



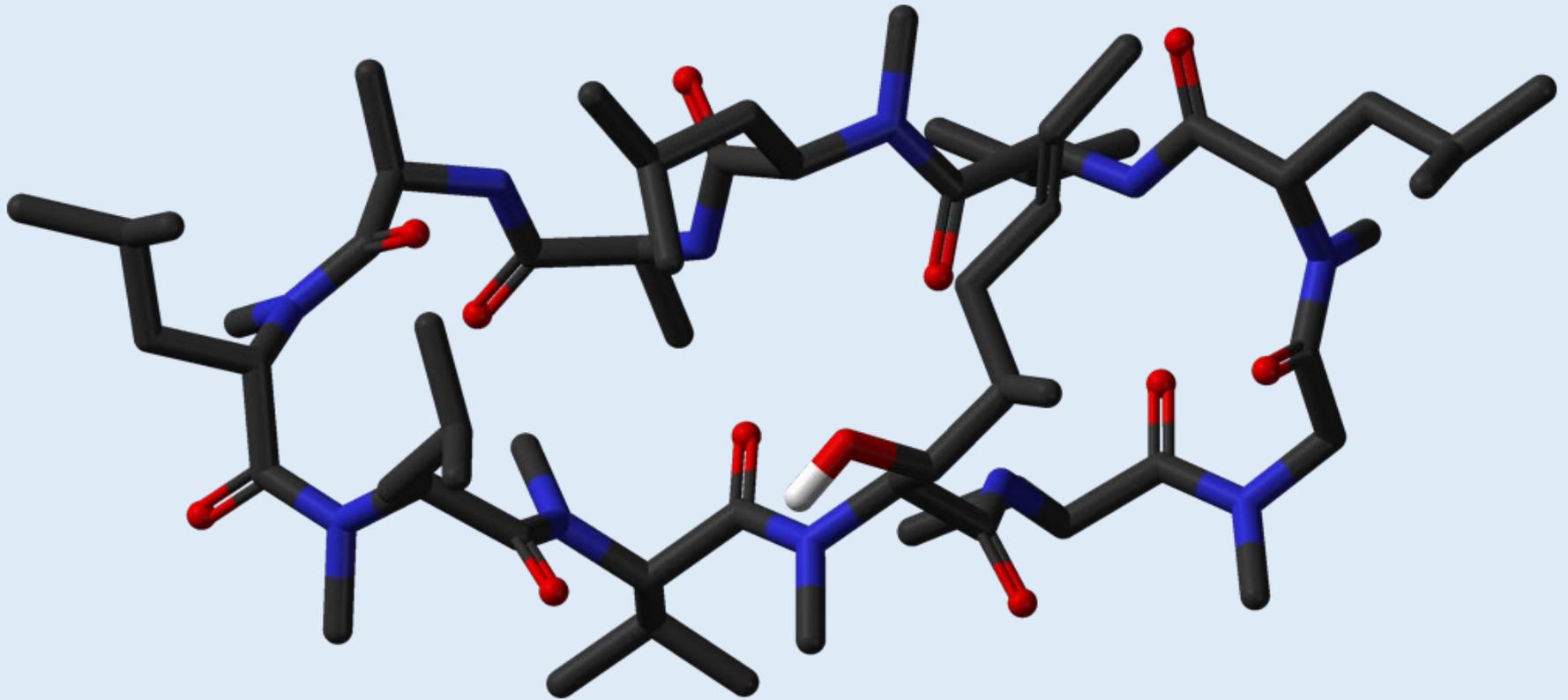
Tokyo Tech

# ペプチド医薬品の膜透過性向上に 貢献する大規模シミュレーション

東京工業大学  
情報理工学院  
教授 秋山 泰

- 1) **ペプチド創薬への期待**
- 2) **機械学習じゃダメなんですか？**
- 3) **分子動力学法による予測への挑戦 (第一次)**
- 4) **分子動力学法による予測への挑戦 (第二次)**
- 5) **まとめ**
- 6) **今後の展望**

