

201428005A

厚生労働科学研究費補助金

化学物質リスク研究事業

Q S A R 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26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27(2015)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. 構造類似性に基づく毒性予測法の開発

栗原 正明 7

2. 化学物質の構造の精密化・高度化

出水 庸介 117

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

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となってきた。本研究では、現在ある予測法を評価するとともに、様々な手法により、より高度な新しい予測システムを開発する。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。特に構造類似性のみに基づいた新規予測法の開発を行った。また、化学物質の構造の精密化、高度化を行った。

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

構造類似性に基づく毒性予測法の開発

【データの整理】

昨年と同様に以下の手順で有効母集団 5067 化合物を設定した。

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。

任意の 214 件を抽出し検証集団とした。

【構造類似化合物の過半数による予測法】
検証化合物のAmes試験を予測するときに構造類似化合物のAmes試験の過半数の結果を予測値とする。

構造類似化合物の過半数による予測については以下の4つの手法で各214件の検証化合物の類似化合物を有効母集団5067件から抽出し、検証した。

- ① 90%以上類似度母集団：母集団の最も多くのAmes試験結果（90%）
- ② 最も類似した化合物のAmes試験結果（TOP1）
- ③ 最も類似した3つの化合物から最も多くのAmes試験結果（TOP3）
- ④ 最も類似した5つの化合物から最も多くのAmes試験結果（TOP5）

【組み合わせによる予測法】

上記の4つの手法の結果を組み合わせる方法も行った。

【類似学習母集団からの予測モデル作成（QSAR）について】

上記4つの手法で予測結果が一致しなかった83件に対し、以下の方法によって、予測モデル作成を行いました。

（1）パラメータ発生：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。

（2）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。

（3）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（ $|t\text{-statistics}| > 2.0$ かつ $p\text{-value} < 0.05$ ）に絞りました。最終的にパラメータを47個に絞った。

（4）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、高度化に関する研究

検証母集団の214化合物について化学計算による化合物の精密な構造データの取得を行った。方法は昨年度と同じ方法で行った。分子力学計算（MMFF, OPLS2005 力場等）を用いたコンフォメーション探索により最安定構造を計算した。ソフトウェアはシュレディンガー社のMacromodelを用いた。

（I）初期座標の入力

グラフィカルユーザーインターフェース（GUI）であるMaestro 9.1を用いて初期構造を入力した。構造異性体を含むと考えられる化学物質については、妥当と考えられる代表的な構造を用いた。分子力学計算で安定化を行い、構造を整えた。（ローカルミニマム）

（II）配座探索により最安定構造（グローバルミニマム）を求める

配座探索（Conformational Search）の方法はMacromodelのMCMM（Torsion sampling）法を用いた。力場はMMFF, OPLS 2005を用いた。1000個の初期座標を発生させ、それぞれの安定配座を計算し、最もエネルギーの低いものを最安定コンフォーマー（最安定構造）とした。

（III）ファイルフォーマットの変換

最安定構造の座標をmol2ファイルに変換した。mol2ファイルは生物活性値とデータセットを作る際の必要となる。

（IV）最安定構造の表示

分子表示ソフトウェアDS viewerPro 6.0を用いて最安定構造を分子モデルで表示した。

C. 研究結果

構造類似性に基づく毒性予測法の開発

【構造類似化合物の過半数による予測法】
構造類似化合物の過半数による予測の結果については表1に示した。

- ① 90%以上類似度母集団：母集団の最も多くのAmes試験結果（90%）
- ② 最も類似した化合物のAmes試験結果（TOP1）
- ③ 最も類似した3つの化合物から最も多くのAmes試験結果（TOP3）
- ④ 最も類似した5つの化合物から最も多くのAmes試験結果（TOP5）

【組み合わせによる予測法】

上記の4つの手法が一致した時のAmes試験結果を採用した場合の検証結果は表2の「4」で示した。検証化合物214件中83件が未確定で、131件の予測結果が確定した。確定した予測の内誤分類が6件で、95.4%（125/131）の高予測精度を達成できた。但し、カバー率が61.2%（131/214）で90%以下の低い数字となった。3つが一致した時のAmes試験結果を採用した場合の検証結果は表2の「3」で示した。検証化合物214件中191件の予測結果が確定し、その内誤分類が17件で、91.1%（174/191）の予測精度を達成できた。カバー率が89.3%（191/214）となった。

