

機械学習に基づく嗅覚受容体のリガンド予測

Ligand Prediction for the Olfactory Receptors based on Machine Learning

小坂 宏四郎 上野 恵介 峯田 克彦 遠藤 俊徳

Koshiro Kosaka Keisuke Ueno Katsuhiko Mineta Toshinori Endo

1. 研究目的

本稿では、嗅覚受容体とそのリガンドが特異的に結合することに着目し、結合関係を予測する方法を提案する。提案手法では、自己組織化マップ(SOM)[4]を用いて嗅覚受容体とリガンドの物理化学的特性および構造から分類し、双方の分類群を対応付ける。

2. 関連研究

嗅覚受容体とリガンドの結合関係が明らかになっている例は非常に少ない[7]。匂い分子の数は膨大であるため、嗅覚受容体への結合を確認する実験は困難である。そのため、コンピュータを用いた予測は有用である。Schmiedeberg[8]は、Rhodopsin を鋳型としたホモロジーモデリングにより嗅覚受容体構造を予測し、ドッキングシミュレーションを行うことで効果的な予測をしている。また、新村ら[6]は嗅覚受容体の進化について、分子系統解析による分類結果を示した。

3. 提案手法

提案手法では、機械学習のひとつである SOM により分類を行う。SOM は入力ベクトルの組をその類似度に基づいて学習し、クラス分けするためのネットワークである。結果は 2 次元のマップとして得られるため、多次元のデータを 2 次元に写像するために用いられる。この写像においてトポロジーは維持されるため、多次元空間において近くに位置するデータは 2 次元のマップにおいても近くに位置するように写像される。つまり、この手法の利点はマッピングの結果を容易に可視化できる点にある。さらに、高次元データを教師なしにクラスタリングできる。SOM にはデータの入力順序に依存する性質がある。そのため本手法では、その性質を取り除くための変更が加えられた Batch Learning SOM(BL-SOM)を用いる。以下に、BL-SOM に用いる嗅覚受容体およびリガンドの入力ベクトルの詳細を述べる。

3. 1 嗅覚受容体の特性

嗅覚受容体については、新村らによって報告されている偽遺伝子化していない 388 種類のヒト嗅覚受容体遺伝子を採用した。また受容体構造が解明されていないために、高精度の構造情報は取得することが困難である。そのため、分子量、疎水性、等電位点、pKa(COOH)、pKa(NH₂)というアミノ酸の物理化学的特性を入力ベクトルに用いた。本手法においては、アミノ酸特性を嗅覚受容体全領域ではなく、リガンド結合に関与していると推定される局所的領域[8]のみ適用した。

3. 2 匂い分子の特性

最初に、SAFC の提供する匂い分子カタログから 170 種類(花, 果実, 悪臭)の匂い分子の情報を取得した。取得した匂い分子の原子の 3 次元座標値を構造発生システム CORINA を用いて計算する。その後、プログラム PETRA により電荷分布と分極率のような物理化学的性質の計算を行う。本研究では、全電荷、 σ 電気陰性度、 π 電気陰性度、有効原子分極率を求めている。そしてこれら 3 次元情報と物理化学的な原子特性が、Gasteiger ら[2]らにより考案された動径分布関数(RDF)に基づく、分子のコード化に用いられる。RDF コード化のための式を以下に示す。

$$g(r) = f \sum_{i=1}^{N-1} \sum_{j>i}^N p_i p_j \exp\{-B(r-r_{ij})^2\}$$

f はスケールリングファクター、 N は分子中の原子数である。 p_i と p_j は原子 i と j の物性で、 B は平滑化パラメータである。そして、 r_{ij} は原子 i と j 間の距離である。 $g(r)$ は通常定義済みの間隔内で、多くの離散的なポイントで計算される。動径分布関数は、個別の原子間距離の一種の確率分布である。つまり、全 N 原子に対する RDF は、半径 r の球における原子の確率密度と解釈できる。指数の項は追加的に平滑化パラメータ B を含む。平滑化パラメータは原子の運動を定義する温度因子として考えられる。RDF コードの計算に特徴ある原子特性 p_i と p_j の包含によって異なる要求に対し、この情報を適応させることが可能になる。

以下、嗅覚受容体とリガンドについて BL-SOM により分類した結果を述べる。

4. 嗅覚受容体の分類

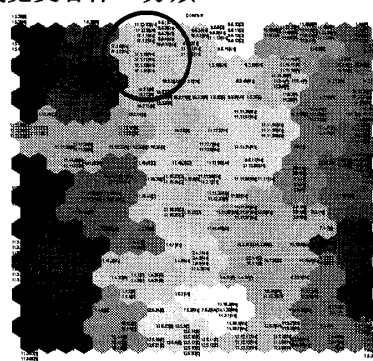


図 1 BL-SOM の適用結果

本手法による嗅覚受容体の分類と分子系統解析による分類を比較すると、83.6% 一致していた。また、分子系統解析で未分類の嗅覚受容体 96 個のうち、58 個、再分類することができた。

結合するリガンド既知の嗅覚受容体 HsOR17.1.2, 17.1.6, 17.1.7, 17.1.11 [1, 3, 7, 8] が同一グループに分類

された(図1の円領域)。また、HsOR17.1.6, 17.1.7, 17.1.11には共通のリガンドである Helional が存在する[3, 8]。つまり、結合するリガンドについても物性、また、構造が類似している可能性が高い。この確認のため、以下にリガンドについて同様の分類を行った。

