

ペプチドの線維形成に関する実測データから得られる主成分得点と分子記述子の相関を用いた毒性配列予測

(甲南大学 FIRST ^{1,3} 神戸大学院医 ²) ○寺尾泰晟¹・小林もえ¹・武川公²・梅谷智弘³・栢森史浩¹・臼井健二¹

Prediction of cytotoxic peptide sequences using principal component scores and molecular descriptors (¹*Graduate School of Frontiers of Innovative Research in Science and Technology Konan University*, ²*Graduate School of Medicine, Kobe University*, ³*Department of Intelligence and Informatics, Konan University*) ○Taisei Terao¹, Moe Kobayashi¹, Akira Takekawa², Tomohiro Umetani³, Fumihiro Kayamori¹, Kenji Usui¹

Effective peptide designing studies using phage display or artificial intelligence (AI) have been actively studied. However, it is difficult to design sequences with amino acids or functional groups that do not include those in the training data. In this study, we attempted to design sequences by analyzing measured data on the fibril formation of selected short peptides using statistical analysis and molecular descriptors. As a result, we succeeded in designing sequences with an artificial functional group applicable to cell culture substrates.

Keywords : *Immobilized peptide, β -Sheet structure, Statistical analysis, Molecular descriptors, Cell adhesion.*

近年機械学習や AI を活用した、ペプチドの特性を予測する研究や、特定の機能（特性）を有するペプチドの配列設計が盛んに研究されている 1)。しかしこれらの手法は、教師データに無いアミノ酸や官能基を有する配列の設計を行うことが難しい 2)。そこで本研究では、短鎖の線維形成ペプチドを題材とし、数十種類の配列の線維形成に関する実測データを統計解析と分子記述子を用いて解析することで、ライブラリ中には含まれない官能基を有する人工の線維形成ペプチド配列を設計し、それらについて細胞毒性の予測を試みた 3)。初めに線維を形成し細胞毒性を有する短鎖ペプチドをコア配列として選定した。コア配列と、コア配列の N 末端に標準アミノ酸を 1 残基ずつ付加した 21 種類の配列を作製した。種々のペプチド配列と線維形成の相関を解析するため、線維形成に関する実測データを得、主成分分析（PCA）を行った。PCA の結果得られた主成分得点と、計算科学分野で使用される分子記述子を比較し、最も相関の高い分子記述子を用いて PC1、PC2 の解釈を行った。次に種々のペプチド配列の細胞毒性や細胞接着など特性の実測データを得、PCA の二次元座標上に示し、線維形成と特性の傾向を得た。以上の主成分得点と分子記述子及び毒性や接着など特性のデータを考慮し、21 種類の配列に含まれない人工の官能基などを付与した、ペプチド配列についてその細胞毒性や細胞接着の予測を試みた。

1) Tran, D. P. *et al.*, *Sci. Rep.*, **11**, 10630 (2021).

2) Batra, R. *et al.*, *Nat. Chem.*, **14**, 1427-1435 (2022).

3) K. Usui *et al.*, *Mol. BioSyst.*, **2**, 417-420(2006).