# 1) ペプチド創薬への期待



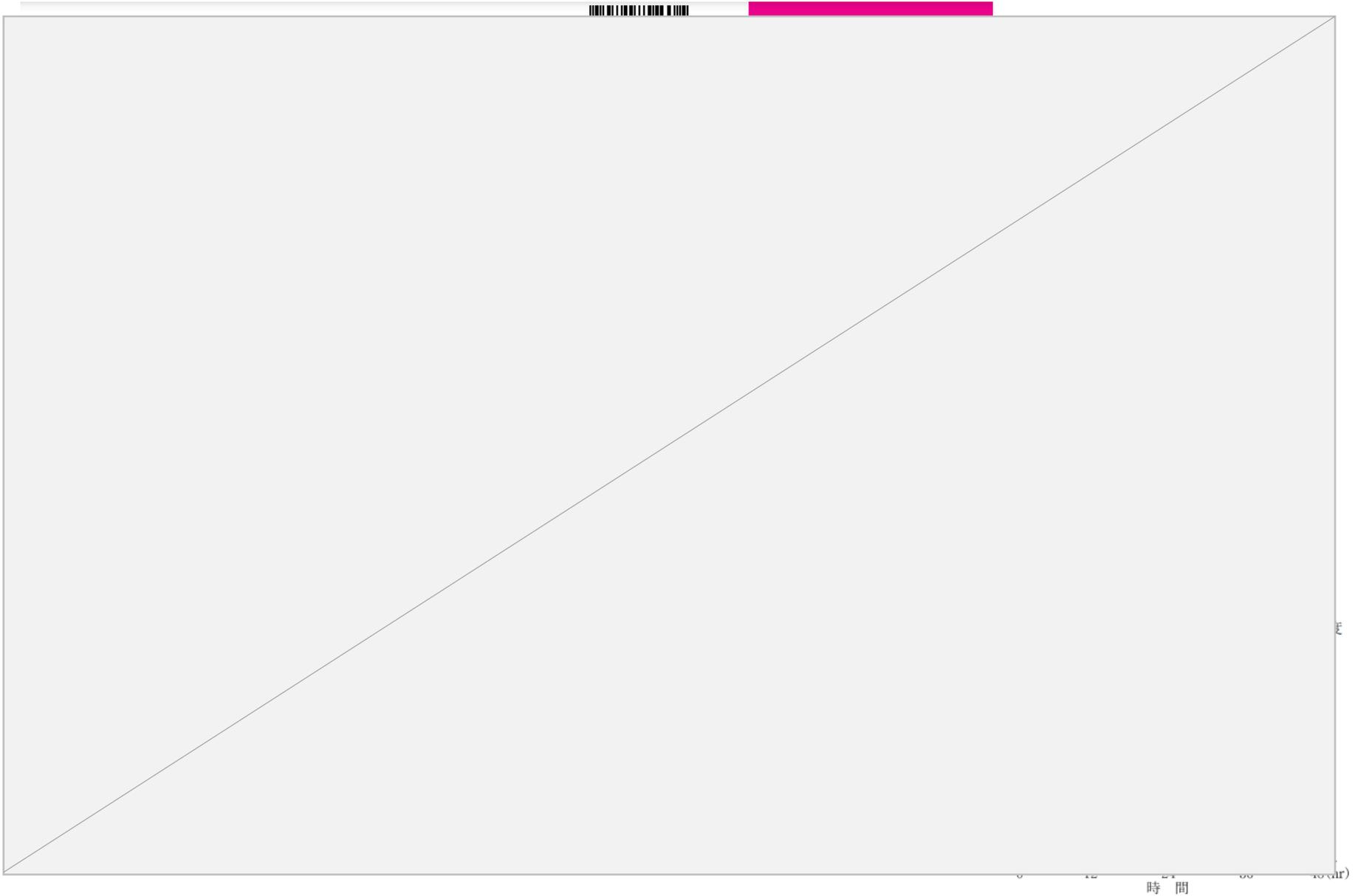
# 創薬におけるコストの上昇

スライドの事後公開に際して、引用図表の一部を処理しています。

- 上市薬 1 案件あたり **3500億円** 程度まで上昇 (1ドル=140円)
- 上市薬 1 案件あたり **10年**

- 従来、製薬企業で一貫して行われてきた創薬の限界
- 単純な標的分子は調べ尽くされてしまった  
→ **生命・細胞システムの熟慮が必要な難しい標的分子が対象に**

# 従来型の低分子医薬品の例



# 低分子と抗体の利点を併せ持つ中分子医薬品

	低分子医薬品	中分子医薬品		高分子医薬品
	有機化合物 ~1,000	ペプチド医薬 600~2,500	核酸医薬 1,000~30,000	抗体医薬 150,000~
分子量	~1,000	600~2,500	1,000~30,000	150,000~
医薬の対象となる標的の数	◎ 多い (ブロックバスター)	○ 多い (アンメットニーズ)	○ 多い (アンメットニーズ)	△ 少ない
標的特異性	△	○	◎	◎
PPI阻害能	△	○	○	○
体内持続性 (体内安定性)	○	△	△	◎
細胞膜透過性	○	△	△	×
標的探索	○	△	△	◎
安全性・毒性	△	△	△	○
製造コスト	◎ 安価	○	○	× 高価

(ご参考) 皮膚がん・肺がん等向け抗体医薬のオプジーボは、2018年4月の国内薬価引下げ後も患者一人当たり年間1000万円超の費用となる。製造には動物由来細胞を利用。

