

# 自己組織化マップを用いた薬物骨格構造と活性地図の作成 Structure-activity Map of Molecular Framework of Drugs using Self-Organizing Mapping

山根 勁  
Tsuyoshi Yamane

高橋 由雅  
Yoshimasa Takahashi

豊橋技術科学大学大学院 工学研究科 情報・智能工学専攻  
Department of Computer Science and Engineering, Toyohashi University of Technology

It is a very important to know the relationships between molecular framework of chemicals and their biological action in drug design and discovery. This paper describes an approach to visualization of chemical space of molecular frameworks with drug actions to do this. We employed Self-Organizing Mapping (SOM) method for making a map of NTGs (Non-Terminal vertex Graph) obtained from molecular graphs and the related drug actions. The NTG is defined as a graph that doesn't have any terminal vertex or any isolated vertex on itself. The topological fragment spectra (TFS) method reported by the authors was used for describing an NTG of a chemical structure into the numerical pattern representation. The computational trials with the NTGs of several types of drug molecules suggested that the approach is successfully applicable to visualize the relationships between the NTGs and drug actions.

## 1. はじめに

合理的な新薬開発研究を進める上では、薬物の骨格構造と活性との関係を知ることが極めて重要となる。こうした構造活性相関知識の獲得に向けた計算機援用技術に関する研究が活発に行われている。化合物の構造特徴を計算機上で扱う際には、離散数量化を行い構造特徴空間に変換するが、一般には高次元データ空間となることから、相互の関係を理解することが困難になってしまう。そこで本研究では自己組織化マップ(SOM)を用いて、多次元構造特徴空間の2次元写像を行うことで、薬物骨格構造の近接性と活性の関係を表す活性地図の作成とその応用について議論する。

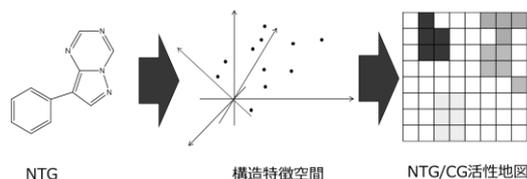


図1 本研究の流れ

## 2. Self-Organizing Map (SOM)

SOMとはKohonenによって提案された人工ニューラルネットワークアルゴリズムの一種であり、高次元データの近接性を維持したまま、低次元空間に非線形写像を行う教師なし学習の一種である。

SOMは入力ベクトルを受け取る入力層と、入力に対する重みベクトルをもつユニットが格子状に配置されている競合層からなる。入力データに対する最大類似度をもつユニット(BMU, Best Match Unit)が入力データの競合層上の写像位置となる。

## 3. 分子骨格と特徴表現

### 3.1 NTG(Non-Terminal vertex Graph)

NTG[Takahashi 04]とは、化学構造における環構造に注目し

た基本骨格で、次数1の頂点および孤立頂点を持たないグラフである。NTGには注目する情報によって5つの表現レベルが定義されており、表現レベル間には階層関係も存在する。図2にDiazepamの化学構造に対応した分子グラフから抽出したNTGの例と、そのNTGにおける各種表現レベルを示す。

