

代表的構造－活性相関手法

1. Hansch-Fujita 法
2. Free-Wilson 法
3. CoMFA法
4. SARDE法
5. AIによる手法
6. パターン認識法
7. LEAD RECONSTRUCTION 法
8. LEAD RETRIEVAL法
9. 3次元化合物検索
10. ドラグレセプター理論によるアプローチ
 - ・レセプターマッピング
 - ・レセプターフィッティング
11. 類似度算出によるアプローチ

Hansch-Fujita 法

基本的特徴

薬物の生体内における挙動をシミュレートするアプローチ

- ・ 膜透過性 $\text{Log } P$
- ・ レセプターサイトでの反応性 $\text{ハメット } \sigma$
- ・ レセプターサイトでの立体的要因 $\text{Taft}' S E s$
 STERIMOL

基本とする薬理効果の種類

- ・ 全構造効果（化合物全体により定まる効果）
- ・ 置換基効果（置換基の種類により定まる効果）

実行上での手続き

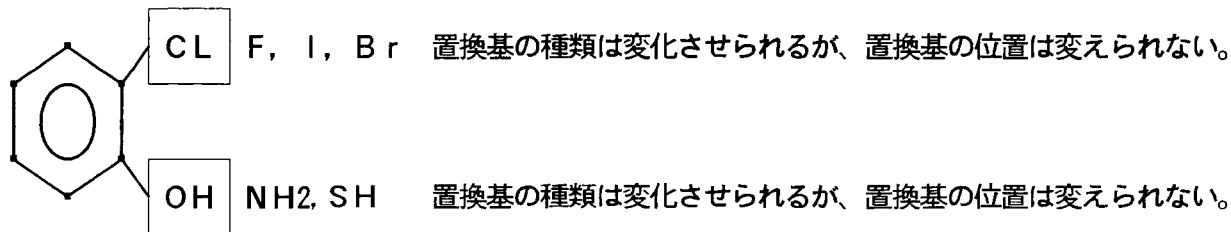
- ① 種々パラメータの收拾（Hansch-Fujita パラメータ）
- ② パラメータを用いた線形及び非線形重回帰の実行
- ③ 回帰式の検討及び新規薬物の設計

手法上での特徴／制限事項等

特徴：定量的構造－活性相関を実現する。

- ・ 連続データを扱うことが可能。

制限：Hansch-Fujita 法で扱う化合物はその基本骨格は置換基パターンも含めて厳しく制限される。置換位置は動かすことは出来ないレベルでの制限。



解析結果

解析精度：定量的レベル

要因解析：定量的レベルで解析可能

構造—活性相関事例 (1 — 1)

□ Hansch-Fujita 法

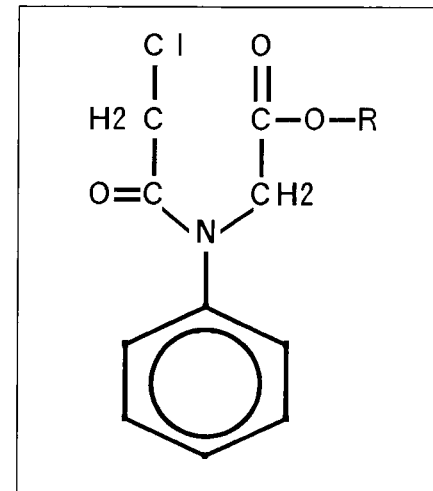
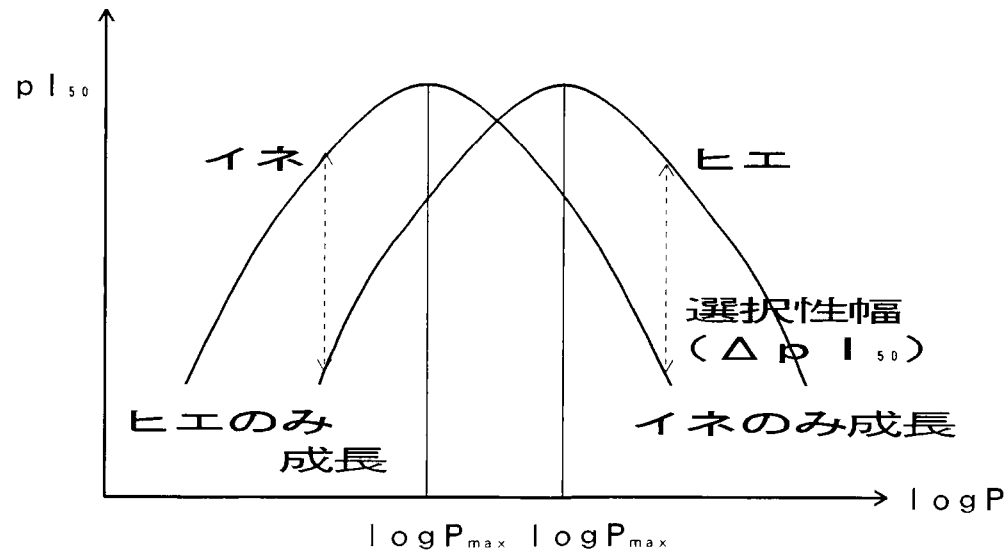
植物生育抑制物質

$$\text{イネ } pI_{50} = -0.28 (\log P)^2 + 1.47 \log P - 1.95 \sigma^* + 4.00$$

n = 22, s = 0.36, r = 0.85

$$\text{ヒエ } pI_{50} = -0.29 (\log P)^2 + 1.88 \log P + 1.00 \sigma^* + 2.89$$

n = 22, s = 0.38, r = 0.85



Log P 値による選択生育の実現

Free-Wilson 法

基本的特徴

Hansch-Fujita 法におけるパラメータのうち、ダミーパラメータを用いるアプローチ

*ダミーパラメータ：0 / 1 で表現される数値データ

基本とする薬理効果の種類

・置換基効果（置換基の種類により定まる効果）

実行上での手続き

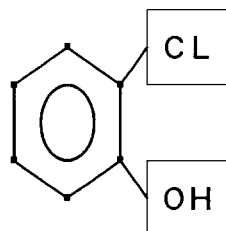
- ①種々パラメータの收拾（ダミーパラメータおよびHansch-Fujitaパラメータ）
- ②パラメータを用いた線形及び非線形重回帰の実行
- ③回帰式の検討及び新規薬物の設計

