

令和 5 年 6 月 16 日現在

機関番号：14401

研究種目：基盤研究(C) (一般)

研究期間：2017～2022

課題番号：17K08235

研究課題名(和文) 多元的非線形相関次元圧縮法の開発と応用

研究課題名(英文) Development and Applications of Nonlinear Dimension Reduction with Weak Supervisors

研究代表者

高木 達也 (TAKAGI, Tatsuya)

大阪大学・大学院薬学研究科・特任教授

研究者番号：80144517

交付決定額(研究期間全体)：(直接経費) 3,700,000円

研究成果の概要(和文)：IFIE(Inter Fragment Interaction Energy)は、FMO法でもたらされる有用な情報であり、どの残基と基質(例えば医薬品)が重要な相互作用をしているかがわかる。医薬品のpIC50との直線回帰は回帰係数として、残基の重要性を示している。然しながらエントロピー項や遠くの静電相互作用を過大評価するFMOの傾向を修正するためには、化学記述子の導入が必要になる。この課題では、化学空間(化学記述子)と生理活性空間(IFIE)を別途PLS回帰し、更に多様体学習により2次元空間にマッピング、リガンドと化学記述子の2次元座標を得、評価指標が最良になるように残基、記述子を選択した。

研究成果の学術的意義や社会的意義

この課題の成果は、COVID-19パンデミックなどもありまだ十分に公開できていないが、成果が広く共有されれば、FMO計算の結果がpIC50とより良い相関を示し、FMO計算から直接、新たな医薬品の開発に繋がる情報を得ることができる。今回のパンデミックで判明したように、医薬品の開発は時として十分な時間を与えられない。そのような場合、医薬品開発の時間短縮に占める本法の意義は大きいと考える。

又、学術的には化学空間と生物空間を別々にPLS回帰し、その後多様体学習を用いることにより、より互いからの「雑音」を消去する形で情報の抽出が可能になったことがあげられる。

研究成果の概要(英文)： Inter Fragment Interaction Energy (IFIE) is a useful piece of information provided by the FMO method, which shows which residues and substrates (e.g. pharmaceuticals) are interacting in an important way. Linear regression with the pIC50 of the drug substance gives the importance of the residue as a regression coefficient. However, the introduction of chemical descriptors is necessary to correct the tendency of FMOs to overestimate the entropy term and move away from electrostatic interactions. In this task, the chemical space (chemical descriptors) and the bioactive space (IFIE) were regressed separately by PLS regression and further mapped into 2D space by manifold learning to obtain the 2D coordinates of the ligands and chemical descriptors, and the residues and descriptors were selected to obtain the best evaluation index.

研究分野：計量薬学

キーワード：FMO法 IFIE 化学記述子 生理活性空間 多様体学習

## 1. 研究開始当初の背景

過去、医薬品設計は受容体(多くはタンパク質)の立体構造を基にした Structure-Based Drug Design (SBDD) と、リガンドのフラグメントや記述子、化学構造 Fingerprintなどを基にした Ligand-Based Drug Design (LBDD) の主として2種類の方向から行われてきた。HIV プロテアーゼ阻害剤の多くはSBDDで、DPP4 阻害剤の一部など、SBDDの後、LBDDで最適化された医薬品も上市されている。しかし従来、これら大別して二つの手法は、それぞれ別個に用いられることが殆どであった。両者が使用される場合でも、多くは、SBDDで大凡の化学構造を導出し、LBDDで最適化するなど、段階的に別個に用いられてきた。しかし、段階的に用いてもそれぞれの長所が生かされているとは限らない等、問題点は指摘されていた。

## 2. 研究の目的

本研究では、医薬品設計の初期段階で用いられる次元圧縮法(パターン分類法)に、SBDDから得られる情報とLBDDからの情報を同時に用いることにより、より効果的なパターン分類法を開発することを主眼に置いた。