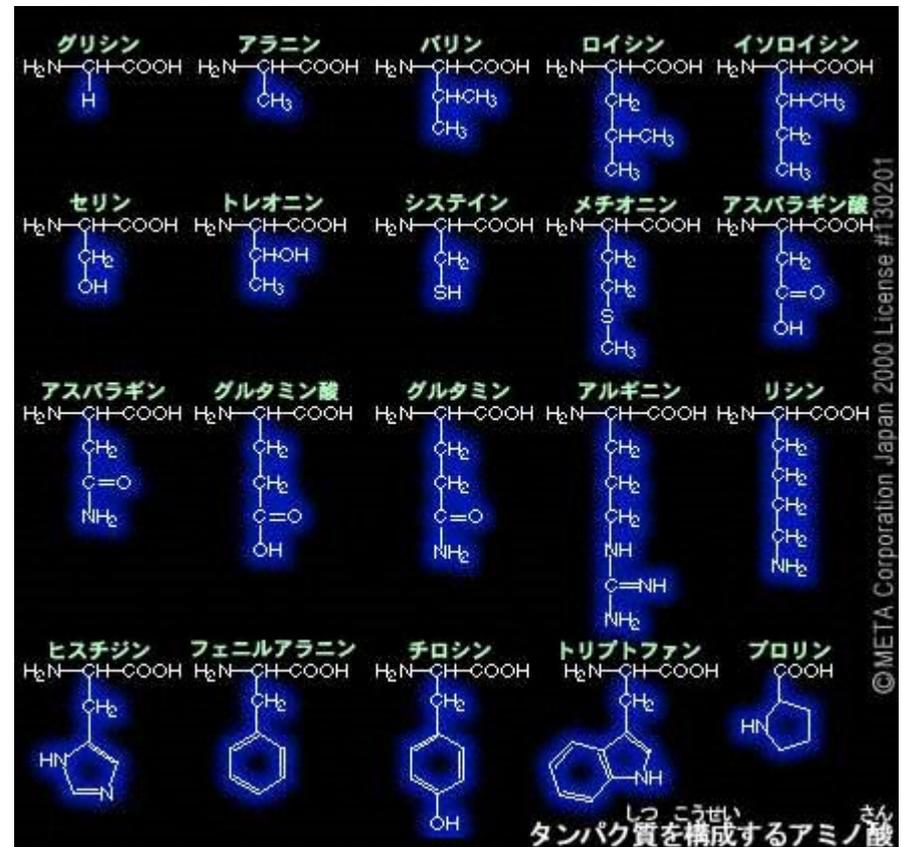
# ペプチドはタンパク質を短くしたもの

タンパク質 = アミノ酸が約50～約3万個 連なった分子



ネスレ株式会社HPより引用

## 20通りの標準アミノ酸



© META Corporation Japan 2000 License #T30201

タンパク質配列の平均アミノ酸数は約300

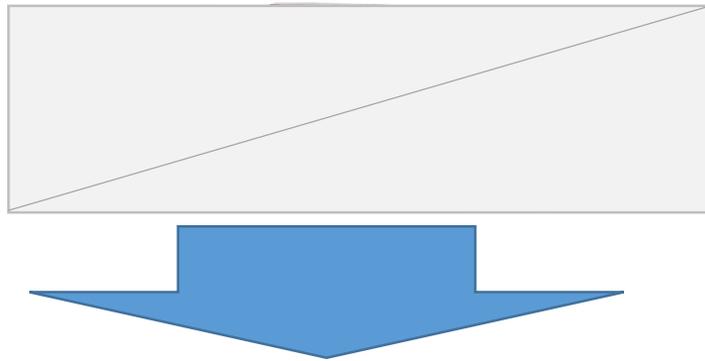
IPA「教育用画像素材集サイト」より引用

# ペプチド創薬

タンパク質 = アミノ酸が約50～約3万個 連なった分子

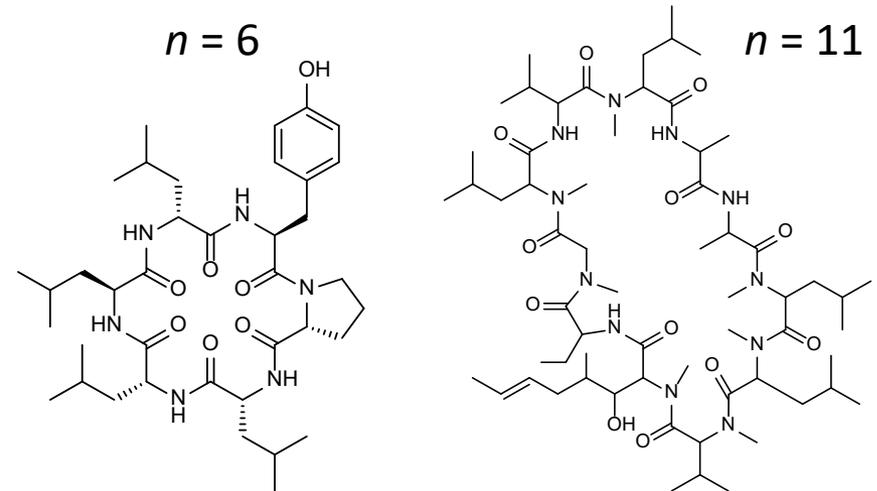
ペプチド = アミノ酸が2個～約50個 連なった分子

- ・ 栄養としてのペプチド
- ・ 機能性ペプチド



## ペプチド創薬

「薬剤設計の標準的骨組み」として  
(特殊環状) ペプチド分子を活用



既存の創薬に  
比べて様々な利点



# 地球の表面全体から、欲しい名刺を探す!?

10

1つの薬剤候補化合物を 名刺サイズ に比べると  
化合物ライブラリーの広さは...



低分子ライブラリー (10<sup>6</sup>種類~)  
= サッカー場の面積  
105m × 68m



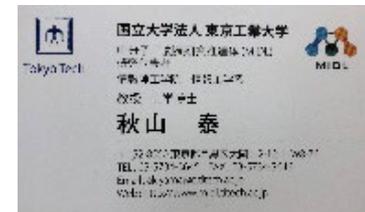
抗体医薬ライブラリー (10<sup>11</sup>種類~)  
= 東京23区の面積



ペプチドライブラリー (10<sup>14</sup>種類~)  
= 日本の面積



創り得る修飾アミノ酸から20種選び、  
n=13の特殊環状ペプチドとして  
合成した際のペプチド配列の多様性  
(20<sup>13</sup>≒10<sup>17</sup>種類)  
= 地球の表面積