手法上での特徴／制限事項等

特徴：定量的構造－活性相関を実現する。

・連続データを扱うことが可能。

制限：Hansch-Fujita 法で扱う化合物はその基本骨格は置換基パターンも含めて厳しく制限される。置換位置は動かすことは出来ない。



0 / 1 0 ; H, 1 ; CL 置換基の種類が決定され、0の時は水素（H）が、1の時はCLが結合していることを示す。

*ダミーパラメータは置換基が存在するか、否かの情報を0 / 1で代表するものである。

0 / 1 0 ; H, 1 ; OH 置換基の種類が決定され、0の時は水素（H）が、1の時はOHが結合していることを示す。

解析結果

解析精度：定量的レベル

要因解析：定量的レベルで解析可能

構造—活性相関事例 (1 - 2)

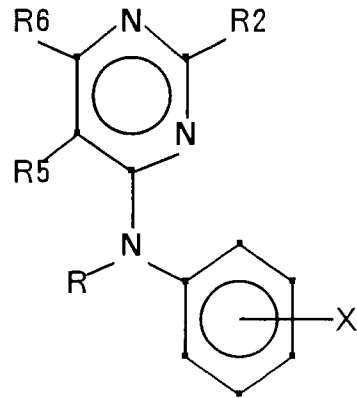
□ Free-Wilson 法

抗うつ活性 4-anilinopyrimidine誘導体

$$\ln(A/100-A) = -1.38\pi_5 + 1.51\sigma_P - 1.40I_{-1} + 1.37I_{-4} + 2.39I_{-5} - 0.76I_{-6} + 1.50I_{-7} - 0.68$$

$n = 59 \quad r = 0.837 \quad s = 0.81$

- π_5 : R5 の π 値
 σ_P : R5 の σ 値 (電子吸引性パラメータ)
 I_{-1} : aniline 部分のベンゼン環に置換基が存在する時 1
 I_{-4} : aniline 部分のパラ位にハロゲンが置換している時 1
 I_{-5} : aniline 部分のパラ位にアルキルが置換している時 1
 I_{-6} : R5 がハロゲンの時 1
 I_{-7} : R5 と R6 が (CH₂)₃ を介して閉環している時 1



- R : H, Me
 R2 : H, NH₂
 R5 : Me, Et, Pr, F, Cl
 R6 : H, Me, Et, Cl
 R5~R6 : (CH₂)_n, n = 3, 4
 X : o⁻, m⁻, p⁻ アルキル, ハロゲン, アルコキシ, CF₃

C o M F A アプローチ

基本的特徴

ドラグレセプターアプローチを基本とし、そのアプローチに重回帰手法による定量的、且つ客観的な解析手法を可能とした。

基本とする薬理効果の種類

- “場” の効果（化合物の周囲の状況により定まる効果）

実行上での手続き

- ① 3次元構造式の決定／算出
- ② 化合物同士の重ね合わせ
- ③ 格子化及び格子パラメータの算出
- ④ PLS法の実行、及び最終コンポーネント決定
- ⑤ コンポーネントデータを用いた線形および非線形重回帰の実行
- ⑥ グラフィックによる場の効果の確認

手法上での特徴／制限事項等

特徴：定性（一部定量）的構造－活性相関を実現する。

- * 手法的には重回帰手法を用いて連続データを扱うので定量的アプローチに分類されるが、場の効果の項は次元圧縮操作の為に定性的な解析となる。
- 連続データを扱うことが可能。

制限：基本骨格はHansch-Fujita 法程厳しくはない。手続き上、3次元化合物の重ね合わせが妥当な範囲で実行できる程度の構造的な自由度がある。（置換位置や基本骨格の変化は許容されるが、大幅な自由度はない）

解析結果

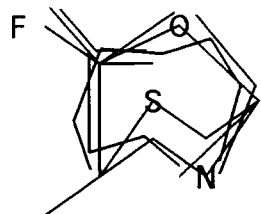
解析精度：定性的レベル（一部定量的解析可能）

□ CoMFAアプローチ

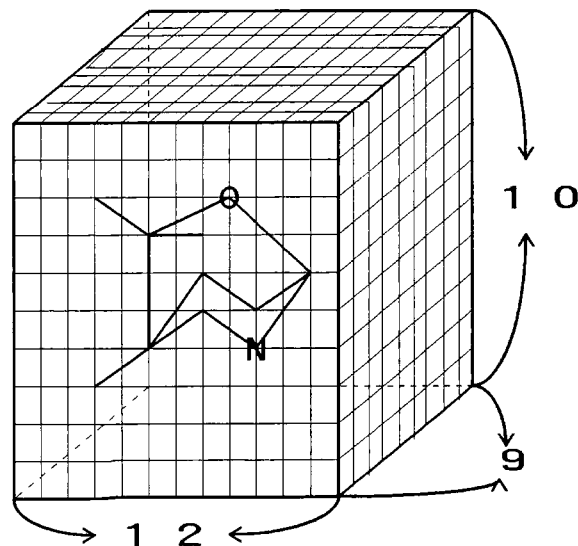
“場”の効果を基本とした3D-QSARアプローチである。
 “場”として化合物の部分領域の電子的および立体的効果が検討されている。

統計的手法：PLSおよび線形・非線形重回帰

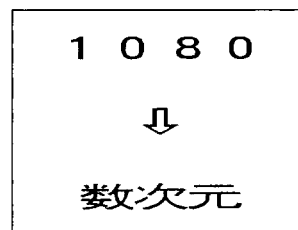
①化合物の重ね合わせ



②化合物周辺空間のキューブによる分割



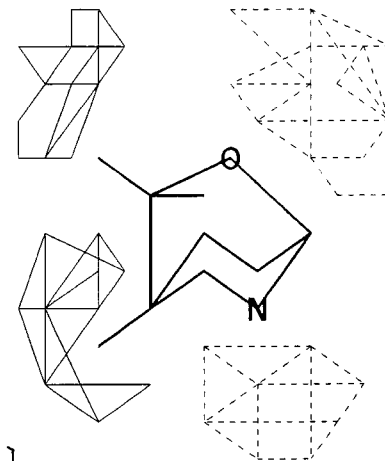
③PLSによる変換及び、次元数減少



④線形重回帰の実行

$$Y = [\text{Electrostatic \& Steric}]$$

⑤等高線図表示



SARDE

基本的特徴

ドラグレセプターアプローチを基本とし、そのアプローチに重回帰手法による定量的、且つ客観的な解析手法を可能とした。

基本とする薬理効果の種類

- ・ “領域” の効果（化合物内の定められた領域により発現される効果）

実行上での手続き

- ① 3次元構造式の決定／算出
- ② 化合物同士の重ね合わせ
- ③ 格子化及びドメインパラメータの算出
- ④ ドメインパラメータを用いた線形および非線形重回帰の実行
- ⑤ ドメインパラメータ情報を基本としたリガンドモデルの構築

