



## 環状ペプチドの膜透過性データベースを公開

－ペプチド創薬へのAI適用の大幅な加速に期待－

### 【要点】

- ペプチド創薬の最大の課題である膜透過性のデータを網羅的に収集したデータベース CycPeptMPDB を無償公開。
- 7334種類の環状ペプチドのデータを収集・分類・情報解析し、高機能専用データベースを構築。
- データ不足により従来は困難だった、AIによる膜透過性予測の研究の加速に期待。

### 【概要】

東京工業大学 情報理工学院 情報工学系の秋山泰教授、柳澤溪甫助教、李佳男博士後期課程学生らのグループは、環状ペプチドの膜透過性に関する実験データを文献等から網羅的に収集した CycPeptMPDB データベースを構築し、無償公開した。

環状ペプチド創薬で重要となる膜透過性のデータを収載したデータベースはこれまで存在せず、異なる測定方法の間では数値を単純に比較できない等の注意点もあり、専門的なデータベースの構築が望まれていた。CycPeptMPDB は、論文や特許から 7334 種類の環状ペプチドの膜透過性データを収集した上で、化学構造や三次元配座データ等も統一的形式で提供している。

環状ペプチドは次世代創薬の主役の一つとして期待されているが、薬剤が細胞内に到達するために必要な膜透過性が一般に低いことが最大の課題となっている。ペプチドを合成して標的への結合親和性などを調べる実験をする前に、膜透過性を予測できれば、後から膜透過性の低さで苦しむことなく、効率的な薬剤候補探索が可能となる。

研究グループは物理化学シミュレーションによる独自の膜透過性予測法も発表しているが、AIを使った予測には大量のデータが必要とされているため、研究分野の発展のために収集したデータを使いやすいデータベースとして無償公開した。当成果は、*Journal of Chemical Information and Modeling* (JCIM)誌の電子版に3月17日に掲載された。

## 環状ペプチドの膜透過性を収集したデータベース

環状ペプチドは従来は不可能と考えられてきた標的にも結合可能なため、次世代の薬として期待されています

しかし、環状ペプチドの薬としての有効性を左右する「膜透過性」を収集したデータベースが存在していませんでした

**CycPeptMPDB: Webから利用できる世界初の環状ペプチド膜透過性データベース**

7,334種の環状ペプチドのデータを集積

- 立体配座
- 精密な配座表現  
PEPT1([DL])  
\$PEPT1, PEPT1,  
1:R1-6,R2SSS
- 各種実験で計測された膜透過性データ
- 物性値インデックス (脂溶性など)
- 環状ペプチドを構成する部分構造 (モノマー) の同定
- 閲覧・検索および可視化の強力なツール
- 無償で利用やダウンロードが可能

**CycPeptMPDBは、創薬研究者らに利用されるだけでなく、膜透過性予測のための機械学習モデルの訓練データとして活用できます**

CycPeptMPDB: A Comprehensive Database of Membrane Permeability of Cyclic Peptides

Li et al. (2023) | *Journal of Chemical Information and Modeling* | 10.1021/acs.jcim.2c01573



### ●背景

低分子を用いた創薬に代わって、ペプチド（用語1）を用いるペプチド創薬が期待されている。特に、直鎖状のペプチドに比べて体内での安定性が改善できる**環状ペプチド**（用語2）を用いた創薬に期待が集まっており、我が国では標的への**結合親和性**（用語3）が高い環状ペプチドを迅速かつ自動的に探し出すための技術の開発も長年進められてきた。

しかし一方で、標的への結合親和性が高いだけでは薬として使うことができない。その薬剤は副作用がなく、体内で安定に存在でき、標的が存在する部位まで届く必要がある。こうした条件の中で、環状ペプチド創薬において最も重要な課題とされるのが、薬剤が細胞内に到達するために必要な細胞膜透過性である。細胞は薄い脂の膜（**脂質二重膜**）（用語4）に包まれており、細胞内部のタンパク質等を標的として薬剤を設計する場合には、この膜を効率的に通過するのに十分な膜透過性を持たせる必要がある。しかし環状ペプチドは従来の低分子薬よりサイズが大きく、膜透過性が一般に低い傾向がある。ただし自然界では膜透過性の高い環状ペプチドも発見されており、環状ペプチドの全てが細胞膜を通過できないわけではない。

結合親和性が優れた環状ペプチド薬が開発できても、後からの微修正で膜透過性を高めることは困難であり、創薬が行き詰るケースが多数あった。これを解決するため、膜透過性を高める原理を解明しようと世界中で研究が続けられている。しかし、実験データが多数の文献に散在しているために、新たに研究者が参入する上でデータ収集が大きな障壁となる状況が何年間も続いていた。