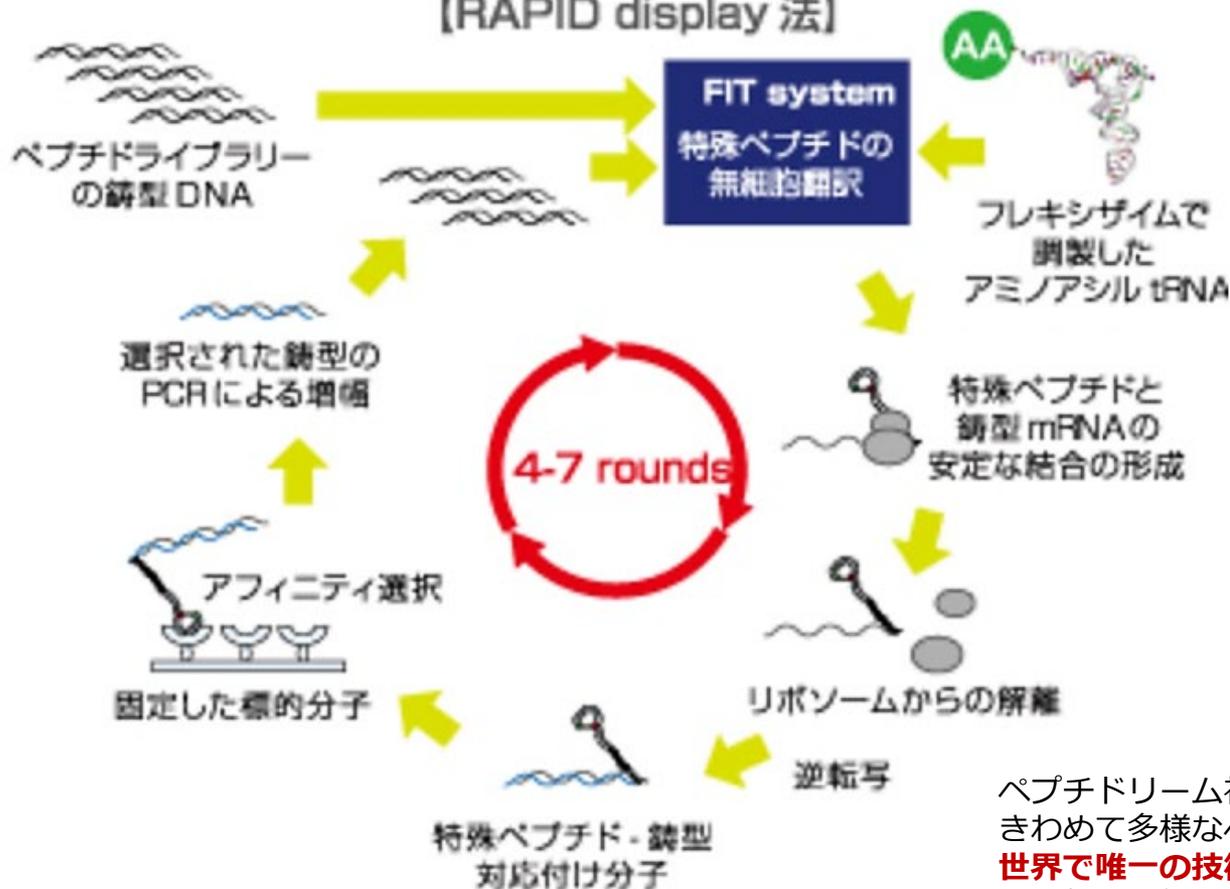


90mm × 60mm

# 迅速にヒットを同定できるPDディスプレイ

## Peptide Discovery Display System (PDDS)

[RAPID display 法]



ペプチドリーム(株) HP  
(旧ページ) より引用

ペプチドリーム社が開発した**PDディスプレイ (PDDS)** は、きわめて多様なペプチドライブラリーを高速で評価できる**世界で唯一の技術**です。

PDディスプレイにより、他のどのディスプレイ技術と比較しても短期間で、ヒットペプチドの同定を行うことができ、しかも極めて高い再現性を有し、かつ、実験者間のデータのばらつきも実質的にありません。

# 東大発ベンチャー企業 ペプチドリーム社

2006年創業

第2回「日本ベンチャー大賞」  
内閣総理大臣賞 受賞



PDPs技術を世界の創薬企業が利用

時価総額 2144億円 (2022/12/2)

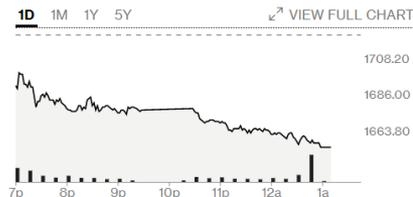
4587:JP Tokyo  
PeptiDream Inc COMPANY INFO

**1,649.00** JPY -74.00 -4.29% ▼

OPEN	PREV CLOSE	VOLUME
1,710.00	1,723.00	701,300
MARKET CAP	DAY RANGE	52 WEEK RANGE
214.387B	1,649.00 - 1,716.00	1,204.00 - 2,944.00

+ ADD TO WATCHLIST

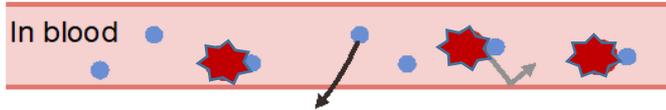
● MARKET CLOSED  
AS OF 12/02/2022 EDT



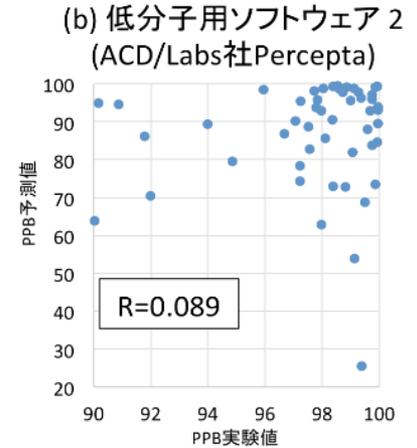
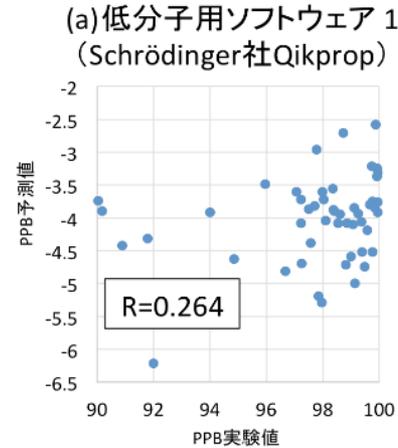
ペプチドリーム株式会社 ホームページより引用

# 最強のはずだったペプチド創薬が持つ問題

## 問題点① 体内持続性が低い場合がある

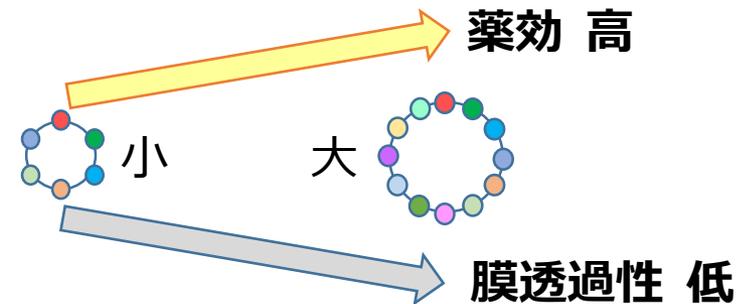
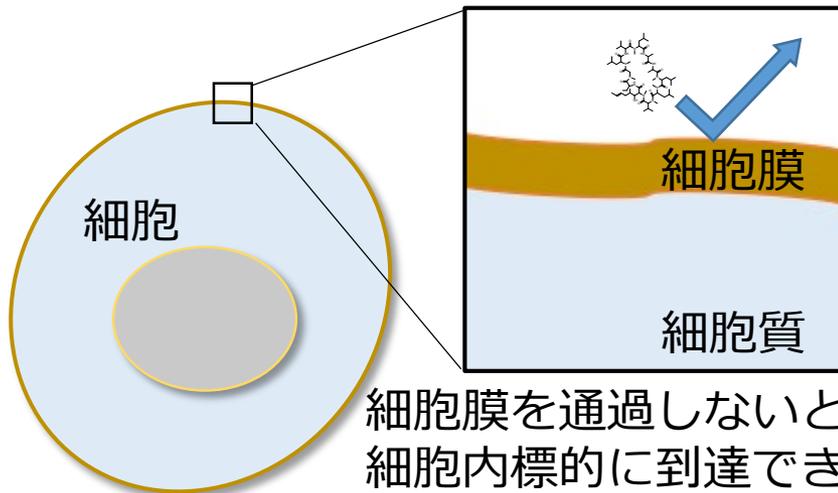