手法上での特徴／制限事項等

特徴：定量的構造－活性相関を実現する。

* 領域効果を数値データ化するドメインパラメータを用いることで重回帰手法の利用を可能としている。

- ・ 連続データを扱うことが可能。
- ・ リガンドモデルの直接的な構築が可能。

制限：基本骨格はHansch-Fujita 法程厳しくはない。手続き上、3次元化合物の重ね合わせが妥当な範囲で実行できる程度の構造的な自由度がある。
(置換位置や基本骨格の変化は許容されるが、大幅な自由度はない)

解析結果

解析精度：定量的レベル

□ SARDEアプローチ

3D-QSARアプローチである。

統計・パターン認識による解析とドラグレセプター理論によるアプローチをの長所を互いに結び付けたアプローチ。

“領域効果”を基本としたアプローチである。

斜面によるドメイン化

自由曲線によるドメイン化

ブロックによるドメイン化

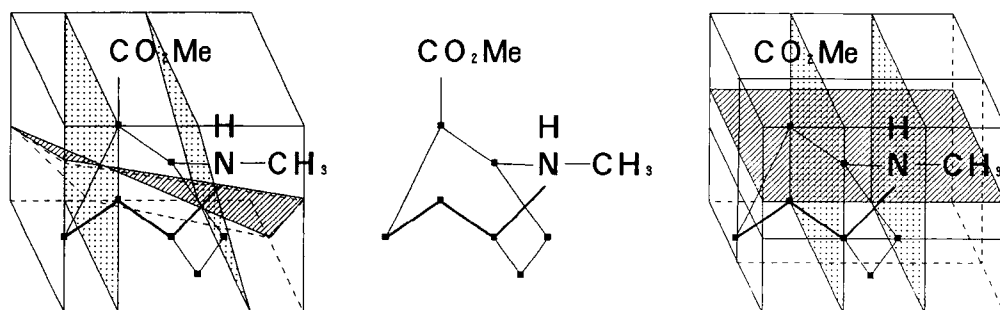
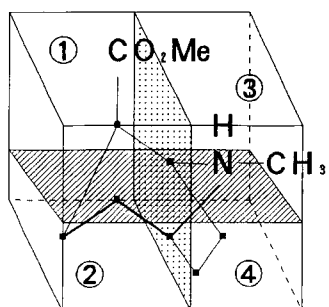


図1. ドメインパラメータの領域化例



領域① 親水性	$\begin{array}{c} \text{CO}_2\text{Me} \\ \\ \text{C} \end{array}$	領域③ 立体的特性	$\begin{array}{c} \text{C} \\ \diagdown \quad \diagup \\ \text{C} \end{array}$
領域② 立体的特性	$\begin{array}{c} \text{C} \\ \diagdown \quad \diagup \\ \text{C} \quad \text{C} \end{array}$	領域④ 電子的特性	+ N-Me

図2. 基本3次元化合物構造式とドメイン化によるドメイン構造式とその特性例

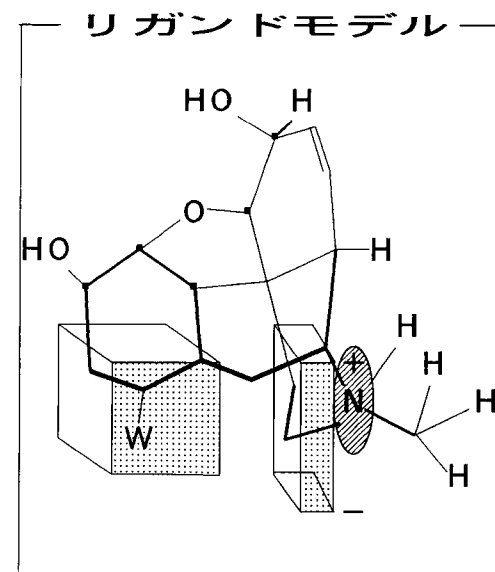


図3. リガンドモデルの構築

A I 利用による構造－活性相関

基本的特徴

人間が行った過去の経験やノウハウ等を盛り込んだ形で構造－活性相関を行う。
このノウハウ情報の種類や扱い方により様々なアプローチが可能である。

基本とする薬理効果の種類

特に無し（システムの設計方法により変わる）

実行上での手続き

- ① ノウハウのシステムへの入力（システムによりノウハウの種類が異なる）
- ② 解析目的情報をシステムに与える
- ③ 実行結果の出力及びその解析

手法上での特徴／制限事項等

特徴：人間のノウハウをシステムに盛り込むので、ノウハウ情報をシステムに盛り込む過程での技術的問題の克服が必要。

制限：

解析結果

解析精度：システムの仕様により異なる。

□ その他のアプローチ

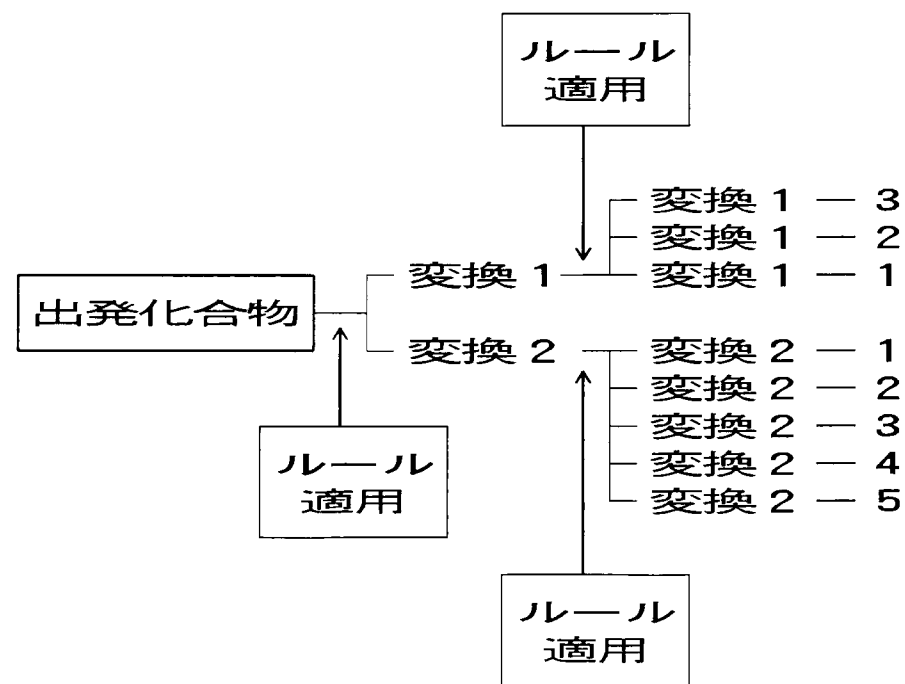
A I (人工知能) 技術を用いたアプローチ

過去に展開された開発履歴 (化合物構造式の展開履歴) をシステムに入力し、その履歴情報を用いて新たに入力された化合物の構造展開を行う。

化合物変換ルールデータベース

入力化合物への変換ルールの適用

薬理 A	薬理 B	薬理 X
変換事例 1 $A \Rightarrow B$ 活性 向上 毒性 低下	変換事例 1 $A \Rightarrow B$ 活性 向上 毒性 低下	変換事 $A \Rightarrow B$ 活性 毒性
変換事例 2 $B \Rightarrow C$ 活性 向上 毒性 強化	変換事例 2 $B \Rightarrow C$ 活性 向上 毒性 強化	変換事 $B \Rightarrow C$ 活性 毒性
.....		
変換事例 2 $Y \Rightarrow Z$ 活性 向上 毒性 強化	変換事例 2 $Y \Rightarrow Z$ 活性 向上 毒性 強化	変換事 $Y \Rightarrow Z$ 活性 毒性



