

機械学習を用いたアミノ酸構造情報に基づくタンパク質の可溶性予測

鈴木皓大・榊原一紀・中村正樹・篠田優（富山県立大学）
松井大亮（立命館大学）・浅野泰久（富山県立大学）

1 概要

タンパク質のアミノ酸配列（一次構造）は、タンパク質の立体構造（二次構造）や性質を決定していると考えられていることから、機能性タンパク質の一部を変異させて所望の機能を発現させようとする実験や研究が行われている。化学産業では可溶化することで、触媒としての産業利用が期待される酵素が大量生産されている。しかし大量生産において不溶性を発現するような酵素が生産される問題がある。その解決のために、不溶性酵素にランダムな変異を加えることで可溶化を試みている [1]。しかし変異の組み合わせは膨大となるため、検証実験に膨大な時間と費用を必要となることから、可溶性化する特徴を予測・可視化することが強く求められている。本研究ではタンパク質のアミノ酸及び構造情報を機械学習の入力として可溶性を予測するモデルを作成し、得られたモデルの解析によりタンパク質の可溶性に影響を与える要因を明らかにする。

2 LSTM を用いた可溶性予測

再帰型ニューラルネットワークの拡張型である LSTM を用いてタンパク質の可溶性を予測する。入力にはタンパク質のアミノ酸配列である一次構造と立体構造を表す二次構造を情報として与える。各タンパク質の可溶性をラベルデータとして学習を行い、未知のタンパク質の可溶性を予測するような回帰モデルを作成する。得られたモデルに基づき、アミノ酸配列内の一つのアミノ酸を自身以外のアミノ酸に置換する（一部位変異）を行い、学習済みのモデルに予測させることで可溶性と変異するアミノ酸性質の傾向を分析する。また二次構造の一つである α ヘリックスについて、構造内のアミノ酸の配置関係を見るために、 α ヘリックス上のアミノ酸を変異した際にヘリカルホイールを描写する。これにより親水性・疎水性面を安定させるような変異をすることで可溶性が向上するというヘリックスルール [1] に基づくような傾向が見られるか、検証を行う。

3 予測結果の分析

タンパク質の一次構造と二次構造を入力に用いた回帰モデルの評価指標である決定係数 R^2 は 0.58 と比較的高い値が見られるため、タンパク質の可溶性に影響を与える要因を分析するためのモデルが作成できたと考えられる。

次に一次構造と二次構造を入力とした回帰モデルを用いて、一部位変異によって得られた新たな変異タンパク質に対して可溶性を予測する。最も可溶性が向上した上位 100 個抽出する。その中でも変異させた位置が α ヘリックスの

場合、変異させた箇所を中心に数残基取り出してヘリカルホイールを描写した図を図 1 に示す。図 1 からは変異前の親水面における疎水性アミノ酸が親水性アミノ酸に置換されているような結果が得られた。図 1 から親水面・疎水面を安定させるようなヘリックスルールに基づくような変異が確認できた。図 2 は PyMol と呼ばれるグラフィックスツールを使用してタンパク質を表示した。変異で可溶性が大きく向上した変異部分を紫色、大きく低下した変異部分を黄色でそれぞれ示している。図 2 の赤の円で示した部分は紫色の変異部分が集中している箇所であり、可溶性に影響していると考えられる部分が可視化できた。また、変異が集中している箇所は α ヘリックスの立体構造上に集中していることが可視化によって確認できた。以上のことから、回帰モデルがヘリックスルールに基づくような特徴を捉えていると考えられる。

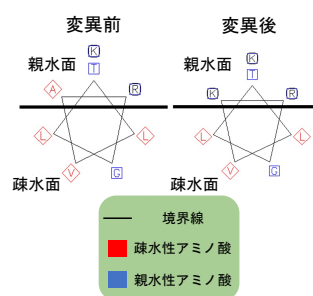


図 1: 一部位変異によるヘリカルホイール

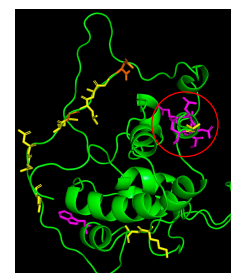


図 2: 一部位変異による可溶性が向上する部分および低下する部分の表示

4 まとめ

回帰モデルの決定係数の値が比較的高いことから、分析可能なモデルが作成できたと考えられる。また作成したモデルによって得られた一部変異タンパク質の可溶性予測とヘリカルホイール、PyMol からヘリックスルールに基づくような傾向を持つこと確認できた。

今後はより精度の高いモデルの作成のために回帰モデルの分析を行う必要がある。また、変異導入実験を行い、モデルの評価を行う必要がある。

5 謝辞

本研究は「くすりのシリコンバレー TOYAMA」創造コンソーシアムの助成を受けたものです。

参考文献

[1] D. Matsui, et al.: Rational identification of aggregation hotspots based on secondary structure and amino acid hydrophobicity, *Scientific Reports* 7, 9558, 2017.