血中タンパク質結合率 (PPB) が従来薬よりも異常に低く、短時間で体内から分解・排泄される場合がある



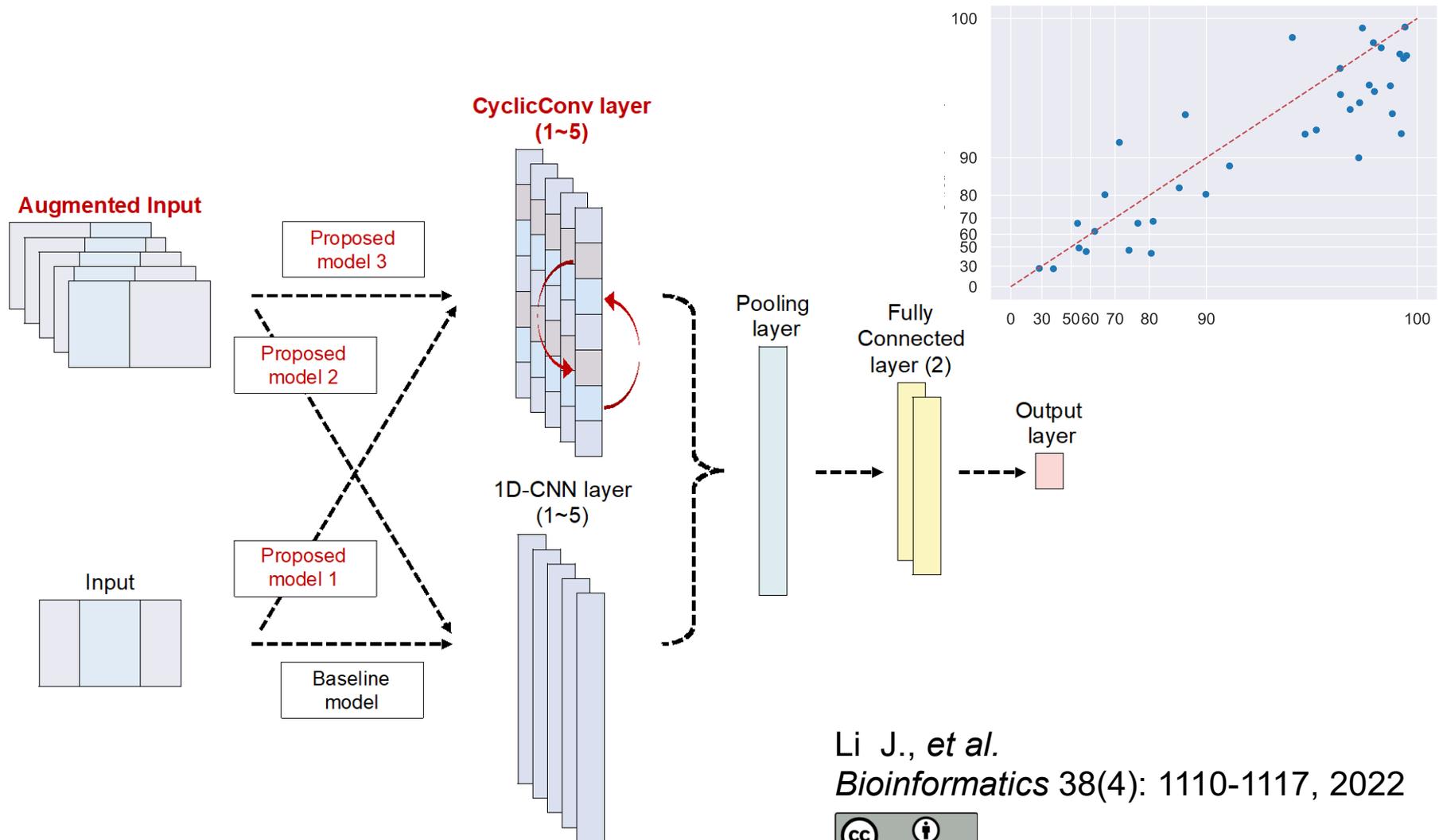
しかも、既存手法では予測ができず、合成してみても問題点が発覚する

## 問題点② サイズが大きいため、細胞膜を通過しにくい



しかも、薬効の高い大型ペプチドほど膜透過性の良いものは見つけにくい

# 2) 機械学習じゃダメなんですか？

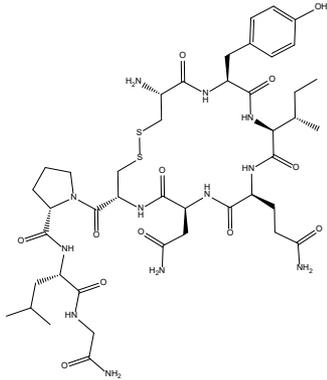


Li J., et al.  
*Bioinformatics* 38(4): 1110-1117, 2022



# 機械学習による予測モデルの構築

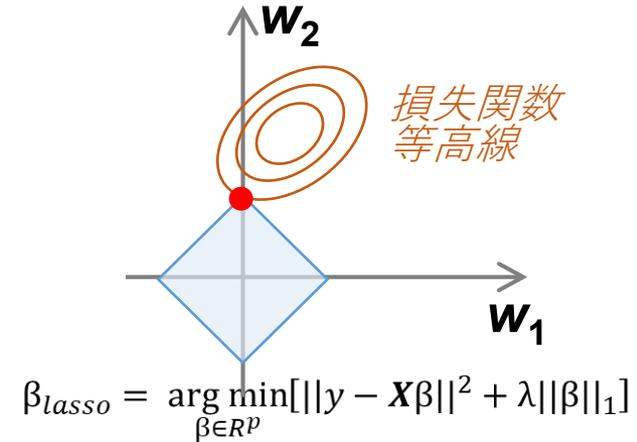
## ペプチド構造



## 特徴ベクトル

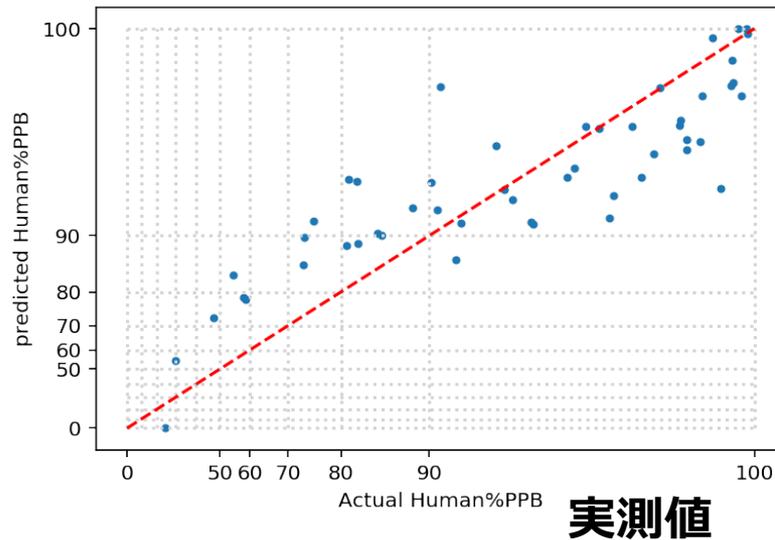
MW  
LogS  
LogP  
#Ring  
TPSA  
⋮