構造－活性相関で何を狙うのか？

□A I（人工知能）の技術を利用した構造－活性相関

定義：特に無し

アプローチ：人工知能の技術を利用し、人間が持っているノウハウを何らかの形でシステムに導入して解析を行う。

例 1：EMIL システム

過去に行われた化合物の変換パターンを知識として蓄え、その知識を用いて新たな化合物の構造式変換を行う。

特徴：過去に試行錯誤で行われた貴重な構造式変換ノウハウをシステムが蓄え、その知識を用いて新たな構造式変換を行う。

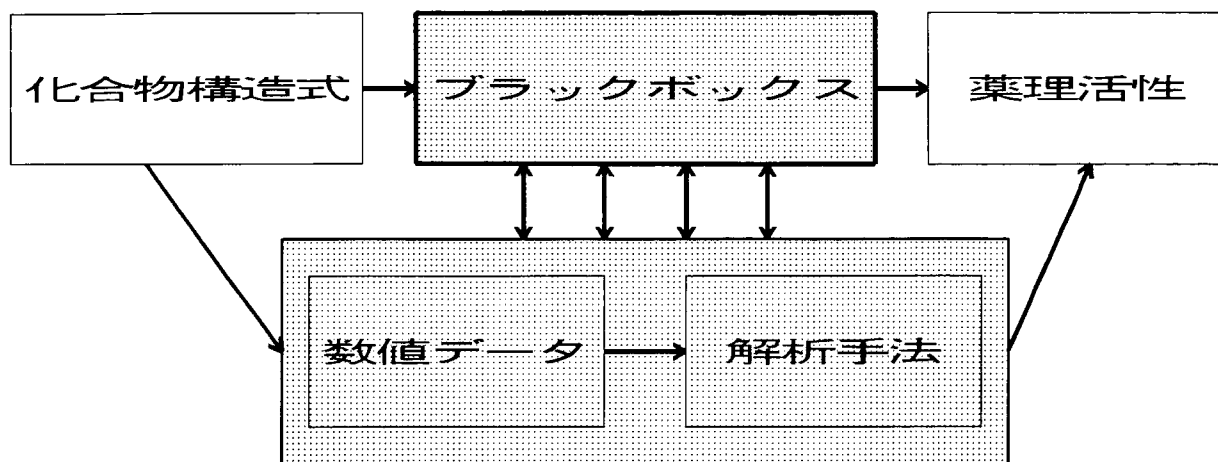
シリーズの異なる（薬理活性／対象分野（医薬／農薬））化合物への適用により新規のLEAD候補化合物を創出出来る。

例 2：KARMA システム

Hansch-Fujita 法で利用される回帰式を様々な検索条件から探し出し、その取り出されたデータに関連する情報等もインヘリタンス（獲得情報）として自動的に呼び出す。

パターン認識法

基本原理図



基本的特徴

化合物に関する情報をパターン認識手法を用いて様々な観点から解析する。

基本とする薬理効果の種類

特に無し

手法上での特徴／制限事項等

特徴：

制限：構造上での制限は緩い。アプローチによっては解析母集団として構造式の全く異なる化合物群を用いても構わない。

解析結果

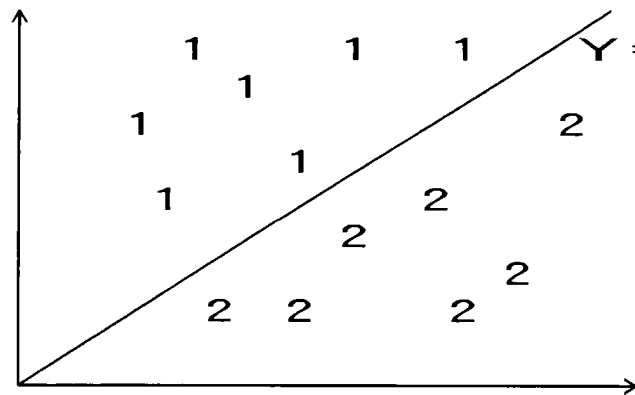
解析精度：定性的レベル（クラス分類等）

□ パターン認識法

種々のパターン認識手法を駆使しながら活性に関する要因解析や活性予測等を行う。

解析手法：パターン認識手法全般

クラス分類（2クラス）

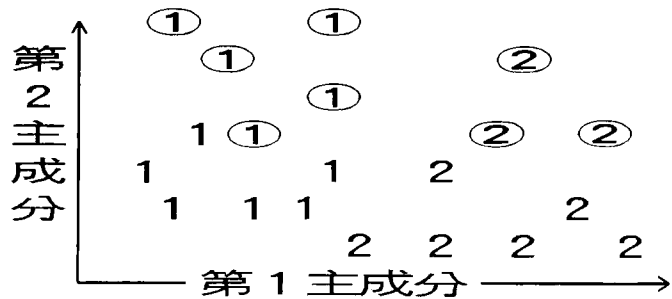


$$Y = a x_1 + b x_2 + \dots + n x_n + C$$

- 1 : 活性化化合物
- 2 : 不活性化化合物

- 係数（ $a \sim n$ ）が、
 正の時 \Rightarrow 活性クラスに寄与する要因
 負の時 \Rightarrow 不活性クラスに寄与する要因
- 活性予測が可能

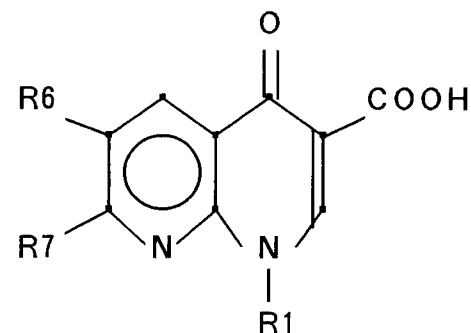
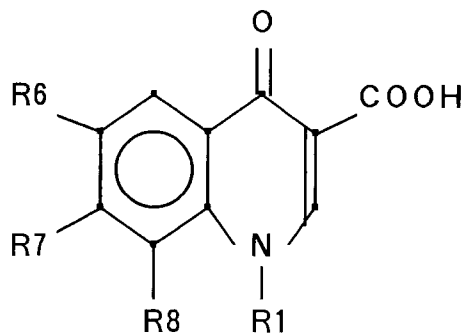
主成分分析



