphatic Amines	Esters		8	unpredictable
Aliphatic Amines	Esters	thiols	1	unpredictable
Aliphatic Amines	Esters(phosphate)		1	unpredictable
Aromatic Amines	Phenols		4	Phenols
Esters	Phenols		2	Esters
Esters	cyano		2	unpredictable
Esters(phosphate)	thiols		3	Esters(phosphate)

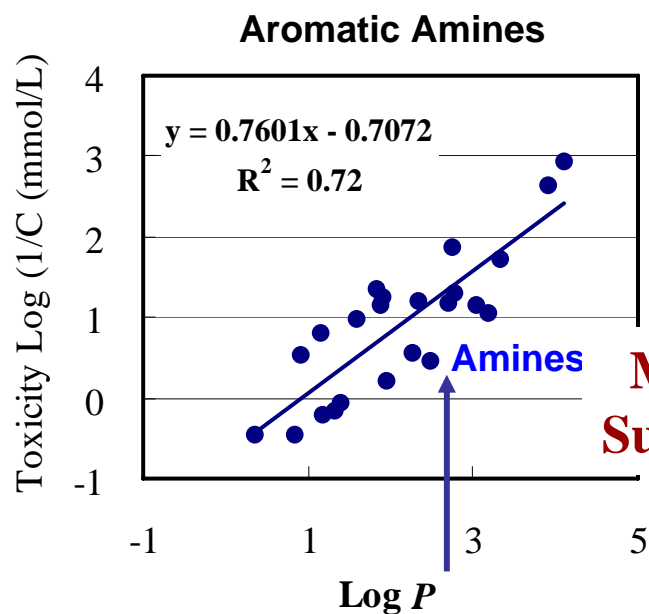
Existing Chemicals

Num of category	Chemicals	Ratio
Base line	6,457	34.2%
1	9,959	52.7%
2	1,940	10.3%
3	450	2.4%
4	Not Predict 79	0.42%
5	8	0.04%
6	2	0.01%
Sum	18,895	100.0%