高次元のデータから有用な知見を得るデータマイニングは薬学分野でも注目が高まっている。データマイニングに用いられる手法の1つである次元圧縮はデータの特徴抽出や機械学習における教師なし学習、ビッグデータ、ミドルデータの可視化など様々な分野で用いられる。一般的に用いられる次元圧縮として主成分分析(Principal component analysis: PCA)があげられる。これまで、薬学分野でもPCAを用いた報告は多くされている。しかし、PCAはデータが低次元の線形空間に埋め込まれていると仮定しているため、実際には線形空間ではない低次元に埋め込まれているデータを正しく抽出できない場合がある。一方で、多様体学習は、高次元に存在するデータが実際には低次元空間(低次元多様体)が埋め込まれている状態であるとし、その低次元のパラメータに非線形的に次元圧縮する方法である。多様体学習を用いることで従来の線形圧縮では正しく抽出できなかったデータの次元圧縮が行えるという利点がある。近年、Isomapを用いたbiomedical spaceの可視化や次元圧縮法を用いたhERG毒性や他の化学物質が示す毒性に相関する数値の算出をはじめとし、多様体学習を用いた新たな解析手法が薬学分野にも用いられるようになった。しかし、実際のin silico医薬品分子設計ではタンパク質-リガンド間相互作用を解析し、新規化合物の設計および薬理活性の向上を図ることを目的とした、タンパク質-リガンド間相互作用を取り入れた機械学習や統計学を用いたQSAR(定量的構造活性相関解析)モデルは少数であり、特に多様体学習を用いた報告は未だ行われていない。そこで、当課題では、タンパク質リガンド間相互作用の計算結果を包括する多様体学習を用いた新規手法の開発を目的とする。

本課題ではFMO(Fragment Molecular Orbital)法により計算された、セクレターゼと阻害剤の相互作用を例にあげた。その理由は、第一に、アルツハイマー病の根本的治療法として、アミロイドの産生抑制が目標として掲げられており、医薬品開発のターゲットとして重要な位置を占めること(セクレターゼはアミロイド産生に重要な働きを見せる)、第二に当該相互作用が、FMO計算の結果と生理活性指標であるpIC<sub>50</sub>の相関が良好でないため、機械学習を用いた修正が望まれているからである。以下、セクレターゼを例として、計算結果を表示したい。

## 3. 研究の方法

提案手法のフローチャートを図1に示す。

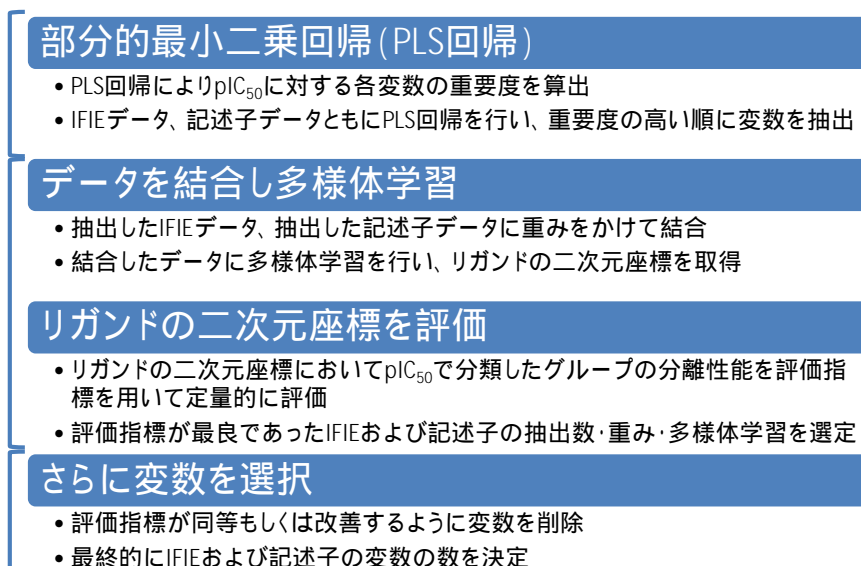


図1 提案手法のフローチャート

の変数選択ではステップワイズ変数減少法を用い、最終的な2次元マップを得、変数を減少しても分類がうまく行われていることを確認する。こうすれば、分子設計に重要でない IFIE や記述子を削除することが出来る。