- 第1主成分
 活性の有無を識別するのに重要な情報
- 第2主成分
 薬物群を1 2 と①②とに分離するのに重要な情報

構造—活性相関事例

□キノロン系化合物による抗菌活性に関するパターン認識による解析



クラス分類結果 (2クラス)

	クラス1	クラス2	トータル
線形学習機械法	100.0	100.0	100.0
線形判別分析 (最少二乗法)	100.0	100.0	100.0
線形判別分析 (シンプレックス法)	100.0	100.0	100.0
Bayes線形判別分析	96.3	93.1	94.6
Bayes非線形判別分析	88.9	96.6	92.9
最近隣法 K=1	85.2	86.2	85.7
K=3	81.5	79.3	80.4

クラス1 (活性) = 27 クラス2 (不活性) = 29 トータル = 56化合物

クラス	係数	説明	寄与クラス
+	0.1876	Total Number of Paths/Number of Atoms	クラス1
-	0.6912	Number of Nitrogen Atoms in Structure	クラス2
+	0.3183	Number of Basis Rings in Structure	クラス1
-	0.3004	Minimum Nucleophilic Superdelocalizability	クラス2
+	0.5428	Tortional Strain Energy of Molecule	クラス1
+	0.0302	Constant Term	—

種々のクラス分類手法による分類率

- 解析に用いた判別関数のよしあし
- 活性予測に対する信頼性の指標

- 係数が正⇒活性に寄与する要因
この値を大きくすると活性が強まる
- 係数が負⇒不活性に寄与する要因
この値を大きくすると活性が弱まる

LEAD RECONSTRUCTION

基本的特徴

パターン認識による解析結果の情報（部分構造記述子）を受け、その情報に基づいて新規LEAD候補化合物を創出する。

基本とする薬理効果の種類

・ 特になし

実行上での手続き

- ① 部分構造記述子を用いたパターン認識による解析実行
- ② 最終判別式による、活性に寄与する部分構造の決定
- ③ 基本構造と部分構造の組み合わせによる新規化合物の創出
- ④ 新規創出化合物の活性予測
- ⑤ 活性有りの化合物を最終候補化合物とする

手法上での特徴／制限事項等

特徴：新たなLEAD化合物を設計することを目的としたアプローチ。

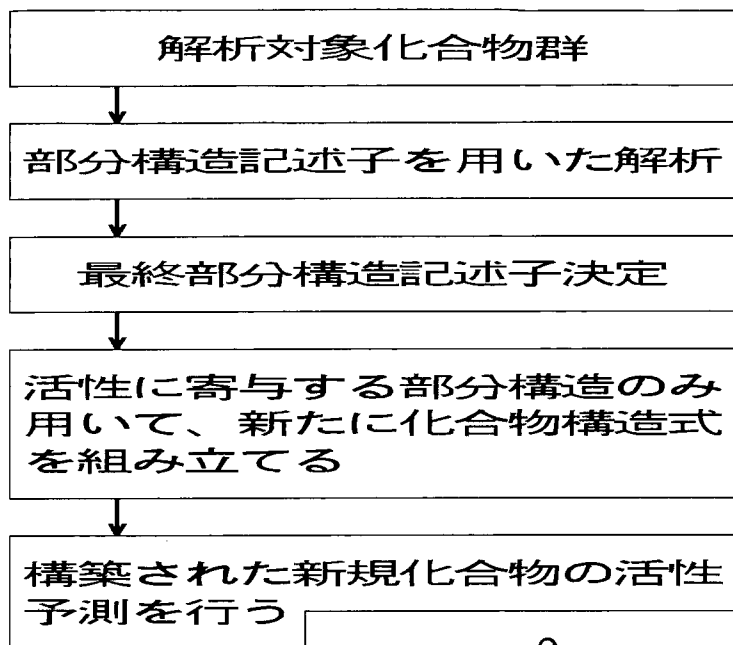
制限：解析対象化合物群は同一の基本構造を持つことが必要である。

解析結果

新規LEAD候補化合物の創出が行われる。 但し、基本構造は保たれる。

□ LEAD RE-CONSTRUCTION

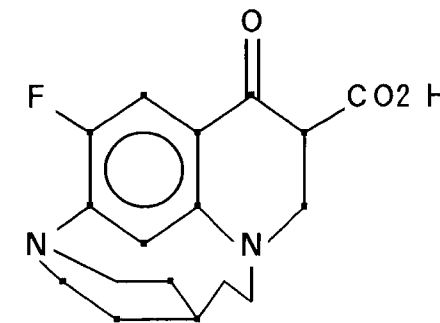
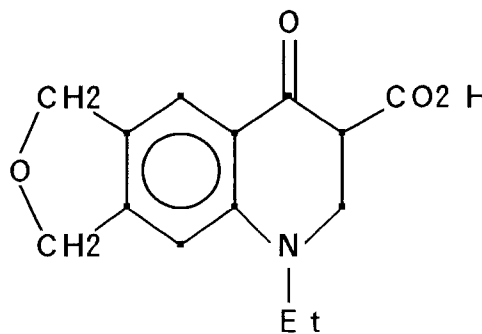
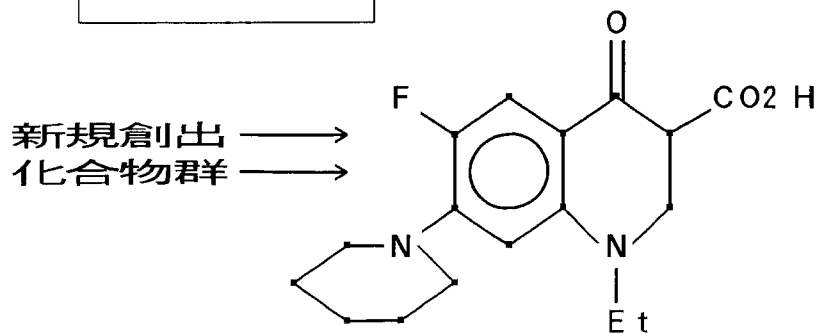
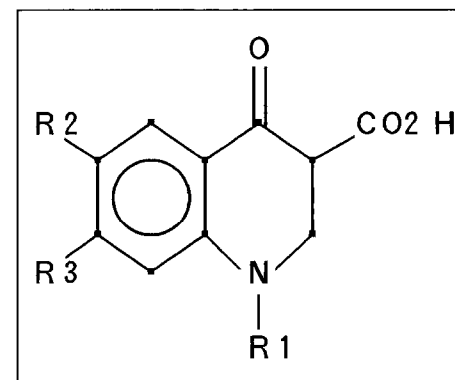
パターン認識で要因解析を行い、その結果得られた記述子（数値データ）情報を用いて化合物データベース中から取り出された化合物の活性予測を行う。