## 機械学習の技術



## 予測値

## 予測モデルの構築



有効なデータが  
多く集まれば  
精度が向上する

# 体内持続性(PPB)予測については開発済

Bioinformatics, 2021, 1-8  
<https://doi.org/10.1093/bioinformatics/btab726>  
Original Paper

OXFORD

Data and text mining

**Plasma protein binding prediction focusing on residue-level features and circularity of cyclic peptides by deep learning**

Li J., *et al.*

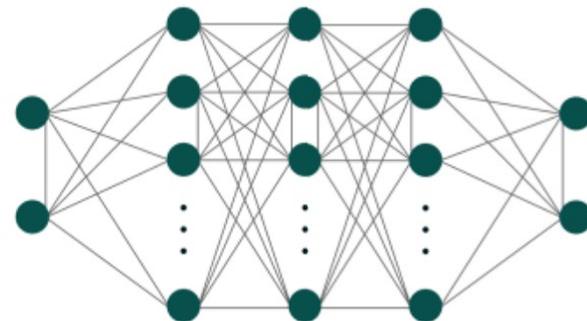
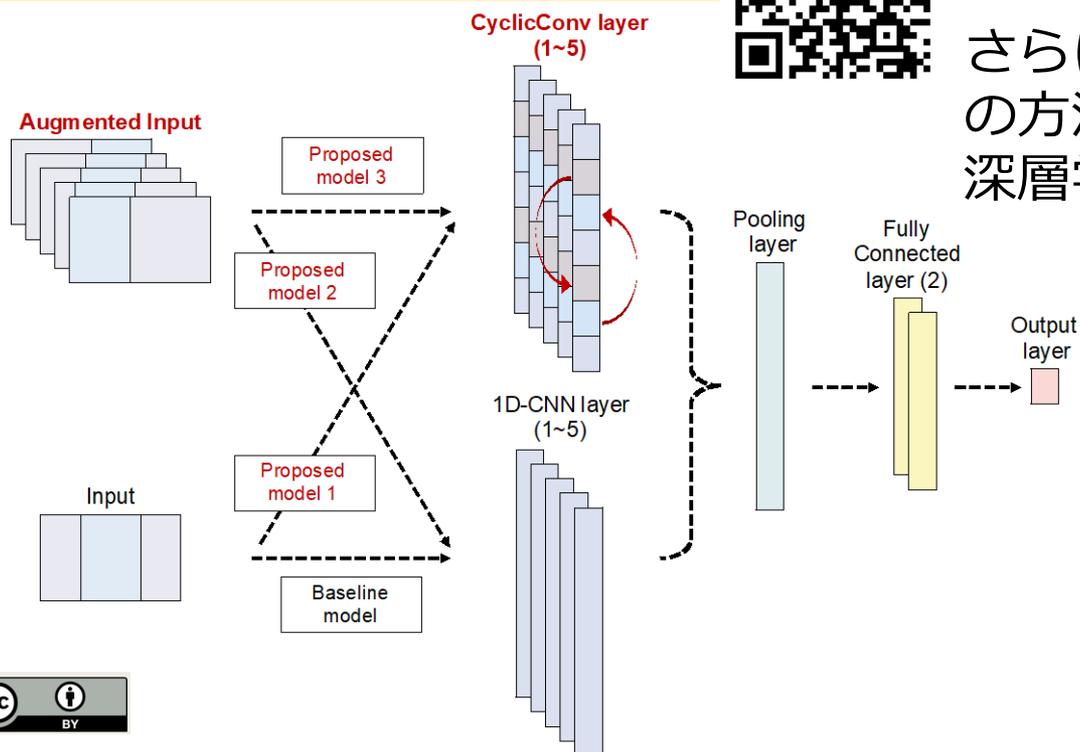
*Bioinformatics* 38(4): 1110-1117, 2022.