【類似学習母集団からの予測モデル作成（QSAR）について】

4つの手法が一致した場合の結果を採用し、一致しなかった場合はQSAR予測モデルを採用した結果は表2の「4+QSAR」で示した。

また、4つの手法の内3つが一致した場合の結果を採用し、2つ以下しか一致しなかった場合はQSAR予測モデルを採用した結果は表2の「3+QSAR」で示した。

化学物質の構造の精密化、高度化に関する研究

本プロジェクトで用いた、検証母集団の214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

予測精度95%以上かつカバー率90%以上を目指すために、TOP1で誤分類された28件の化合物（表2を参照）の共通点を検証し、その共通点（部分構造など）が存在すれば、予測に誤分類される化合物を予め予測対象から除外できるため、TOP1手法の予測精度向上を見込める。

現在のところ2次元の構造類似性による予測法を行っているが、3次元構造による構造類似性による予測法を検討することは非常に興味深い問題である。特に、最も構造類似性の高い化合物と活性が不一致のものについては化学構造及び立体構造による考察が必要であると考えられる。

E. 結論

通常のQSAR法を使わずに、最も構造類似性の高い化合物のAmes試験結果を予測値とするだけで、86%の予測精度が得られた。また、カバー率を下げるならば、ある種の化合物に対して非常に高い精度で予測を行えることが明らかとなった。このことは、この手法の可能性を示すことができた。

検証母集団の214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)

T. Misawa, Y. Demizu, M. Kawamura, N. Yamagata, M. Kurihara; Structural development of stapled short helical peptides as vitamin D receptor (VDR)-coactivator interaction inhibitors; *Bioorg. Med. Chem.*, 2015, 23, 1055-1061

Y. Demizu, H. Yamashita, T. Misawa, M. Doi, M. Tanaka, M. Kurihara; Effects of D-Leu residues on the helical secondary structures of L-Leu-based nonapeptides; *Chem. Pharm. Bull.* 63, 218-224 (2015)

H. Yamashita; Y. Demizu; T. Misawa; T. Shoda; M. Kurihara; Synthesis of a bis-cationic, α -disubstituted amino acid (9-amino-bispidine-9-carboxylic acid) and its effects on the conformational properties of peptides; *Tetrahedron*, 71, 2241-2245 (2015)

T. Shoda, K. Okuhira, M. Kato, Y. Demizu, H. Inoue, M. Naito, M. Kurihara; Design and synthesis of tamoxifen derivatives as a selective estrogen receptor down-regulator *Bioorg. Med. Chem. Lett.*, 24, 87-89 (2014)

H. Yamashita, Y. Demizu, T. Shoda, Y. Sato, M. Oba, M. Tanaka, M. Kurihara; Amphipathic short helix-stabilized peptides with cell-membrane penetrating ability *Bioorg. Med. Chem.* 22, 2403-2408 (2014)

K. Zaima, D. Wakana, Demizu, Y. Kumeta, H. Kamakura, T. Maruyama, M. Kurihara, Y. Goda; Isoheleproline: a new amino acid-sesquiterpene adduct from *Inula helenium*; *J Nat Med*, 68, 432-435 (2014)

Y. Demizu, H. Yamashita, Y. Sato, M. Doi, M. Tanaka, M. Kurihara; Helical Screw-sense Control of LD-Peptides Containing

Equal Amounts of L- and D-Amino Acids *Peptide Sciences* 2013, 271-272 (2014)

N. Yamazaki, Y. Demizu, Y. Sato, M. Doi, M. Kurihara; Development of Helical Foldamer Containing a Combination of Cyclopentane-1,2-Diamine and 2,2-Dimethylmalonic Acid; *Peptide Sciences* 2013, 273-274 (2014)

T. Sugiyama, K. Kuwata, Y. Imamura, Y. Demizu, M. Kurihara, M. Takano, A. Kittaka; Sequence-Specific Cleavage of DNA by β -Peptide Acid Bearing σ Pendant Metal Complex; *Peptide Sciences* 2013, 369-370 (2014)

S. Yamada, Y. Kotake, Y. Demizu, M. Kurihara, Y. Sekino, Y. Kanda; NAD-dependent isocitrate dehydrogenase as a novel target of tributyltin in human embryonic carcinoma cells; *Sci Rep.* 2014; 4: 5952.