特徴：LEAD EVOLUTIONN 的アプローチ

1. 部分構造記述子情報を利用した構造創出
2. 基本構造を保持した新規化合物のデザイン

解析母集団化合物



LEAD RETRIEVAL

基本的特徴

パターン認識による解析結果（最終記述子：構造に直接関係した記述子は除く）を用いて、化合物データベース中より新規LEAD候補化合物群を探し出す。

基本とする薬理効果の種類 特に無し

実行上での手続き

- ①記述子（部分構造を除く）を用いたパターン認識による解析実行
- ②最終判別式による、活性に寄与する記述子の決定
- ③化合物データベース中の化合物を記述子情報へと変換する
- ④創出された記述子を用いて、活性予測をする
- ⑤活性有りの化合物を最終候補化合物とする

手法上での特徴／制限事項等

特徴：新規LEAD化合物の発見。
LEAD GENERATION を目指す。

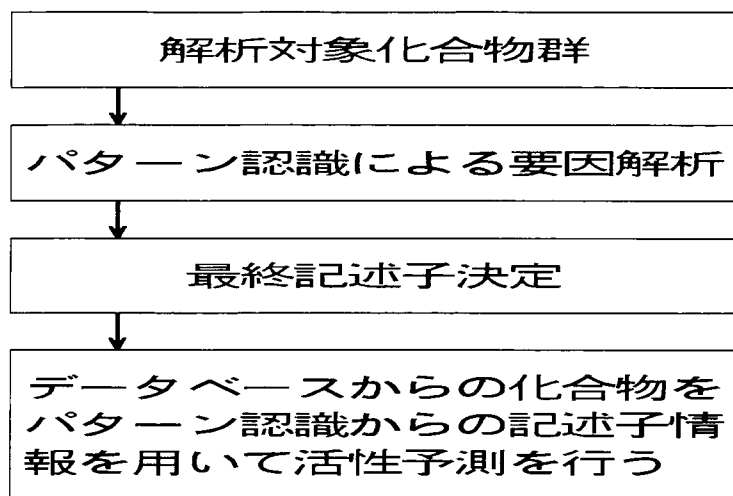
制限：特にない。パターン認識による解析が行えるならば問題はない。
従って、化合物構造式が特に制限されることはない。

解析結果

新規LEAD候補化合物（LEAD GENERATION レベル）の化合物データベースからの取り出し。

□ LEAD RETRIEVAL

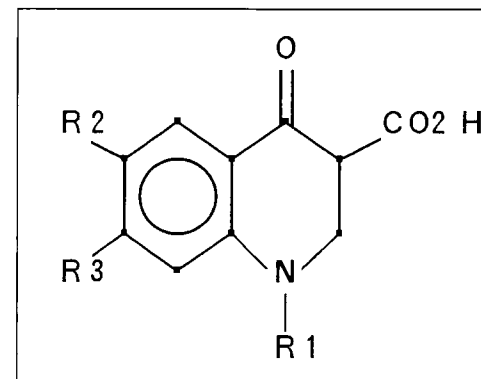
パターン認識で要因解析を行い、その結果得られた記述子（数値データ）情報を用いて化合物データベース中から取り出された化合物の活性予測を行う。



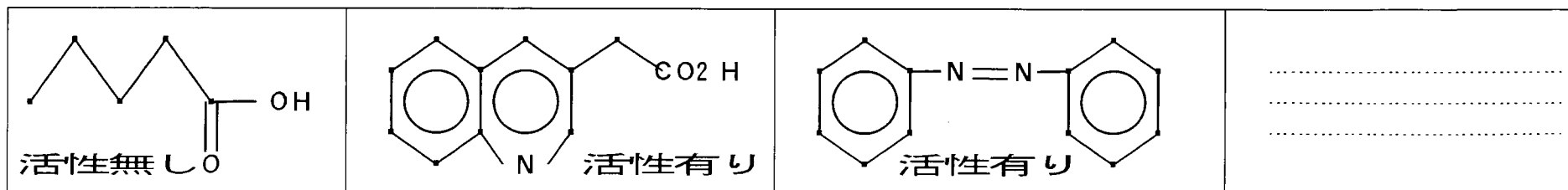
特徴：LEAD GENERATION に最も近いアプローチ

1. 記述子情報を基本として活性予測する

基本構造
(母集団化合物)