ペプチドリーム社との共同研究により、  
学習用のデータ量を確保できた

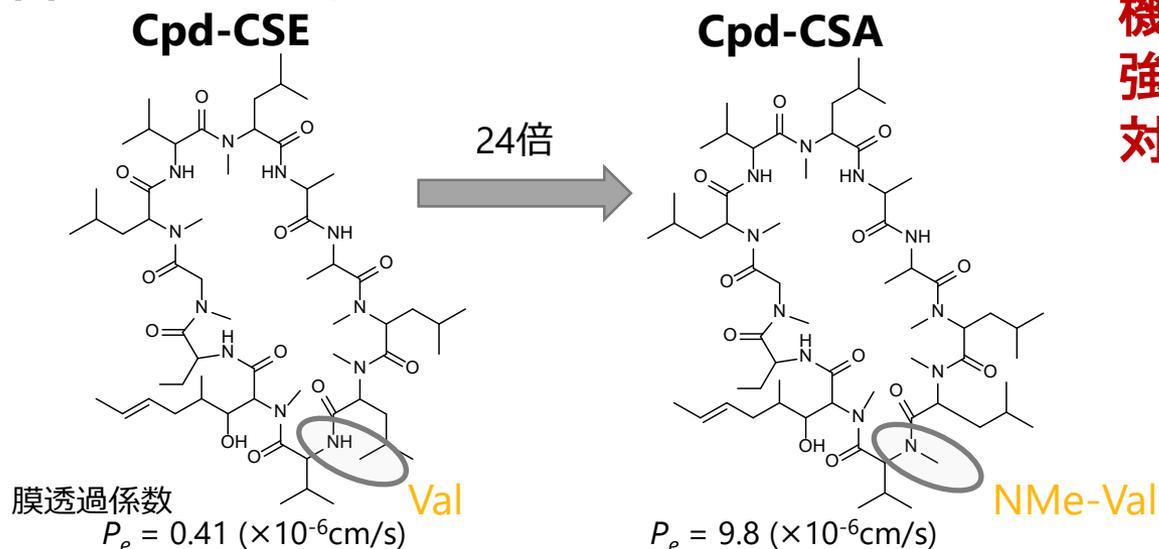
**環状ペプチドのPPBデータを  
500件近く収集**

さらにデータ拡張（データ水増し）  
の方法を開発したことにより、  
深層学習の適用が可能になった。



# 機械学習を阻む、化学修飾の多様性

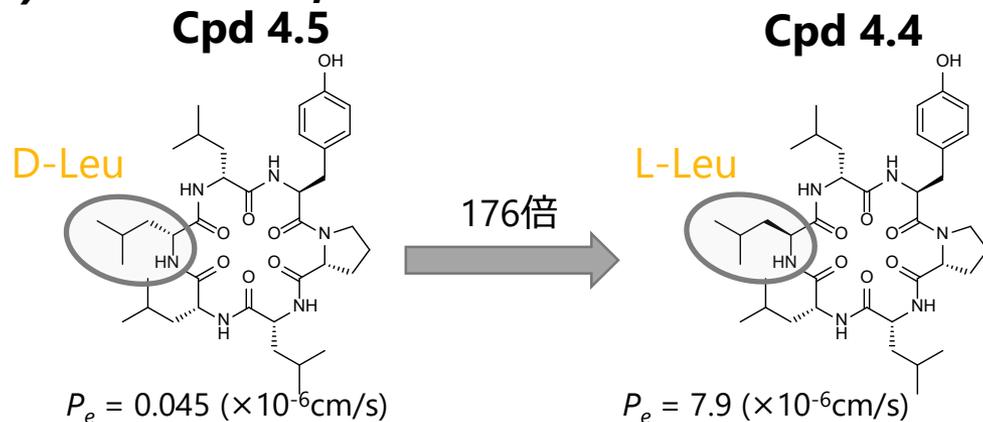
(a) Ahlbach CL, *et al. Future Med. Chem.* 2015