N. Hirata, S. Yamada, T. Shoda, M. Kurihara, Y. Sekino, Y. Kanda; Sphingosine-1-phosphate promotes expansion of cancer stem cells via S1PR3 by a ligand-independent Notch activation; *Nature Communications* 5, Article number: 4806

Imai, K., Nakanishi, I., Ohno, A., Kurihara, M., Miyata, N., Matsumoto, K., Nakamura, A., Fukuhara, K.; Synthesis and radical-scavenging activity of a dimethyl catechin analogue; *Bioorg. Med. Chem. Lett.*, 24, 2582-4 (2014)

Oba, M.; Takazaki, H.; Kawabe, N.; Doi, M.; Demizu, Y.; Kurihara, M.; Kawakubo, H.; Nagano, M.; Suemune, H.; Tanaka, M.; Helical peptide-foldamers having a chiral five-membered ring amino acid with two azido functional groups; *J. Org. Chem.*, 79, 9125-9140 (2014)

M. Oba, N. Kawabe, H. Takasaki, Y. Demizu; M. Doi; M. Kurihara; H. Suemune; M. Tanaka; Conformational studies on peptides having chiral five-membered ring amino acid with two azido or triazole functional groups within the sequence of Aib residues; *Tetrahedron*, **70**, 8900-8907(2014)

T. Nagakubo, Y. Demizu, Y. Kanda, T. Misawa, T. Shoda, K. Okuhira, Y. Sekino, M. Naito, M. Kurihara; Peptides as Inhibitors of Estrogen Receptor-Mediated Transcription; *Bioconjugate Chem.*, 2014, 25, 1921–1924

2. 学会発表

なし

H. 知的財産権の出願・登録状況

なし

表 1 : 構造類似性に基づく毒性予測 (1)

	90%	TOP1	TOP3	TOP5
TN	95	123	121	120
TP	59	63	61	62
FN	13	10	12	11
FP	44	18	20	21
TOTAL	211	214	214	214
NG= FN+FP	57	28	32	32
未確定	3	0	0	0
カバー 率	0.986	1.000	1.000	1.000
予測率	0.730	0.869	0.850	0.850
感度	0.819	0.863	0.836	0.849
特異度	0.683	0.872	0.858	0.851

表 2 : 構造類似性に基づく毒性予測 (2)

	4	3	QSAR	4+QSAR	3+QSAR
TN	77	115	42	118	123
TP	48	59	14	62	64
FN	2	6	7	9	9
FP	4	11	13	17	17
TOTAL	131	191	76	206	213
NG= FN+FP	6	17	20	26	26
未確定	83	23	8	8	1
カバー 率	0.612	0.893	0.905	0.963	0.995
予測率	0.954	0.911	0.737	0.874	0.878
感度	0.960	0.908	0.667	0.873	0.877
特異度	0.951	0.913	0.764	0.874	0.879