3つのカテゴリーの組み合わせのルール化

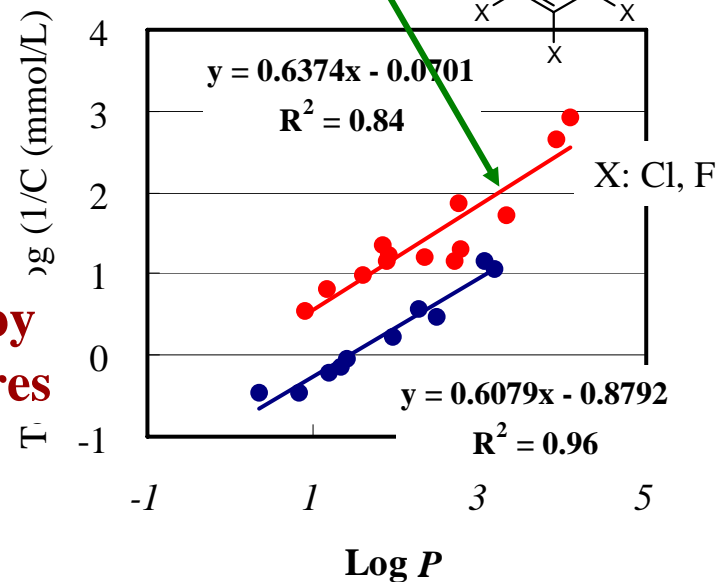
骨格構造と隣接官能基による修飾作用を考慮したサブフラグメント方式の概念

電子吸引性、共役性、反応性、立体障害



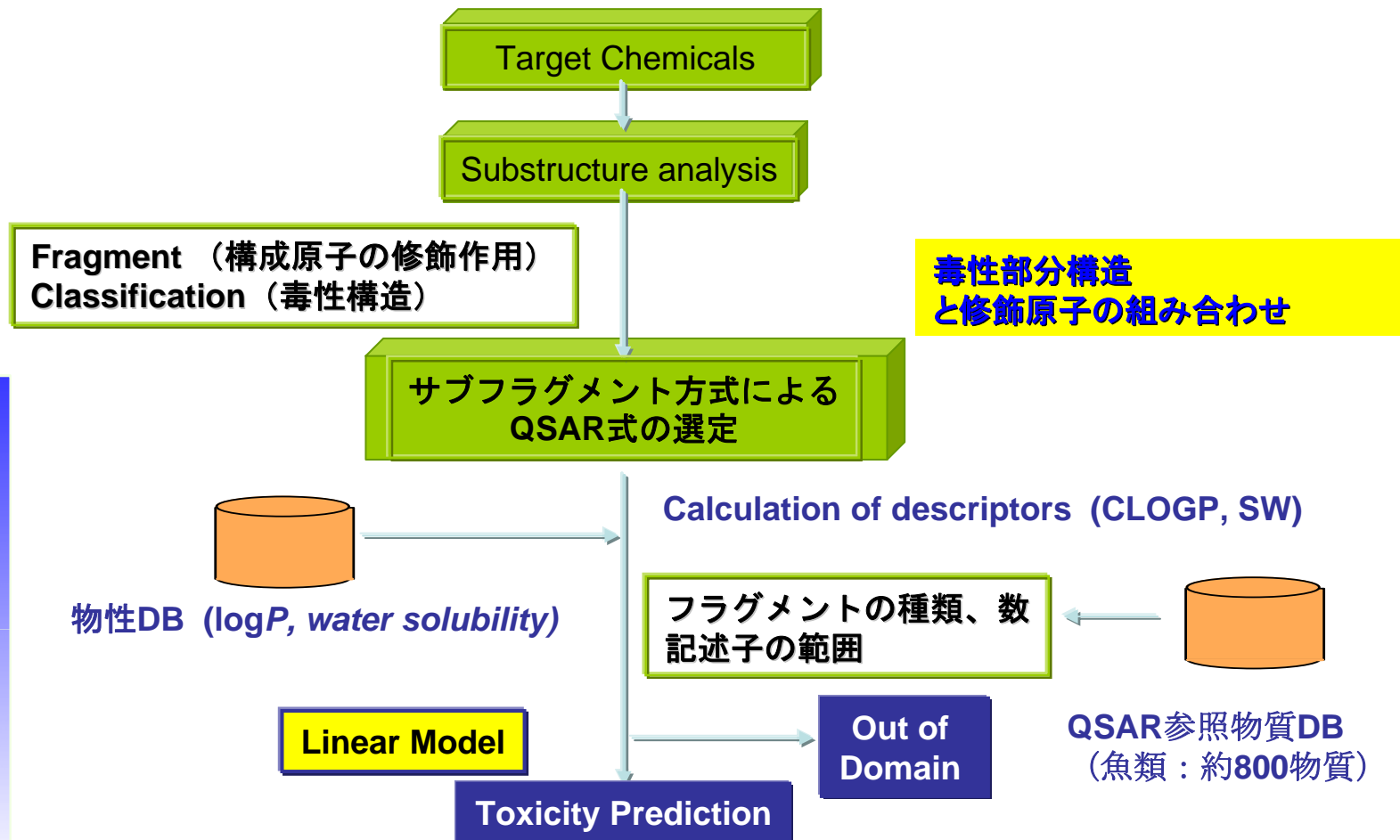
Modified by
Substructures

Strengthen toxicity

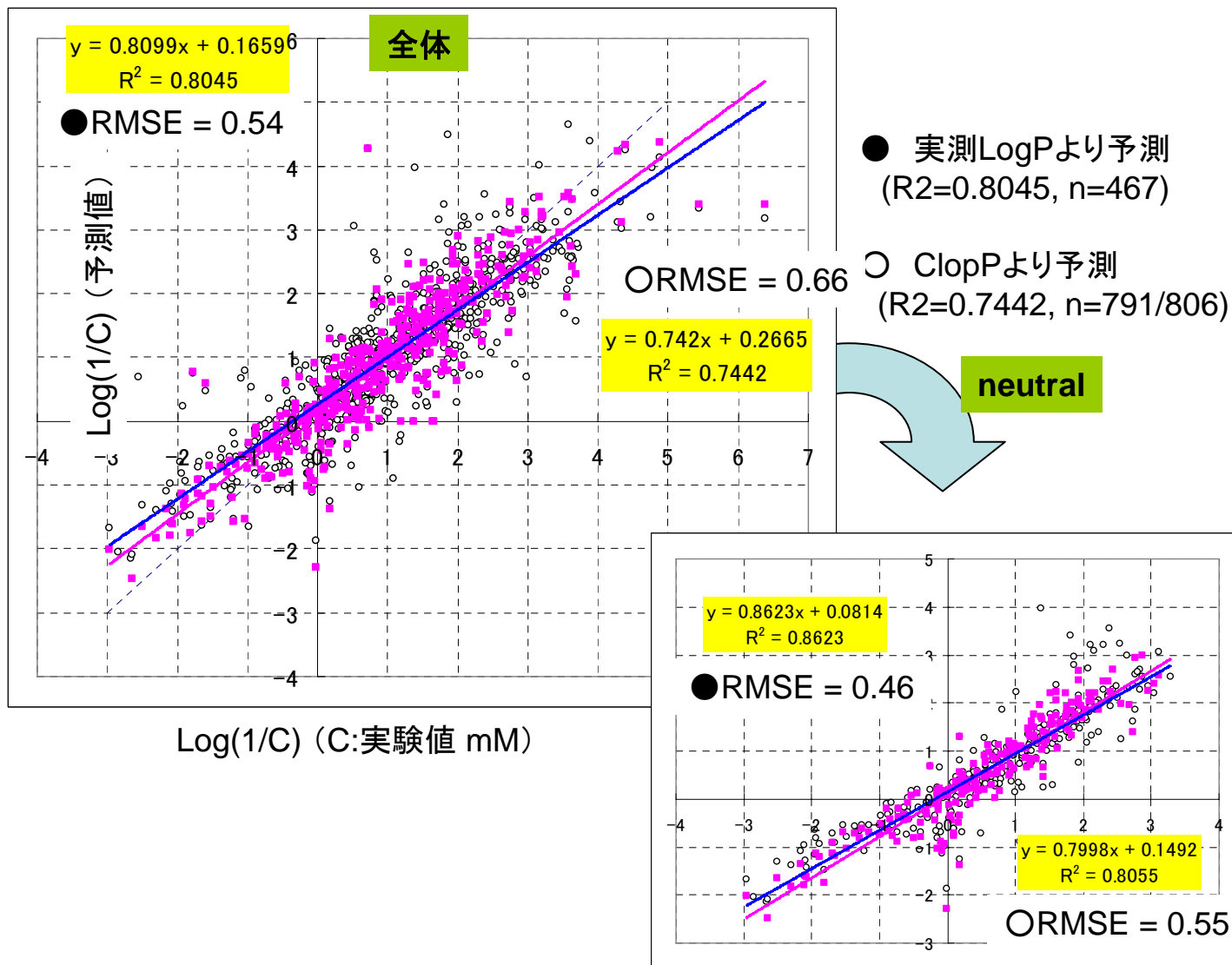


Base Line Toxicity (Narcosis)

生態毒性予測システムの基本的な考え方



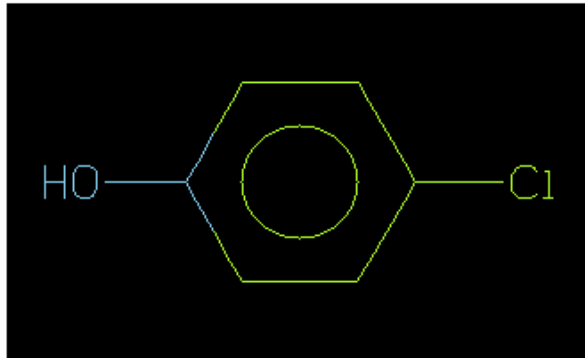
QSAR システムの検証 (開発用 α バージョン)



Web版 2 結果表示画面

[Top](#) > [Login/Logout](#) > [input chemical](#) > [view result](#) > [verify QSAR](#)

魚類急性毒性の予測結果



CAS番号	106-48-9
物質名称	P-CHLOROPHENOL
SMILES	Oc1ccc(Cl)cc1
分子量	128.560
水溶解度(mg/L)	実測値(input) <input type="text"/>
LogP	実測値DB 2.390
	予測値ClogP 2.485
	実測値(input) <input type="text"/>
色の説明	水色: 毒性が懸念される部分構造 緑色: 参照物質と共通 赤色: 参照物質と不一致 白色: 未検索

QSARクラス: Phenols

// [結果ファイル6\(RED\)の参照](#) // [更新](#)

魚類96時間LC50の予測: $\text{Log}(1/\text{C}[\text{mM}]) = 0.72 * \text{logP} - 0.77$

ClogP = 2.485 より予測: 0.0957[mmol], 12.30 [mg/L]