データベースより取り出された化合物



3次元化合物検索

基本的特徴

化合物の3次元の特徴を基本とし、化合物データベースより検索する

基本とする薬理効果の種類

特に制限なし

実行上での手続き

3次元化合物検索用の特殊な検索キーが必要となる。
化合物データベースにも3次元情報が必要となる。

手法上での特徴／制限事項等

特徴：3次元化合物検索を実現する。

- ・ 化合物の立体的な情報に基づいた化合物検索
- ・ LEAD GENERATION や LEAD EVOLUTION 的なアプローチが可能

限界：化合物の活性構造と安定構造とで大きな差異がある時は、検索精度は落ちる

検索結果

検索精度：定性的レベル

構造—活性相関事例

□ 3次元化合物検索によるアプローチ

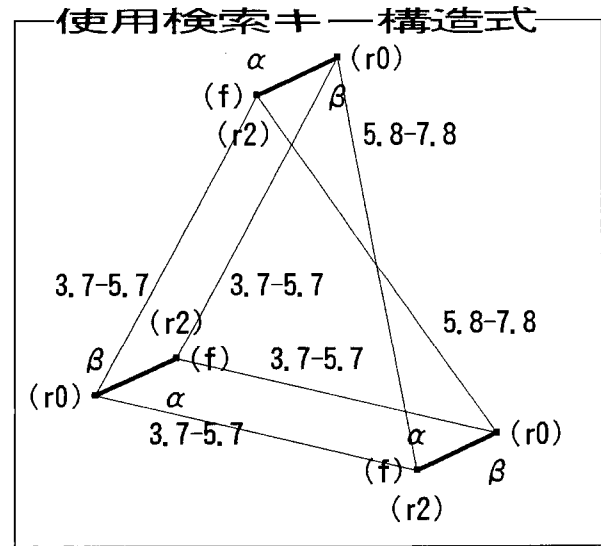
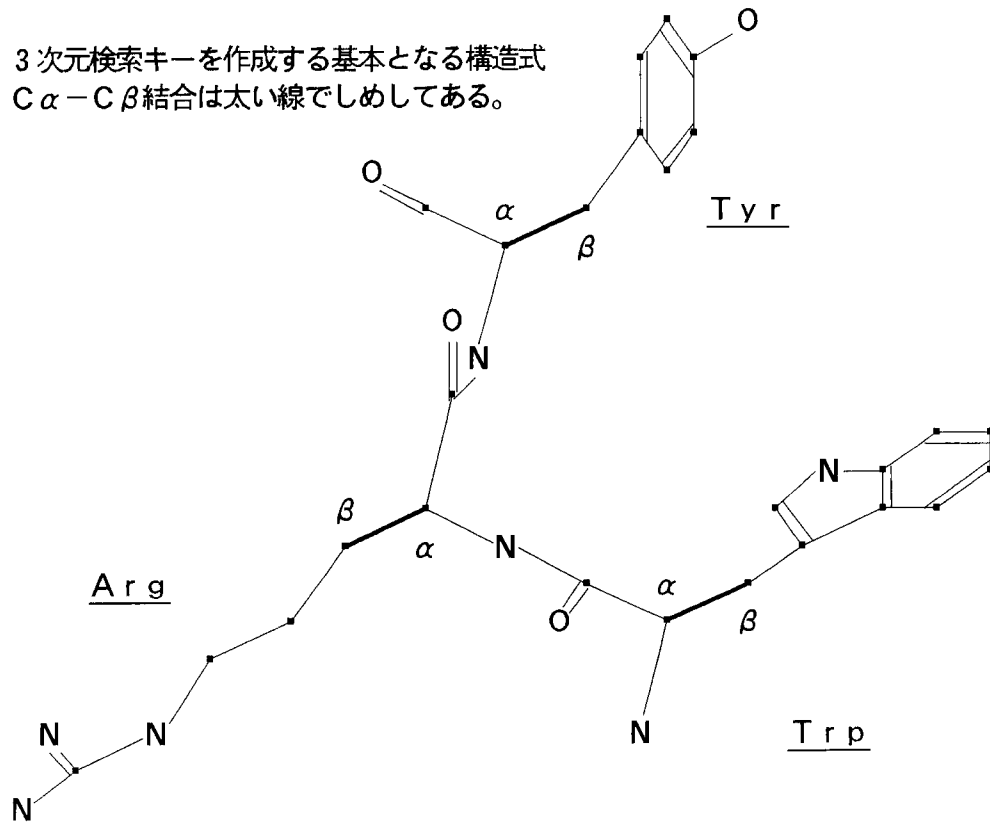
3次元化合物検索機能を用いた化合物検索による構造—活性相関

- ・ 構造—活性相関を意識する時、3次元化合物検索の検索キー構造式は受容体モデルから得られる情報を基本にして構築されることが多い。

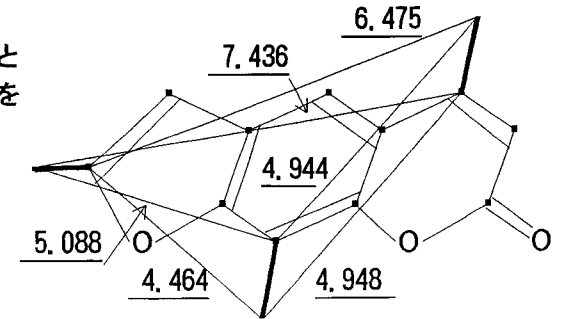
□ α -アミラーゼ阻害作用

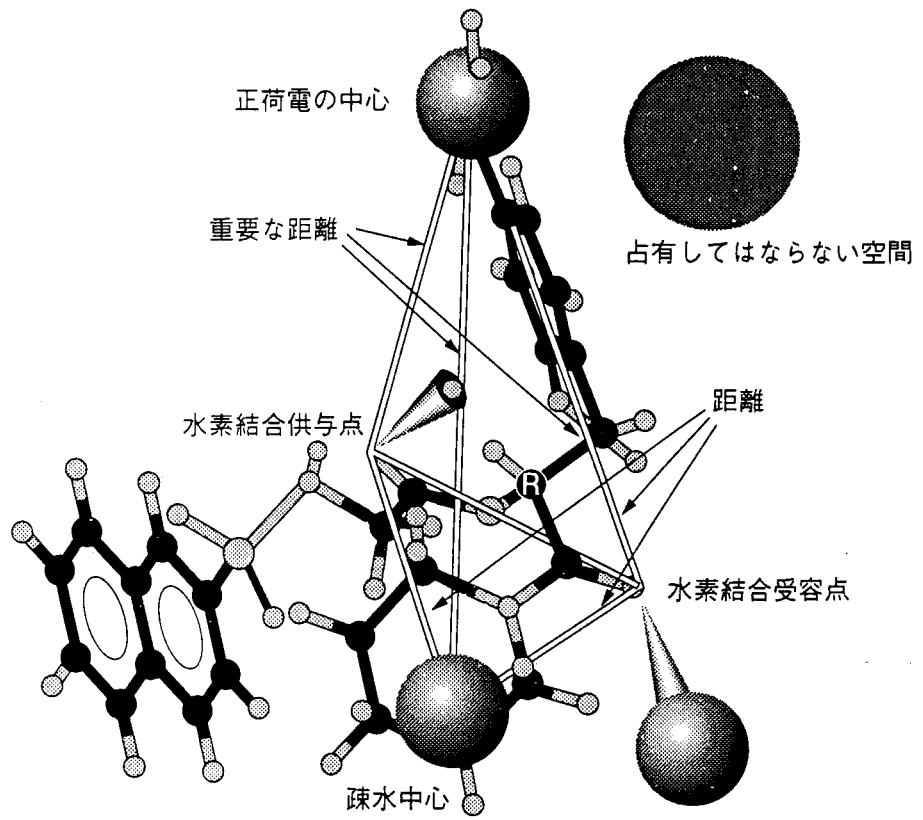
[基本化合物構造式]

3次元検索キーを作成する基本となる構造式
 $C\alpha-C\beta$ 結合は太い線でしめしてある。



最終ヒット化合物と
 3次元検索キーとを
 重ね合わせ表示





ドラグレセプター理論を基本としたグラフィック主体のアプローチ

基本的特徴

- ① KEY & LOCK の理論を実践する
- ② 視覚等に訴えたアプローチが中心で、最終的な判断は研究者のスキルに大きく依存する部分が多くなる

基本とする薬理効果の種類

特に無し

実行上での手続き

- ① 化合物の 3 次元構造式創出 (データベースより求めても良い)
- ② レセプターサイトの情報
 - 有り⇒リガンド化合物とのフィッティング実行
 - 無し⇒リガンド化合物群からレセプターサイトの情報を取り出す
レセプターマッピングの実践