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：構造類似性に基づく毒性予測法の開発

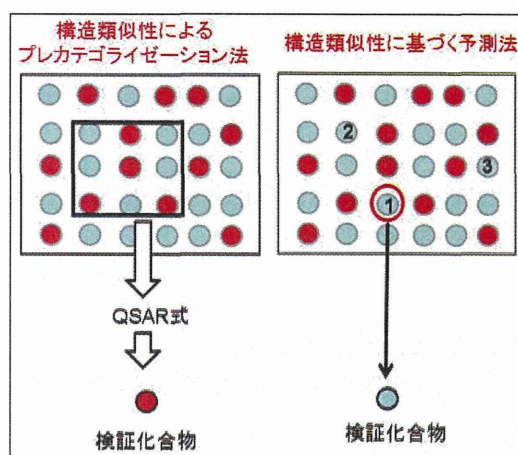
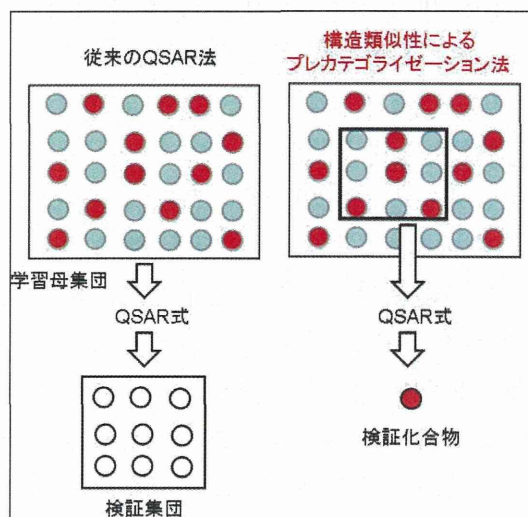
研究分担者：栗原正明 国立医薬品食品衛生研究所有機化学部 部長

研究要旨

昨年度までに構造類似学習母集団を用いた QSAR モデルの作成を行った。すなわち有効母集団 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。今年度は、構造類似性のみに基づいた新規予測法の開発を行った。有効母集団の中からターゲット化合物に対する構造類似性が高い化合物を抽出し、構造類似性の高い化合物の活性値によりターゲット化合物の活性値を予測する方法論の開発を行った。

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。昨年度は構造類似性によるプレカテゴリー化法の開発を行った。本年度は構造類似性のみに基づいた新規予測法の開発を目的とした。



B. 研究方法

データの整理

昨年と同様に以下の手順で有効母集団 5067 化合物を設定した。

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 (図 1)

その結果, 5067 件を有効母集団とした.

任意の 214 件を抽出し検証集団とした.

構造類似化合物の過半数による予測法

検証化合物の Ames 試験を予測するときに構造類似化合物の Ames 試験の過半数の結果を予測値とする.

構造類似化合物の過半数による予測については以下の 4 つの手法で各 214 件の検証化合物の類似化合物を有効母集団 5067 件から抽出し, 検証した.

- ① 90%以上類似度母集団: 母集団の最も多くの Ames 試験結果 (90%)
- ② 最も類似した化合物の Ames 試験結果 (TOP1)
- ③ 最も類似した 3 つの化合物から最も多くの Ames 試験結果 (TOP3)
- ④ 最も類似した 5 つの化合物から最も多くの Ames 試験結果 (TOP5)

組み合わせによる予測法

上記の 4 つの手法の結果を組み合わせで予測する方法も行った.

類似学習母集団からの予測モデル作成 (QSAR) について

上記 4 つの手法で予測結果が一致しなかった 83 件に対し, 以下の方法によって予測モデル作成を行った.

(1) パラメータ発生: 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.

(2) 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.

(3) 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 ($|t\text{-statistics}| > 2.0$ かつ

$p\text{-value} < 0.05$) に絞った. 最終的にパラメータを 47 個に絞った.

(4) 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

構造類似化合物の過半数による予測法

構造類似化合物の過半数による予測の結果については表 1 ~ 表 4 に示した.

- ① 90%以上類似度母集団: 母集団の最も多くの Ames 試験結果 (90%) (表 1)
- ② 最も類似した化合物の Ames 試験結果 (TOP1) (表 2)
- ③ 最も類似した 3 つの化合物から最も多くの Ames 試験結果 (TOP3) (表 3)
- ④ 最も類似した 5 つの化合物から最も多くの Ames 試験結果 (TOP5) (表 4)

組み合わせによる予測法

上記の 4 つの手法が一致した時の Ames 試験結果を採用した場合の検証結果は表 5 の「4 つの結果が一致」で示した. 検証化合物 214 件中 83 件が未確定で, 131 件の予測結果が確定した. 確定した予測の内訳分類が 6 件で, 95.4% (125/131) の高予測精度を達成できた. 但し, カバー率が 61.2% (131/214) で 90%以下の低い数字となった. 3 つが一致した時の Ames 試験結果を採