分配係数の手入力はされていません

ClogPの結果によれば 難分解・低濃縮性であれば第三種監視化学物質相当ではない/溶解度を確認してください

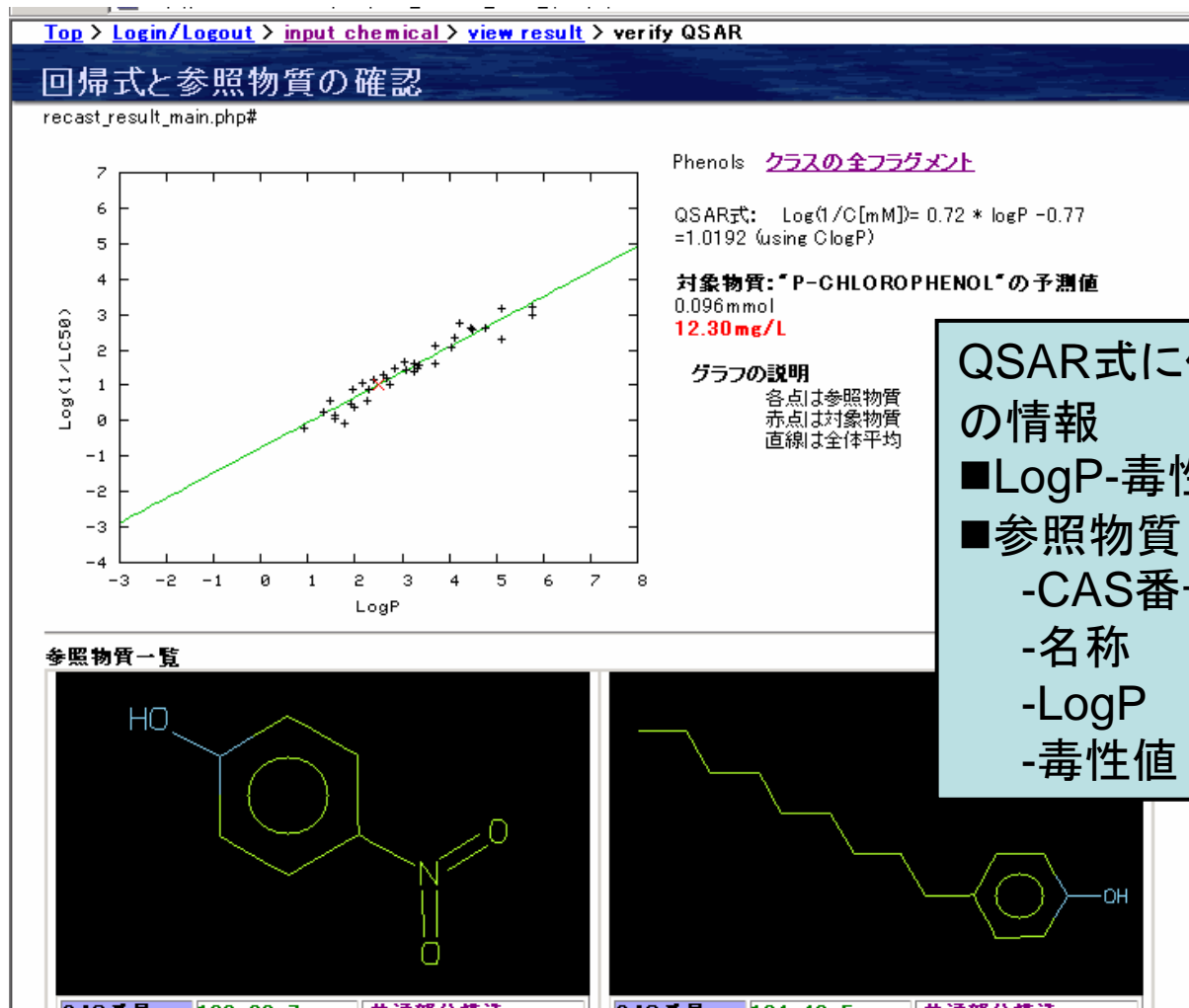
logPが参照物質の範囲内であることを確認してください

■ フラグメント表示(対象物質) - 物質名称[P-CHLOROPHENOL]

ID	名称	個数	比較
2040	Phenols	1	○
0130	AromaticCarbon	6	○
0199	-Cl[chlorine,aromaticattach]	1	○
0285	-OH[hydroxy,aromaticattach]	1	○