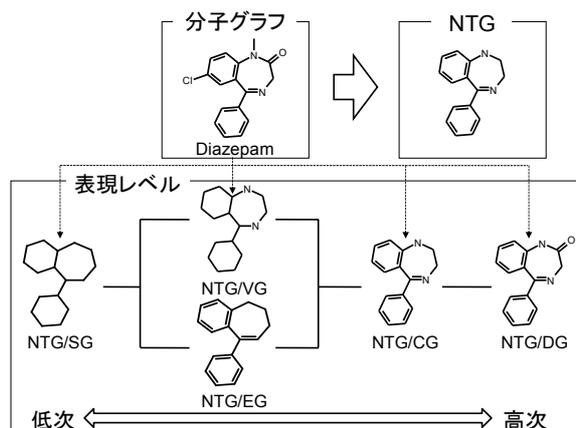


図2 NTGの5種類の表現レベル

### 3.2 TFS(Topological Fragment Spectra)

化学構造の類似性を定量的に比較、評価するためには構造情報を離散・数量化を行う必要がある。本研究ではNTGの構造特徴の記述にはTopological Fragment Spectra (TFS)法[Takahashi 98]を用いた。TFSは当研究室で考案された化学構造情報の定量的な記述手法である。その生成手順は、(1)対象とする化学構造式から可能なフラグメントをすべて列挙し、(2)列挙したそれぞれのフラグメントに対して数値的な特徴付けを行う。(3)その特徴付けの値と出現頻度のヒストグラムを生成する。このヒストグラムがTFSであり、これを多次元パターンベクトルとして用いることで、化学物質の構造特徴を数値的・定量的に表すことができる。本研究ではサイズ5までの部分構造を列挙し、特徴付けには各フラグメントの構造構成原子の質量数の和を用いた。

連絡先: 高橋由雅, 〒441-8580 愛知県豊橋市天伯町雲雀ヶ丘1-1 豊橋技術科学大学 情報・智能工学系, Tel: 0532-44-6878, taka@mis.cs.tut.ac.jp

#### 4. 化学構造特徴空間地図

本研究では, SOM を用いて化学構造特徴空間の近接性を反映した地図が作成できるか確かめるために, 化学構造特徴空間地図の作成を行った. データセットとして, 能登らが作成した薬化合物 NTG 辞書[能登 08]より単純グラフ同型な NTG/CG の数が上位 20 種の NTG/SG を用いた. SOM の学習には佐々木らによって開発された SOMSAR[佐々木 04]を使用した. マップサイズを 10x10 で地図を作成したものを図 3 に示す.

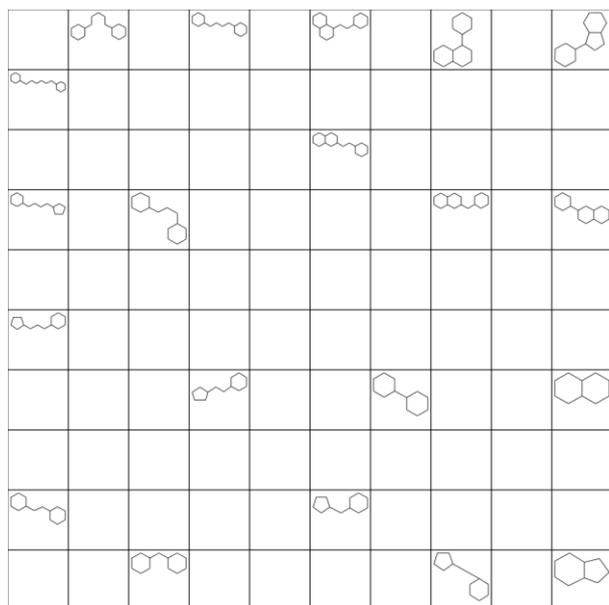


図 3 化学構造特徴空間地図

右上に三環式の構造がクラスタを形成している. また, 環と環の間を結ぶリンカの長さでクラスタが形成されており, 右下には縮合環のみの分子がクラスタを形成している.

このように, SOM によって NTG/SG のデータセットを写像することで化学構造の類似性を反映した地図を作成することが可能である.

#### 5. 薬物骨格構造-活性地図

ここでは化学構造特徴空間地図に薬理活性を関連付けることで薬理活性によるクラスタが形成されるか確認するため薬物骨格構造-活性地図の作成を行った.

データセットとしては, NTG 辞書に含まれる活性クラスの中から, 降圧剤, 抗がん剤, 抗炎症薬, 抗精神病薬の 4 つの薬理活性に注目し, これらの活性をもつ NTG/SG を基本骨格としてもつ NTG/CG を用いた. SOM の学習には前章と同じく SOMSAR を使用した. 図 4 に使用した基本骨格を示す. また, 図 5 にマップサイズ 20x20 で作成した薬物骨格構造-活性地図を示す.