#### 4. 研究成果

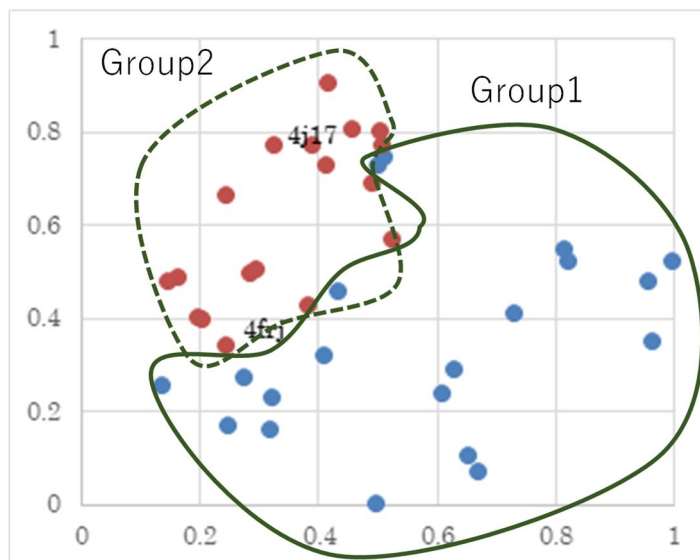


図2 IFIE、記述子の値を用いた多様体学習 (LLE) による高活性化化合物 (Group2) と低活性化化合物 (Group1) の分類 (n\_IFIE = 60, n\_Des = 80, p = 0.8, method = LLE(k=18))

多様体学習による基質 (化合物) の分類は、上図の通り、pIC50 による分類ができるようになりガンドの二次元座標が得られたと言える (Group1 が低活性グループ、Group2 が高活性グループ)。以下、この座標を用いて変数を選択した。

表1 選択した各残基との IFIE の詳細 (一部)

| 残基   | IFIE(kcal/mol)の最小値 | ポケットに位置 | 報告文献   | 極性アミノ酸 |
|------|--------------------|---------|--------|--------|
| Q12  | -7.20              | ○       | 34     | ○      |
| L30  | -3.64              | ○       | 34     |        |
| D32  | -48.8              | ○       | 34, 35 | ○      |
| G34  | -12.4              | ○       | 34, 35 |        |
| V69  | -3.95              | ○       | 34, 35 |        |
| Y71  | -9.96              | ○       | 34, 35 |        |
| Q73  | -3.10              | ○       | 35     | ○      |
| K75  | -3.27              | ○       | 34, 35 | ○      |
| W76  | -12.3              | ○       | 34, 35 | ○      |
| F108 | -5.77              | ○       | 34, 35 |        |
| W115 | -4.79              | ○       | 34, 35 | ○      |
| I118 | -4.59              | ○       | 34, 35 |        |
| D228 | -53.9              | ○       | 34, 35 | ○      |
| T231 | -16.0              | ○       | 34, 35 | ○      |
| R307 | -5.94              | ○       | 34, 35 | ○      |
| V336 | -0.97              | ○       | 34     |        |