用した場合の検証結果は表5の「3つの結果が一致」で示した。検証化合物214件中191件の予測結果が確定し、その内誤分類が17件で、91.1% (174/191) の予測精度を達成できた。カバー率が89.3% (191/214) となった。

類似学習母集団からの予測モデル作成 (QSAR) について

4つの手法が一致した場合の結果を採用し、一致しなかった場合は QSAR 予測モデルを採用した結果は表6の「4+QSAR」で示しています。

また、4つの手法の内3つが一致した場合の結果を採用し、2つ以下しか一致しなかった場合は QSAR 予測モデルを採用した結果は表6の「3+QSAR」で示した。

D. 考察

予測精度95%以上かつカバー率90%以上を目指すために、TOP1で誤分類された28件の化合物(表2を参照)の共通点を検証し、その共通点(部分構造など)が存在すれば、予測に誤分類される化合物を予め予測対象から除外できるため、TOP1手法の予測精度向上を見込める。

また、28件の一部の共通部分構造に対し、全体の有効母集団5067件での出現頻度は表7に示しています。

E. 結論

通常 QSAR 法を使わずに、最も構造類似性の高い化合物の Ames 試験結果を予測値とするだけで、86%の予測精度が得られた。また、カバー率を下げるならば、ある種の化合物に対して非常に高い精度で予測を行えることが明らかとなった。このことは、この手法の可能性を示すことができた。さ

らなる検討が必要である。

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)

T. Misawa, Y. Demizu, M. Kawamura, N. Yamagata, M. Kurihara; Structural development of stapled short helical peptides as vitamin D receptor (VDR)-coactivator interaction inhibitors; *Bioorg. Med. Chem.*, **23**, 1055–1061 (2015)

Y. Demizu, H. Yamashita, T. Misawa, M. Doi, M. Tanaka, M. Kurihara; Effects of D-Leu residues on the helical secondary structures of L-Leu-based nonapeptides; *Chem. Pharm. Bull.* **63**, 218–224 (2015)

H. Yamashita; Y. Demizu; T. Misawa; T. Shoda; M. Kurihara; Synthesis of a bis-cationic, α -disubstituted amino acid (9-amino-bispidine-9-carboxylic acid) and its effects on the conformational properties of peptides; *Tetrahedron*, **71**, 2241–2245 (2015)

T. Shoda, K. Okuhira, M. Kato, Y. Demizu, H. Inoue, M. Naito, M. Kurihara; Design and synthesis of tamoxifen derivatives as a selective estrogen receptor down-regulator *Bioorg. Med. Chem. Lett.*, **24**, 87–89 (2014)

H. Yamashita, Y. Demizu, T. Shoda, Y. Sato, M. Oba, M. Tanaka, M. Kurihara; Amphipathic short helix-stabilized peptides with cell-membrane penetrating ability *Bioorg. Med. Chem.* **22**, 2403–2408 (2014)

K. Zaima, D. Wakana, Demizu, Y. Kumeta, H. Kamakura, T. Maruyama, M. Kurihara, Y. Goda; Isoheleproline: a new amino

acid-sesquiterpene adduct from *Inula helenium*; *J Nat Med*, **68**, 432-435(2014)

Y. Demizu, H. Yamashita, Y. Sato, M. Doi, M. Tanaka, M. Kurihara; Helical Screw-sense Control of LD-Peptides Containing Equal Amounts of L- and D-Amino Acids *Peptide Sciences* 2013, 271-272(2014)

N. Yamazaki, Y. Demizu, Y. Sato, M. Doi, M. Kurihara; Development of Helical Foldamer Containing a Combination of Cyclopentane-1,2-Diamine and 2,2-Dimethylmalonic Acid; *Peptide Sciences* 2013, 273-274 (2014)

S. Yamada, Y. Kotake, Y. Demizu, M. Kurihara, Y. Sekino, Y. Kanda; NAD-dependent isocitrate dehydrogenase as a novel target of tributyltin in human embryonic carcinoma cells; *Sci Rep.* **2014**; 4: 5952.