手法上での特徴／制限事項等

特徴：定性的構造－活性相関を実現する。

限界：化合物の活性構造と安定構造とで大きな差異がある時は、検索精度は落ちる

制限：構造的な制限は緩い。

解析結果

解析精度：定性的レベル

要因解析：定性的レベルで解析可能

レセプターマッピング / フィッティング

基本的特徴

レセプターサイトの形状や、様々な特性を推測するアプローチ

* レセプターサイトの形状を明確にすることで、そのレセプターサイトにフィットする化合物の設計が現実のものとなる。

レセプターサイトの形状がわかっている時 : フィッティング

レセプターサイトの形状が不明な時 : レセプターマッピング

基本とする薬理効果の種類

- ・ ドラグレセプター理論

手法上での特徴 / 制限事項等

特徴 : 定性的構造 - 活性相関を実現する。

制限 : 同一レセプターサイトに作用することが保証される範囲。
Hansch-Fujita 法ほど構造制限は厳しくない。

解析結果

解析精度 : 定性的レベル

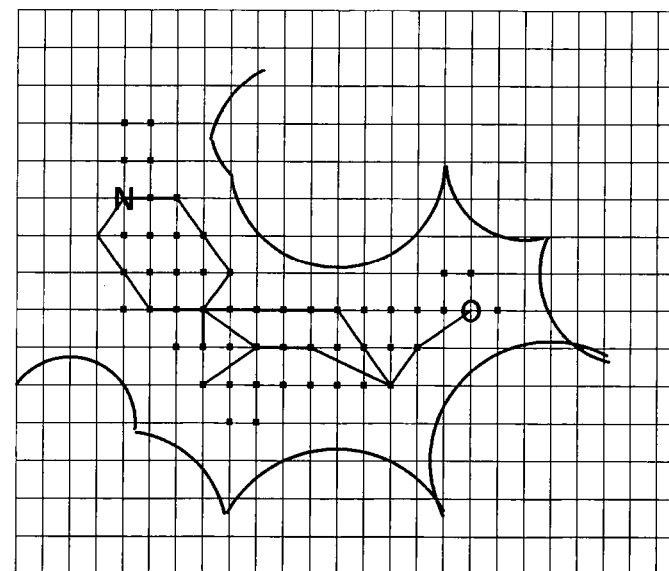
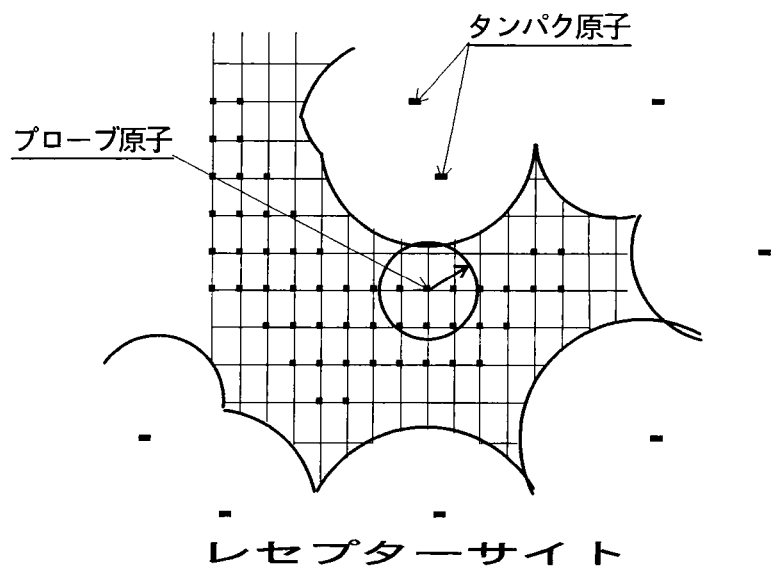
□ ドラッグレセプター理論を基本としたグラフィック主体のアプローチ

1. レセプターサイトの形状が既知の時

2. レセプターサイトの形状が未知の時

▪ フィッティングの実行

▪ フィッティングの実行



- * プローブ原子が存在する範囲のみにリガンド原子が存在する
- * セットされた時の反発エネルギーを算出し、その妥当性をチェックする

- * 総てのリガンドモデル（重ね合わせデータを用いる）について、格子点データを求める
- * 最終格子点モデルから、レセプターモデルを形成する

構造-活性相関で何を狙うのか？

□適用技術 | 類似度による構造-活性相関|

定義：『似た化合物は似た活性を示す』

応用 ①化合物の類似度を計りながら活性を維持／向上させて行く

②類似度測定による類似化合物検索によるLEAD候補化合物の取り出し

類似度計算：谷本の類似インデックス

$$S = \frac{\sum_{i,j} W_i W_j}{\sum_i (W_i)^2 + \sum_j (W_j)^2 + \sum_{i,j} W_i W_j}$$

類似度による構造－活性相関

基本的特徴

『似たものは似た活性を示す』という基本概念を実践するアプローチ

基本とする薬理効果の種類

特に制限無し

実行上での手続き

- ① 化合物と被検索化合物との類似度を算出する
- ② 類似度が高い順に化合物を取り出す

手法上での特徴／制限事項等

特徴：化合物構造式の類似性に着目したアプローチ

限界：基本とする化合物と構造的、或いは定義上で離れた化合物を取り出すことは不可能
類似度を算出するアプローチの違いにより結果が大きく左右される
構造的類似は測定出来ても、機能的類似を計るアプローチが出来ない

制限：基本骨格は最低限同じか、定義上の指定事項を満たすことが必要

解析結果

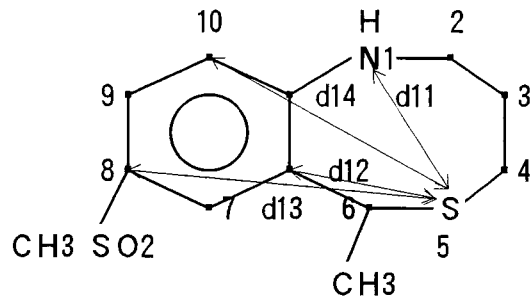
解析精度：定性的レベル

構造－活性相関事例

□ 類似度を基準とした構造－活性相関へのアプローチ

ベンゾチアジン (RS-2039) とジペプチドの類似性の評価

$$DS_i = \frac{2 |d_{1i} - d_{2i}|}{d_{1i} + d_{2i}}$$



d11 = 3.68 Å
 d12 = 2.74 Å
 d13 = 4.72 Å
 d14 = 4.73 Å