表 1 は抽出された残基の一部ではあるが、大半がポケットに位置しており、重要な残基が抽出できたことがわかる。

表 2 選択された化学記述子

| 記述子        | 詳細                                      |
|------------|---|
| ATSC0dv    | 荷電子で重み付けしたラグ 0 の中心 moreau-broto 自己相関    |
| ATS6dv     | 価電子で重み付けされたラグ 6 のモラン係数                  |
| Xpc-4d     | シグマ電子で重み付けされた 4 次の Chi path-cluster     |
| JGI7       | 7 次のトポロジカルチャージの平均                       |
| MPC2       | 2 次のパスカウント                              |
| Zagreb2    | Zagreb index(version 2)                 |
| vsurf_HB1  | Volsurf から得られる記述子の 1 つ                  |
| vsurf_HB2  | Volsurf から得られる記述子の 1 つ                  |
| vsurf_W2   | Volsurf から得られる記述子の 1 つ                  |
| SRW10      | Leg10 での歩行回数 (only self returning walk) |
| nHeavyAtom | 重原子の個数                                  |
| IC1        | 1 次の近隣情報                                |
| MWC06      | Leg6 での歩行回数                             |
| VR1_D      | 距離行列からのランダムな固有ベクトル                      |
| MPC3       | 3 次のパスカウント                              |
| vsurf_HB3  | Volsurf から得られる記述子の 1 つ                  |
| MWC07      | Leg7 での歩行回数                             |
| vsurf_W3   | Volsurf から得られる記述子の 1 つ                  |
| MID        | 分子 ID                                   |
| SlogP_VSA2 | MOE で得られる LogP VSA 記述子 2                |

次に抽出された化学記述子の一部を表 2 に示す。重要である意味が必ずしも取れないものも残されているが、SlogP (薬理活性とは異なり pIC50 の値なので、即医薬品活性に結びつくものではないが、疎水性などの指標として重要なものと考えられる) や重原子の個数など、医薬品分子設計に重要と思われるものも多い。必ずしも意義が不明なものも、医薬品分子設計に利用することが可能なものが殆どで、セクレターゼ阻害剤の設計に用い得る。

以上、タンパク質リガンド相互作用の解析手法として多様体学習を用いたものを提案した。本稿では省略したが、この他、変数削除の過程も含まれており、全体として重要な手法を提案できたと考えている。結果として不要な変数を削除でき、pIC50 にとって重要な変数が得られたと考えられる。また、選択された IFIE および化学記述子だけのデータでも多様体学習を行い、得たリガンドの二次元マップでは pIC50 の予測が凡そ可能であった。しかし、データを組み合わせることにより分離性能の高いマップが得られている。以上より、本提案手法は pIC50 に寄与している説明変数を選定することが可能であり、また得られた二次元マップにより、未知のリガンドの pIC50 の推測を可能にする。したがって in Silico 医薬品設計への応用が可能である。

## 5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件/うち国際共著 0件/うちオープンアクセス 2件）

|   |                         |
|---|-------------------------|
| 1. 著者名<br>Watanabe Kazuki, Watanabe Chiduru, Honma Teruki, Tian Yu-Shi, Kawashima Yusuke, Kawashita Norihito, Fukuzawa Kaori, Takagi Tatsuya                            | 4. 巻<br>94              |
| 2. 論文標題<br>Computational ab initio interaction analyses between neutralizing antibody and SARS-CoV-2 variant spike proteins using the fragment molecular orbital method | 5. 発行年<br>2021年         |
| 3. 雑誌名<br>Bulletin of the Chemical Society of Japan   | 6. 最初と最後の頁<br>1794-1798 |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>10.1246/bcsj.20210104  | 査読の有無<br>有              |
| オープンアクセス<br>オープンアクセスとしている（また、その予定である）   | 国際共著<br>-               |

|  |                         |
|--|-------------------------|
| 1. 著者名<br>Kazuki Watanabe, Chiduru Watanabe, Teruki Honma, Yu-Shi Tian, Yusuke Kawashima, Norihito Kawashita, Tatsuya Takagi, Kaori Fukuzawa   | 4. 巻<br>12              |
| 2. 論文標題<br>Intermolecular Interaction Analyses on SARS-CoV-2 Spike Protein Receptor Binding Domain and Human Angiotensin-Converting Enzyme 2 Receptor-Blocking Antibody/Peptide Using Fragment Molecular Orbital Calculation | 5. 発行年<br>2021年         |
| 3. 雑誌名<br>Journal of Physical Chemistry Letters  | 6. 最初と最後の頁<br>4059-4066 |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>10.1021/acs.jpcllett.1c00663  | 査読の有無<br>有              |
| オープンアクセス<br>オープンアクセスとしている（また、その予定である）  | 国際共著<br>-               |