N. Hirata, S. Yamada, T. Shoda, M. Kurihara, Y. Sekino, Y. Kanda; Sphingosine-1-phosphate promotes expansion of cancer stem cells via S1PR3 by a ligand-independent Notch activation; *Nature Communications* 5, Article number: 4806

Imai, K., Nakanishi, I., Ohno, A., Kurihara, M., Miyata, N., Matsumoto, K., Nakamura, A., Fukuhara, K.; Synthesis and radical-scavenging activity of a dimethyl

catechin analogue; *Bioorg. Med. Chem. Lett.*, **24**, 2582-4(2014)

Oba, M.; Takazaki, H.; Kawabe, N.; Doi, M.; Demizu, Y.; Kurihara, M.; Kawakubo, H.; Nagano, M.; Suemune, H.; Tanaka, M.; Helical peptide-foldamers having a chiral five-membered ring amino acid with two azido functional groups; *J. Org. Chem.*, **79**, 9125–9140(2014)

M. Oba, N. Kawabe, H. Takasaki, Y. Demizu; M. Doi; M. Kurihara; H. Suemune; M. Tanaka; Conformational studies on peptides having chiral five-membered ring amino acid with two azido or triazole functional groups within the sequence of Aib residues; *Tetrahedron*, **70**, 8900-8907(2014)

T. Nagakubo, Y. Demizu, Y. Kanda, T. Misawa, T. Shoda, K. Okuhira, Y. Sekino, M. Naito, M. Kurihara; Peptides as Inhibitors of Estrogen Receptor-Mediated Transcription; *Bioconjugate Chem.*, **25**, 1921–1924(2014)

2. 学会発表

なし

H. 知的財産権の出願・登録状況

なし

解析の結果を別表 1～7 として添付した (p. 16～116)

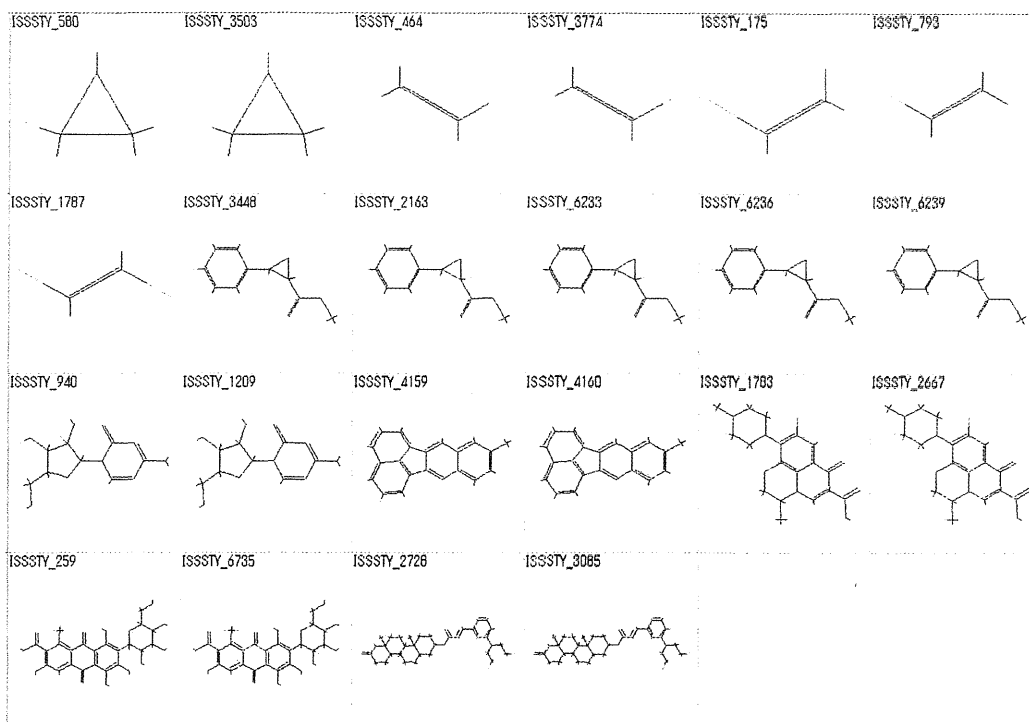


図 1：異なる Ames 試験結果が掲載されている化合物（22 件）

表 1 : 90%以上類似度母集団の過半数を予測値とする

	90%
TN	95
TP	59
FN	13
FP	44
TOTAL	211
NG= FN+FP	57
未確定	3
カバー率	0.986
予測率	0.730
感度	0.819
特異度	0.683