## ●研究成果

研究グループが今回開発した CycPeptMPDB は、最新の論文や特許を 40 件以上網羅的に調査して、環状ペプチドの膜透過性のデータを収集・分類し、それらに高度な情報解析の結果なども付与したデータベースである。現在、7334 種類の環状ペプチドの膜透過性の測定値が、測定に用いられた実験手法や発表グループなどと詳細に紐づけられて記載されており、特定の条件を満たすデータ群を簡便に検索できる。環状ペプチドは複雑な分子構造だが、CycPeptMPDB では **HELM 形式**（用語 5）と呼ばれる柔軟な記法を採用して、収集した全ての環状ペプチドの構造を統一的に表現している。またユーザの利便性を高めるために **SMILES 形式**（用語 6）でも化学構造を記述し、さらには各環状ペプチド分子が取りやすい三次元構造（立体配座）の代表例も計算して記載している。

膜透過性の公開データは数年前までほとんど存在していなかったが、研究分野の急速な隆盛により、膜透過性に関する論文や公開特許は増えつつある。その機会を捉えて網羅的な収集作業を実施しただけでなく、HELM 形式や SMILES 形式等を用いて正確な情報として公開した点が本成果の特徴である。CycPeptMPDB は、データの分布などを示す様々な統計の可視化機能も有し、またデータを一括でダウンロードすることも可能となっている。

研究グループは、環状ペプチドに関連して様々な技術開発を行ってきた。第一著者である李佳男博士後期課程学生は、環状ペプチドの体内持続性を AI の一種である深層学習技術を用いて予測するシステムも既に開発し、論文発表を行った（参考文献 1）。一見すると、同様のアプローチで膜透過性の予測についても実現できそうに考えられるが、膜透過性予測の場合には、まず**膜透過性の測定手法**（用語 7）の種類の違いにより、測定値を単純に比べられないなど注意点が多く存在する。さらに、膜透過現象そのものもはるかに複雑であるため、体内持続性予測の場合よりも大量のデータを細心の注意をもって収集する必要性に直面した。今回の CycPeptMPDB はそうした点を考慮し、これまでに研究上の必要に応じて工夫してきた整理方法のノウハウに基づいて開発された。全データが統一フォーマットで掲載されているため、この分野に新たに参入して情報解析を始めようとする研究者にも簡便に扱えると考えている。

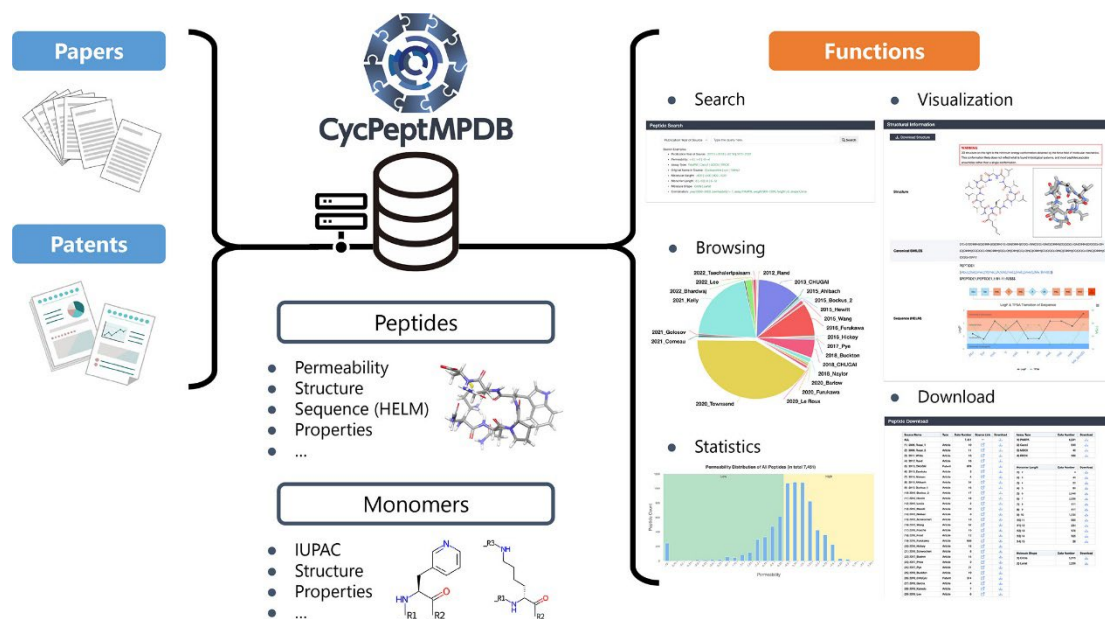


図1 CycPeptMPDB データベースの構成

## ●社会的インパクト

従来の低分子創薬を補う存在として、抗体創薬が広く注目を集めてきた。抗体はきわめて大きな分子であるため、タンパク質の表面に幅広く結合してタンパク質間相互作用を阻害することも可能であることや、狙った標的のみに強く結合する性質により他の分子には干渉せずに副作用が出にくいことなどの優れた利点が知られているが、製造コストに起因する薬価の高さが社会的な問題になっている。こうした高分子と従来の低分子の間にあるのが環状ペプチドなどを用いる「中分子創薬」で、抗体薬に匹敵する優れた薬効と製造コストの大幅低減との二つの利点を併せ持つ、次世代創薬の切り札として期待されている。