結果として, 同じ基本骨格を持つ NTG/CG がクラスタを形成しており, さらに薬理活性によってもクラスタができていることがわかる.

図 5 において, 実線で示した領域が図 4(a)の骨格を持つ降圧剤活性クラスタ, 点線で示した領域が図 4(b)の骨格を持つ抗がん剤活性クラスタ, 一点鎖線で示した領域が図 4(c)の骨格を持つ抗炎症薬活性クラスタ, 破線で示した領域が図 4(d)の骨格を持つ抗精神病薬活性クラスタとなっている.

また, 降圧剤活性クラスタは, 大部分の化合物が抗精神病薬活性を持っている. これは基本骨格である図 4(a)と図 4(d)が類似しているためだと考えられる.

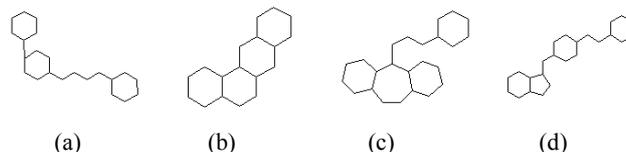


図 4 使用した基本骨格

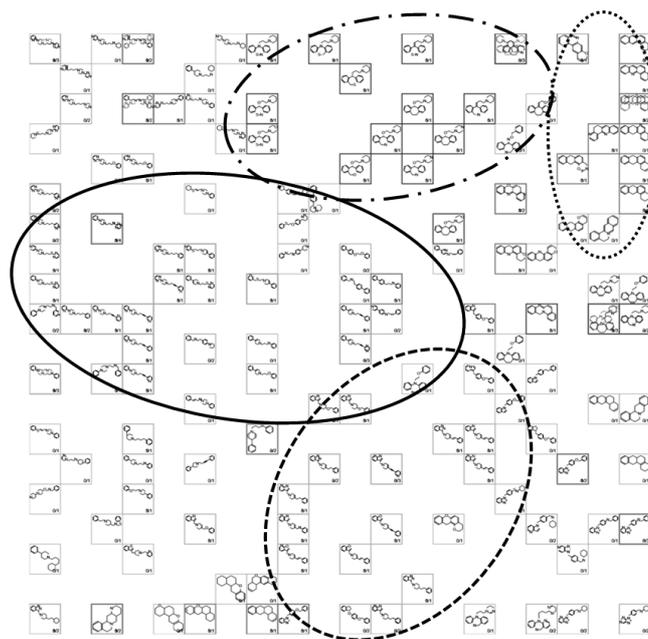


図 5 薬物骨格構造-活性地図

#### 6. まとめ

上記の結果は, SOM により化学構造特徴空間の可視化が可能であるとともに, 薬理活性を関連付けることで化学構造の類似性と薬理活性との関係を視覚的に明らかにする活性地図が作成できることを示している. 今後は, より多くの種類の活性クラスタを持つ規模の大きな活性地図の作成や, 活性地図に関連付ける活性の種類を変えることで, 活性地図の有用性を高めたい. また, 活性予測への応用の検討も行う予定である.

#### 参考文献

- [Takahashi 98] Y. Takahashi, H. Ohoka, Y. Ishiyama, Structural Similarity Analysis Based on Topological Fragment Spectra, *Advances in Molecular Similarity*, 2, 93-104(1998)
- [コホネン 96] T.コホネン著, 得高平蔵他訳, 自己組織化マップ, シュプリンガー・フェアラーク東京 (1996)
- [能登 08] 能登上総, 高橋由雅, 2008 年度人工知能学会全国大会(第 22 回)論文集, 2P2-13 (2008)
- [佐々木 04] 佐々木英史, 高橋由雅, 第 32 回構造活性相関シンポジウム講演要旨集, pp.137-140 (2004)
- [Takahashi 04] Y. Takahashi, Chemical data mining based on non-terminal vertex graph, *Systems, Man and Cybernetics*, 2004 IEEE International Conference, Volume 5, 2004 pp. 4583 - 4587(2004).