5. 匂い分子の分類

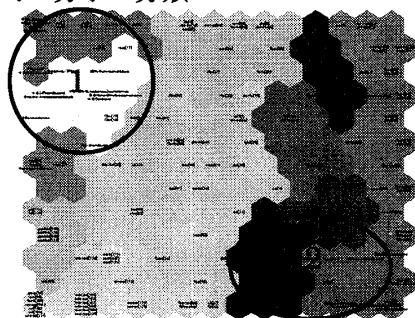


図2 BL-SOMの適用結果

分類の結果、HsOR17.1.2, 17.1.6, 17.1.7, 17.1.11に結合するリガンドは図2の領域1と2に分類された。領域が2つに分かれているのは、領域1はベンゼン環を含む分子群であり、領域2は主に直鎖状の分子群の違いによるものと考えられる。つまり、リガンドについても、それぞれが同じ分類になっている。以上の結果から、本手法により嗅覚受容体とリガンドの分類結果を対応可能であることが示唆された。以下に、本手法を用いることで嗅覚受容体とリガンドの予測にどのように適用できるかを述べる。

6. 嗅覚受容体のリガンド予測への適用

Helionalと嗅覚受容体でドッキングシミュレーションを行い、本手法の利点を述べる。

まず、Rhodopsin (1F88)を鋳型としたホモロジーモデリングから嗅覚受容体の構造を予測した。次に、AutoDock4[5]を用いて、ドッキングシミュレーションを行った。AutoDock4のシミュレーション結果の信頼性は、最終的に結合した立体構造の類似性に依存してくる。そのため、信頼性評価に、10種類の結合結果のうち最もエネルギーが低い立体構造のRMSD値とその他の立体構造のRMSD値を比較し、その差が0.5Å以内の構造を算出する。さらに、構造の類似性だけではなく、AutoDock4では結合結果の相互作用エネルギーが低い程安定した構造と言える。

ドッキングシミュレーションの結果が図3である。図3(b)の円領域は、最も低いエネルギーを示した領域を示しており、結合の安定性が高いと考えられる。また、この領域を(c)で見ると、前述の構造数が最も多い領域であった。以上より、本稿では、この領域に存在する受容体とリガンドが「潜在的に結合可能」とであると定義した。事実、この領域に含まれる受容体は Helional と結合すると報告のある HsOR17.1.6, 17.1.7, 17.1.11 が含まれている。

このドッキングシミュレーションの結果と前述の図1,2の結果を併用することで、嗅覚受容体とリガンドの結合を検索する場合、直観的にもわかりやすく効率的に結合先を検索することが可能になると考えられる。

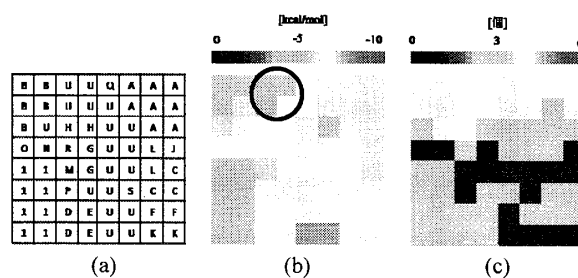


図3 ドッキングシミュレーション結果

(a) 図1の分類から代表的な受容体を選択し、その分類名(新村ら[6])を対応箇所とともに示したマップ。(b) (a)で示した受容体とリガンド(Helional)とのドッキングシミュレーションで算出された相互作用エネルギー。(c) は最もエネルギーが低い構造との誤差が0.5Å以内の構造数。

7. まとめ

本稿では、嗅覚受容体とそのリガンドの結合関係を予測する方法を提案した。自己組織化マップを用いて、嗅覚受容体およびリガンドの分類を行った結果、Helionalをリガンドに持つ複数の嗅覚受容体が同一グループに分類されるなど、双方の分類結果を対応付けできることが示唆された。

さらに、本手法により、直観的にもリガンドが潜在的に結合可能な嗅覚受容体を検索することを容易にできると考えられる。

参考文献

- [1] Doszczak, L., Kraft, P., Weber, H-P, Bertermann, R., Triller, A., Hatt, H. and Tacke, R., Prediction of Perception: Probing the hOR17-4 Olfactory Receptor Model with Silicon Analogues of Bourgeonal and Liliat' *Angew. Chem. Int. Ed.*, 46: 3367-3371, 2007
- [2] Gasteiger, J., Schuur, J., Selzer, P., Steinhauer, L. and Steinhauer, V., "Finding the 3D structure of a molecule in its IR spectrum". *Fresenius Journal Of Analytical Chemistry*, 359(1): 50-55, 1997.
- [3] Jacquier, V., Pick, H and Vogel, H. Characterization of an extended receptive ligand repertoire of the human olfactory receptor OR17-40 comprising structurally related compounds *Journal of Neurochemistry*, 97: 537-544, 2006
- [4] Kohonen, T., Self-organizing maps, Berlin, Springer-Verlag, 1995.
- [5] Morris, G. M., Goodsell, D. S., Halliday, R.S., Huey, R., Hart, W. E., Belew, R. K. and Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and Empirical Binding Free Energy Function *J. Computational Chemistry*, 19: 1639-1662, 1998
- [6] Niimura, Y. and Nei, M., Evolution of olfactory receptor genes in the human genome. *Proc. Natl. Acad. Sci. USA*, 100: 12235-12240, 2003.
- [7] Triller, A., Boulden, E. A., Churchill, A., Hatt, Hanns., Englund, J., Spehr, Marc., and Sell, C. S. "Odorant-Receptor Interactions Odor Percept: A Chemical Perspective" *CHEMISTRY BIODIVERSITY*, 5: 862-886, 2008
- [8] Schmiedeberg, K., Shirokova, E., Weber, H-A., Schilling, B., Meyerof, W. and Krautwurst, D., Structural determinants of odorant recognition by the human olfactory receptors OR1A1 and OR1A1 *Journal of Structural Biology*, 159: 400-412, 2007.