LogP実測濃度及び予測濃度からの毒性値予測
水溶解度実測値と毒性値の比較
抽出フラグメント及び参照物質との比較

Web版 3 参照物質画面



QSAR式に使用した参照物質
 の情報

- LogP-毒性相関図
- 参照物質
 - CAS番号
 - 名称
 - LogP
 - 毒性値

Web版 4 参照物質のフラグメントリスト

生態毒性判定システム - Microsoft Internet Explorer

生態毒性判定システム

■ フラグメントリスト

ID	名称	化合物数	最大個数	比較
2040	Phenols	38	1	○
0130	AromaticCarbon	38	12	○
0133	AromaticNitrogen	1	1	-
0146	-Br[bromine,aromaticattach]	2	3	-
0164	-C(=O)-[carbonyl,onearomaticattach]	1	1	-
0191	-CH[aliphaticcarbon]	1	1	-
0192	-CH2-[aliphaticcarbon]	4	8	-
0193	-CH3[aliphaticcarbon]	16	7	-
0199	-CL[chlorine,aromaticattach]	13	5	○
0205	-COOH[acid,aromaticattach]	2	1	-
0217	-I[aromaticattach]	1	1	-
0259	-NO2[nitro,aromaticattach]	2	1	-
0266	-O-[aliphaticO,twoaromaticattach]	2	1	-
0267	-O-[oxygen,onearomaticattach]	6	1	-
0285	-OH[hydroxy,aromaticattach]	38	1	-
0362	-tertCarbon[3ormorecarbonattach]	2	1	-

○: グループ内の他の物質にもフラグメントが存在する。

×: 自分自身にしかフラグメントがない。特異的である。



Zako2.wmv

参照物質中にあるフラグメントリスト

- 化合物数: 参照物質数
- 最大個数: 物質ごとに含まれるフラグメントの個数の最大値
- 比較: 対象物質に含まれる場合に○

ウィンドウズ版 (αバージョン)

JQSAR: Japan QSAR for Eco-Toxicity Assessment Ver0.06.12.30

ファイル(E) データ管理 ヘルプ プログラム開発 QSAR研究・開発

ID番号 CAS NO.

SMILES

名称

QSAR実行

登録番号=7282 が、ありました。

logPow 実測値

類似化合物検索

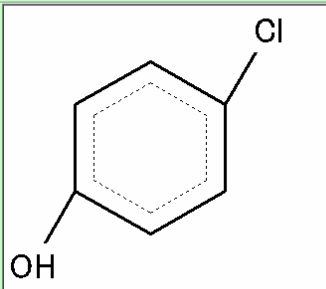
- ・CAS番号からの検索
- ・SMILESからの2次元構造の描画
- ・2次元構造描画ツール

JQSAR-Cal

Return Specific Structure 拡大図表示 ヘルプ

毒性推定結果(概要) 推定詳細 logPow推定詳細

Factor Code No=0 SMILES: Oc(ccc(c1)Cl)c1
***** JQSAR 出力 *****
100 No= 130 n= 6 Frag Aromatic Carbon
1630 No= 236 n= 1 Frag -OH [hydroxy, aromatic attach]
2410 No= 199 n= 1 Frag -Cl [chlorine, aromatic attach]
***** JQSAR Estimate Equation *****
10650 No=2052 n= 1 JQSAR NeutralOrganics
***** 終了 *****



拡大図 再計算 logPow 実測値 logPow 推定
M.W Cal 128.558

Information

- ・毒性構造
- ・フラグメントの抽出
- ・QASR式の判定



現状のまとめ

- 毒性データセットのスクリーニングおよびモデル構築等の作業に供するための開発基盤、枠組みの整備
- 魚類急性毒性についての構造活性相関モデルのサブフラグメント法による構築
- 生態毒性予測システム（WEB版、ウィンドウズ版）の作成と、その公開に向けた準備
- OECD会合、第一回（定量的）構造活性相関についてのアドホックグループ会合（2006年6月、於イタリア、ストレーザ）にて生態毒性予測システムの開発状況を報告

今後の予定

- 魚類モデルの検証（validation）と公開
- 化学物質のカテゴリ分類手法の検討、モデル構築手法の検討、
- 甲殻類、藻類に対する毒性についてのQSARモデル構築