表 2 : 最も類似性の高い化合物による予測結果

	TOP1
TN	123
TP	63
FN	10
FP	18
TOTAL	214
NG= FN+FP	28
未確定	0
カバー率	1.000
予測率	0.869
感度	0.863
特異度	0.872

表 3 : 最も類似した 3 つの化合物から最も多くの Ames 試験結果を予測値とする

	TOP3
TN	121
TP	61
FN	12
FP	20
TOTAL	214
NG= FN+FP	32
未確定	0
カバー率	1.000
予測率	0.850
感度	0.836
特異度	0.858

表 4 : 最も類似した 5 つの化合物から最も多くの Ames 試験結果を予測値とする

	TOP5
TN	120
TP	62
FN	11
FP	21
TOTAL	214
NG= FN+FP	32
未確定	0
カバー率	1.000
予測率	0.850
感度	0.849
特異度	0.851

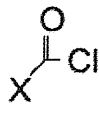
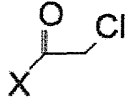
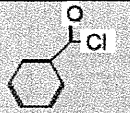
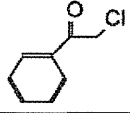
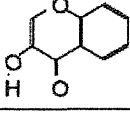
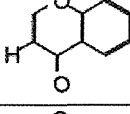
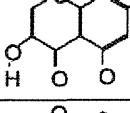
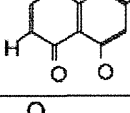
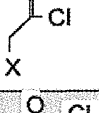
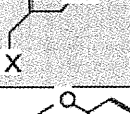
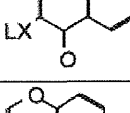
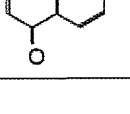
表5：重複した予測値による結果

	90%	TOP1	TOP3	TOP5	4つの結果が一致	3つの結果が一致
TN	95	123	121	120	77	115
TP	59	63	61	62	48	59
FN	13	10	12	11	2	6
FP	44	18	20	21	4	11
TOTAL	211	214	214	214	131	191
NG= FN+FP	57	28	32	32	6	17
未確定	3	0	0	0	83	23
カバー率	0.986	1.000	1.000	1.000	0.612	0.893
予測率	0.730	0.869	0.850	0.850	0.954	0.911
感度	0.819	0.863	0.836	0.849	0.960	0.908
特異度	0.683	0.872	0.858	0.851	0.951	0.913

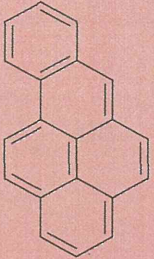
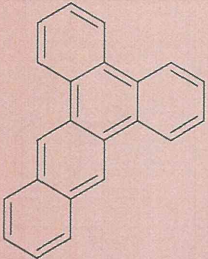
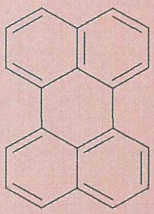
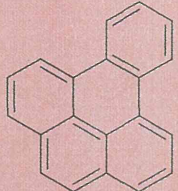
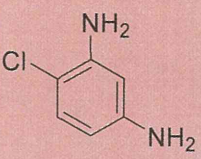
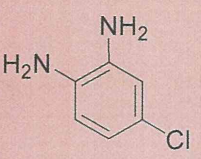
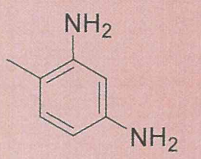
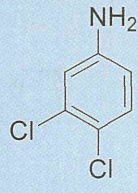
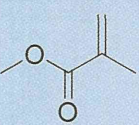
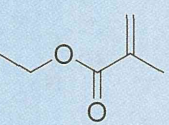
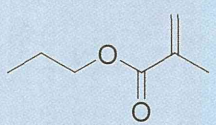
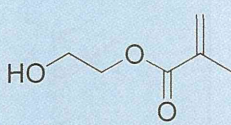
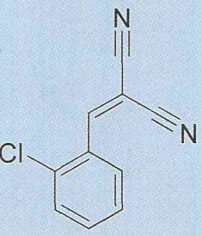
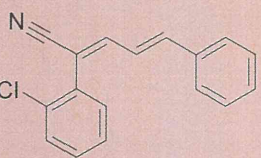
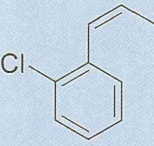
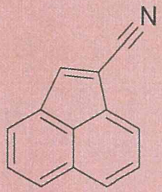
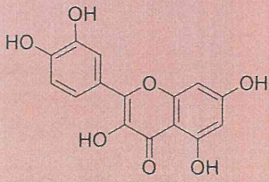
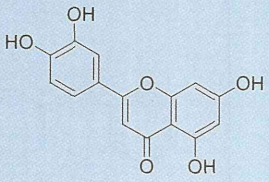
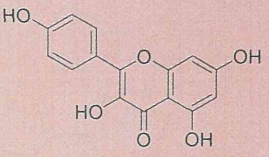
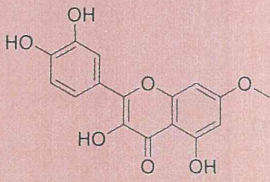
表6：類似学習母集団からの予測モデル作成 (QSAR)