〔学会発表〕 計6件（うち招待講演 3件/うち国際学会 4件）

|                                   |
|-----------------------------------|
| 1. 発表者名<br>高木達也                   |
| 2. 発表標題<br>計量薬学の歴史と未来             |
| 3. 学会等名<br>第49回構造活性相関シンポジウム（招待講演） |
| 4. 発表年<br>2021年                   |

|   |
|---|
| 1. 発表者名<br>渡邊一樹、川嶋裕介、福澤薫、渡邊千鶴、本間光貴、田雨時、川下理日人、高木達也 |
| 2. 発表標題<br>FMO 計算を用いたSARS-CoV-2 のRBD-中和抗体間の相互作用解析 |
| 3. 学会等名<br>第48 回構造活性相関シンポジウム                      |
| 4. 発表年<br>2020年                                   |

|  |
|--|
| 1. 発表者名<br>Yusuke Kawashima, Nanami Mori, Hiroto Mori, Norihito Kawashita, Tian Yu-Shi, Tatsuya Takagi                                     |
| 2. 発表標題<br>Examination of statistical methods for analyzing Fragment Molecular Orbital calculation results on ligand-protein interactions. |
| 3. 学会等名<br>9th Conference of the Asia-Pacific Association of Theoretical and Computational Chemists (国際学会)                                 |
| 4. 発表年<br>2019年  |

|   |
|---|
| 1. 発表者名<br>Naoto Motoyama, Tian Yu-shi, Yuji Hashimoto, Hiroto Mori, Tatsuya Takagi, Norihito Kawashita   |
| 2. 発表標題<br>SBDD of MDM2 and beta-Secretase Inhibitors Using FMO and PLS                                   |
| 3. 学会等名<br>20th International Conference on Computer Aided Drug Designing and Drug Delivery (招待講演) (国際学会) |
| 4. 発表年<br>2018年   |

|   |
|---|
| 1. 発表者名<br>Naoto Motoyama, Tian Yu-shi, Yuji Hashimoto, Hiroto Mori, Tatsuya Takagi, Norihito Kawashita |
| 2. 発表標題<br>SBDD of MDM2 and beta-Secretase Inhibitors Using FMO and PLS                                 |
| 3. 学会等名<br>11th China-Japan Symposium on Drug Design and Development (国際学会)                             |
| 4. 発表年<br>2018年   |

|  |
|--|
| 1. 発表者名<br>Tatsuya TAKAGI  |
| 2. 発表標題<br>SBDD of MDM2 Inhibitors Using FMO and Data Mining Method                                    |
| 3. 学会等名<br>18th International Conference on Medicinal Chemistry & Targeted Drug Delivery (招待講演) (国際学会) |
| 4. 発表年<br>2017年  |

〔図書〕 計1件

|   |                 |
|---|-----------------|
| 1. 著者名<br>Yuji Mochizuki, Shigenori Tanaka, Kaori Fukuzawa(ed), Tatsuya Takagi et al.,                    | 4. 発行年<br>2021年 |
| 2. 出版社<br>Springer  | 5. 総ページ数<br>677 |
| 3. 書名<br>Recent Advances of the Fragment Molecular Orbital Method: Enhanced Performance and Applicability |                 |

〔産業財産権〕

〔その他〕

|  |
|--|
| FMODB<br><a href="https://drugdesign.riken.jp/FMODB/">https://drugdesign.riken.jp/FMODB/</a> |
|--|

6. 研究組織

|       | 氏名<br>(ローマ字氏名)<br>(研究者番号)  | 所属研究機関・部局・職<br>(機関番号) | 備考 |
|-------|----------------------------|-----------------------|----|
| 研究協力者 | 田 雨時<br><br>(Tian Yui-Shi) |                       |    |

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

|         |         |
|---------|---------|
| 共同研究相手国 | 相手方研究機関 |
|---------|---------|