しかし、長年の投資にも関わらず、なかなか環状ペプチド薬の成功例が増えてこなかった。主な原因は、細胞膜透過性の低さにあると分析されている。環状ペプチド創薬の成功率を高めるためには、細胞膜透過性に関して AI やシミュレーションによる先回りの予測技術を導入し、薬としての性質が優れた候補分子だけを効率的に探索することが強く求められている。その点で、今回発表した CycPeptMPDB データベースは AI 研究者らがそのまま利用できる形で、膜透過性に関する実験データを網羅的に整理して提供するものであり、AI を用いた膜透過性予測の研究の進展に大きく貢献すると期待されている。

## ●今後の展開

CycPeptMPDB は、同分野で初めての公開データベースである。今後は取案件数を定期的に増やすことに加えて、提供している情報解析（例えば環状ペプチドの三次元の立体配座の提示や、表面積の計算など）の精度を高めていくことを考えている。

また研究グループでは、AI によるアプローチ以外に、高度な物理化学シミュレーション技術を利用した環状ペプチドの膜透過性予測法を発表してきた（参考文献 2,3）。これ

らのシミュレーションによる膜透過性の予測はきわめて強力だが、1つの環状ペプチドの予測だけでも膨大な計算時間がかかるという課題がある。そこで、ごく短いシミュレーションだけを実施し、その時の動き方の様々な特徴から AI によって膜透過性を簡便に予測するアプローチも考えられる。シミュレーションと AI の融合を進めていくためにも、実験データを収載したデータベースはきわめて重要な第一歩となる。

## ●付記

本研究は、科研費「マルチタスク深層学習によるタンパク質の隠された薬剤結合部位の網羅的予測（課題番号：JP20K19917）」、科研費「部分構造の重複を利用した大規模化合物データベース向けバーチャルスクリーニング手法（課題番号：JP22H03684）」、科学技術振興機構（JST）次世代研究者挑戦的研究プログラム「殻を破るぞ！越境型理工系博士人材育成—総合知の創造と社会変革に貢献する（課題番号：JPMJSP2106）」、科学技術振興機構（JST）創発的研究支援事業「マルチモダリティ創薬を拓くインフォマティクス基盤（課題番号：JPMJFR216J）」、日本医療研究開発機構(AMED)「スーパーコンピュータ資源及び大規模シミュレーションと AI に基づく創薬・生命科学の支援（課題番号：JP22ama121026）」の支援により実施された。

## 【用語説明】

- (1) **ペプチド**：アミノ酸が短い鎖状に結合した分子の総称。
- (2) **環状ペプチド**：ペプチドの両端等が結合されて円環状または投げ縄状になった分子の総称。直鎖状の場合に比べて生体内での分解への耐性が一般的に改善される。
- (3) **結合親和性**：分子間の相互作用の強さ。ここでは標的タンパク質等に対する環状ペプチドの相互作用の強さ。薬剤として持つべき重要な性質の一つである。
- (4) **脂質二重膜**：極性リン脂質が向かい合って二層に並んだ生物の細胞膜の基本構造。
- (5) **HELM 形式**：高分子の構造を文字列として記述するための階層的な表記法。Hierarchical Editing Language for Macromolecules の略。
- (6) **SMILES 形式**：化合物の構造を文字列として記述するための表記法の一つ。Simplified Molecular Input Line Entry System の略。
- (7) **膜透過性の測定手法**：人工膜を用いる PAMPA 法、イヌ腎臓由来細胞を用いる MDCK 法など多数の手法があり、同一分子でも必ずしも同じ結果にならない。

## 【参考文献】

- (1) 同研究グループによる環状ペプチドの体内持続性を AI 予測する論文（2022年）：Plasma protein binding prediction focusing on residue-level features and circularity of cyclic peptides by deep learning.  
(DOI:10.1093/bioinformatics/btab726)
- (2) 「アミド-エステル置換により環状ペプチド膜透過性を大幅に向上」東工大ニュース 2023年3月20日  
(<https://www.titech.ac.jp/news/2023/066204>)



- (3) 「大規模分子シミュレーションによる環状ペプチドの細胞膜透過性予測法を開発」 東工大ニュース 2021年7月13日  
(<https://www.titech.ac.jp/news/2021/061266>)

**【論文情報】**

掲載誌： *Journal of Chemical Information and Modeling*

論文タイトル： CycPeptMPDB: A Comprehensive Database of Membrane Permeability of Cyclic Peptides.

著者： Jianan Li, Keisuke Yanagisawa, Masatake Sugita, Takuya Fujie, Masahito Ohue, and Yutaka Akiyama

DOI： 10.1021/acs.jcim.2c01573

**【問い合わせ先】**

東京工業大学 情報理工学院 情報工学系 教授

秋山 泰

Email: [akiyama@c.titech.ac.jp](mailto:akiyama@c.titech.ac.jp)

TEL: 03-5734-3645

**【取材申し込み先】**

東京工業大学 総務部 広報課

Email: [media@jim.titech.ac.jp](mailto:media@jim.titech.ac.jp)

TEL: 03-5734-2975 FAX: 03-5734-3661