	4つの結果が一致	3つの結果が一致	QSAR	4+QSAR	3+QSAR
TN	77	115	42	118	123
TP	48	59	14	62	64
FN	2	6	7	9	9
FP	4	11	13	17	17
TOTAL	131	191	76	206	213
NG= FN+FP	6	17	20	26	26
未確定	83	23	8	8	1
カバー率	0.612	0.893	0.905	0.963	0.995
予測率	0.954	0.911	0.737	0.874	0.878
感度	0.960	0.908	0.667	0.873	0.877
特異度	0.951	0.913	0.764	0.874	0.879

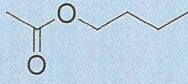
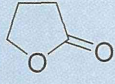
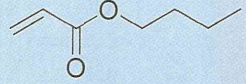
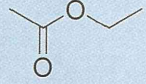
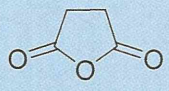
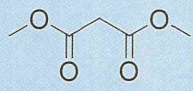
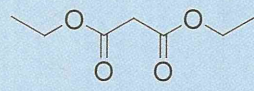
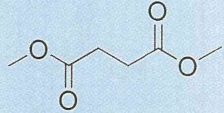
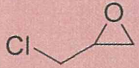


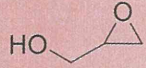


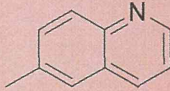
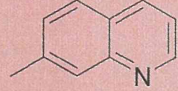
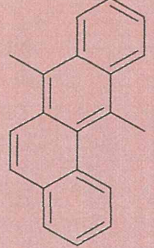
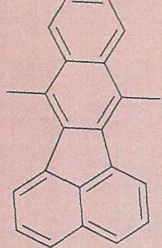
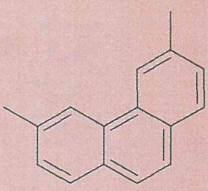
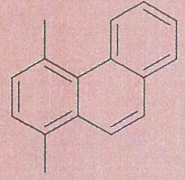
表 7 : 共通部分構造の全体母集団での出現頻度

Structure	SUBID	TOTAL	POS	NEG
	s3554-p	28	22	6
	s303-n	94	69	25
	s3554-p2	16	15	1
	s303-n2	4	2	2
	s54-p4	7	4	3
	s398-n4	17	6	11
	s54-p5	6	4	2
	s398-n5	10	4	6
	s3554-p1	7	5	2
	s303-n1	9	8	1
	p54-LX	9	6	3
	p54-null	17	6	11

別表 1 : 最も構造が類似した 3 つの化合物の Ames 試験結果

番号	化合物	最も類似した構造	2 番目に類似した構造	3 番目に類似した構造
1	 positive			
2				 negative
3				
4				
5				

6				
7				
8				
9				
10				
11				

12				
13				
14				
15				
16				
17	