機械学習は設計シリーズ内では強力だが、新規パターンにまで対応できるとは限らない……

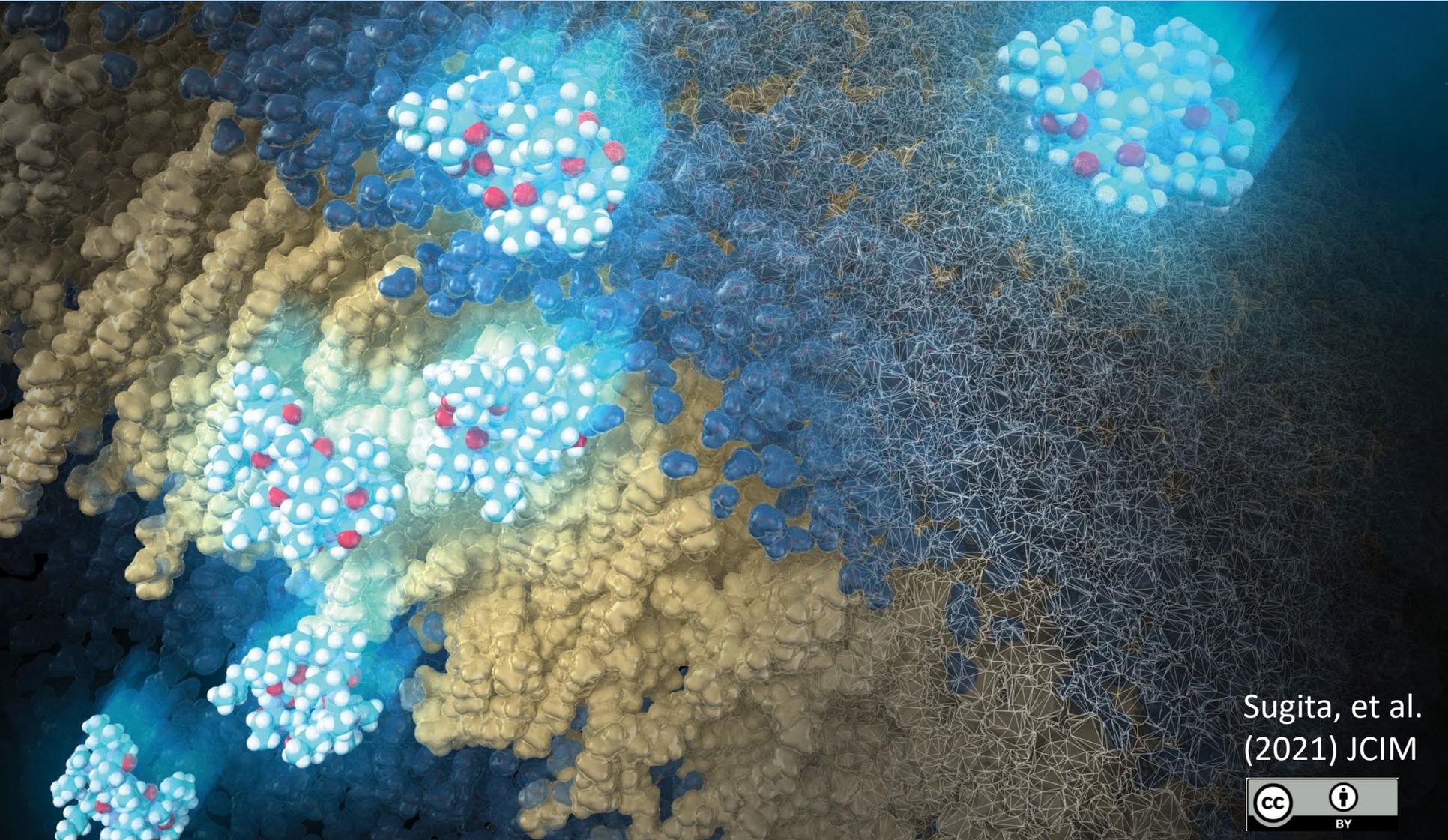
N-メチル化だけで  
**24倍**の膜透過率

(b) Hewitt WM, *et al. J. Am. Chem. Soc.* 2015



L/D 光学異性体で  
**176倍**の膜透過率

# 3) 分子動力学法による予測への挑戦 (第一次)

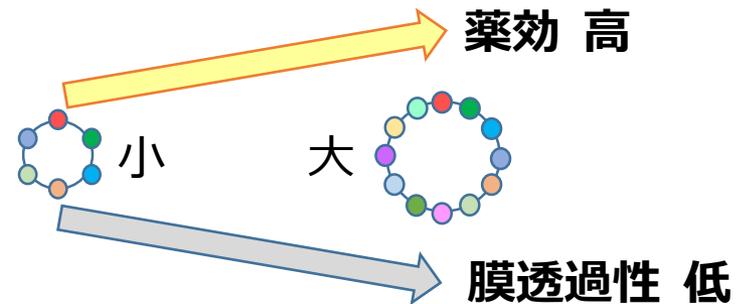
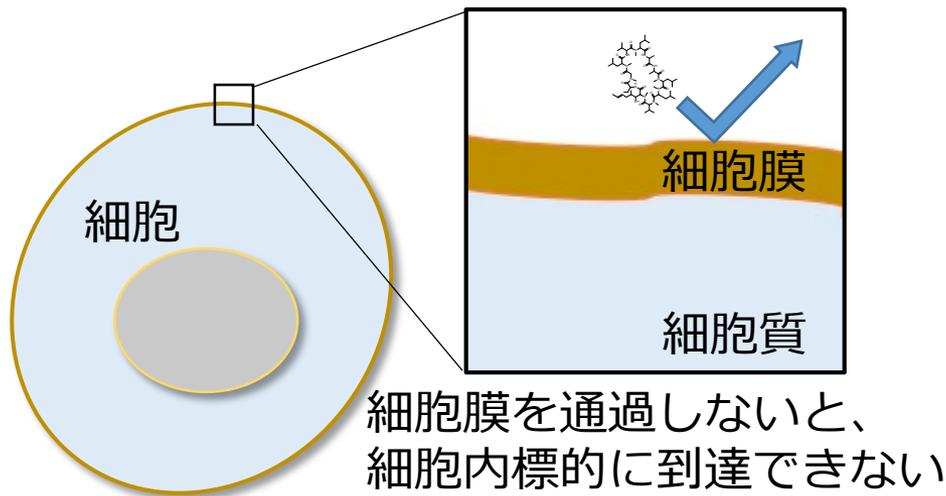


Sugita, et al.  
(2021) JCIM

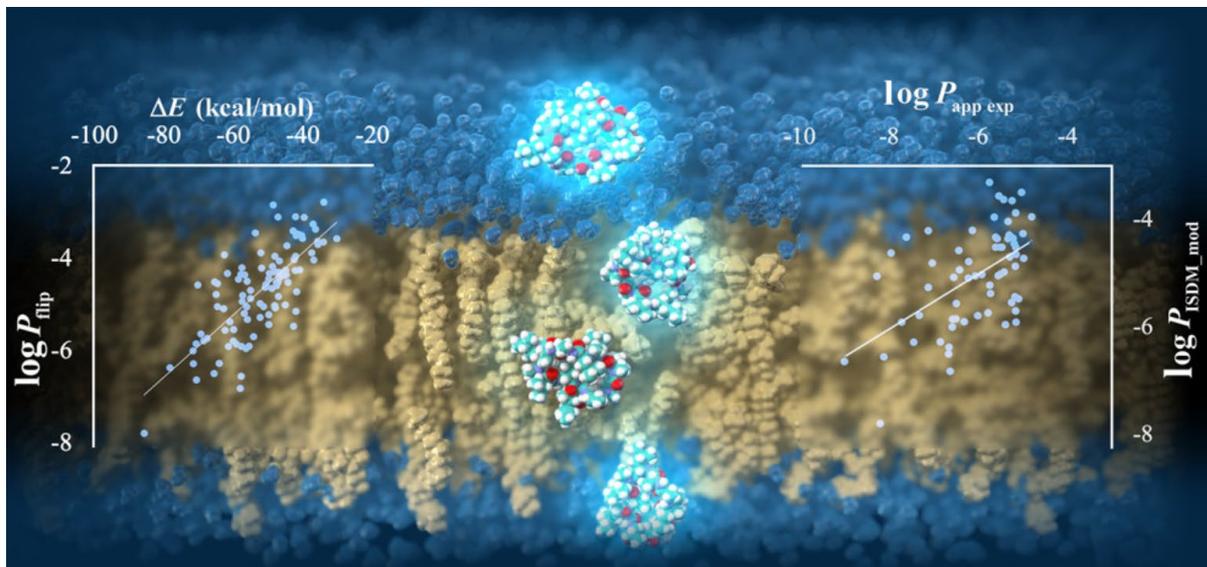


# 膜透過性はペプチド創薬の最大の未解決課題 19

**問題点** サイズが大きいため、細胞膜を通過しにくい



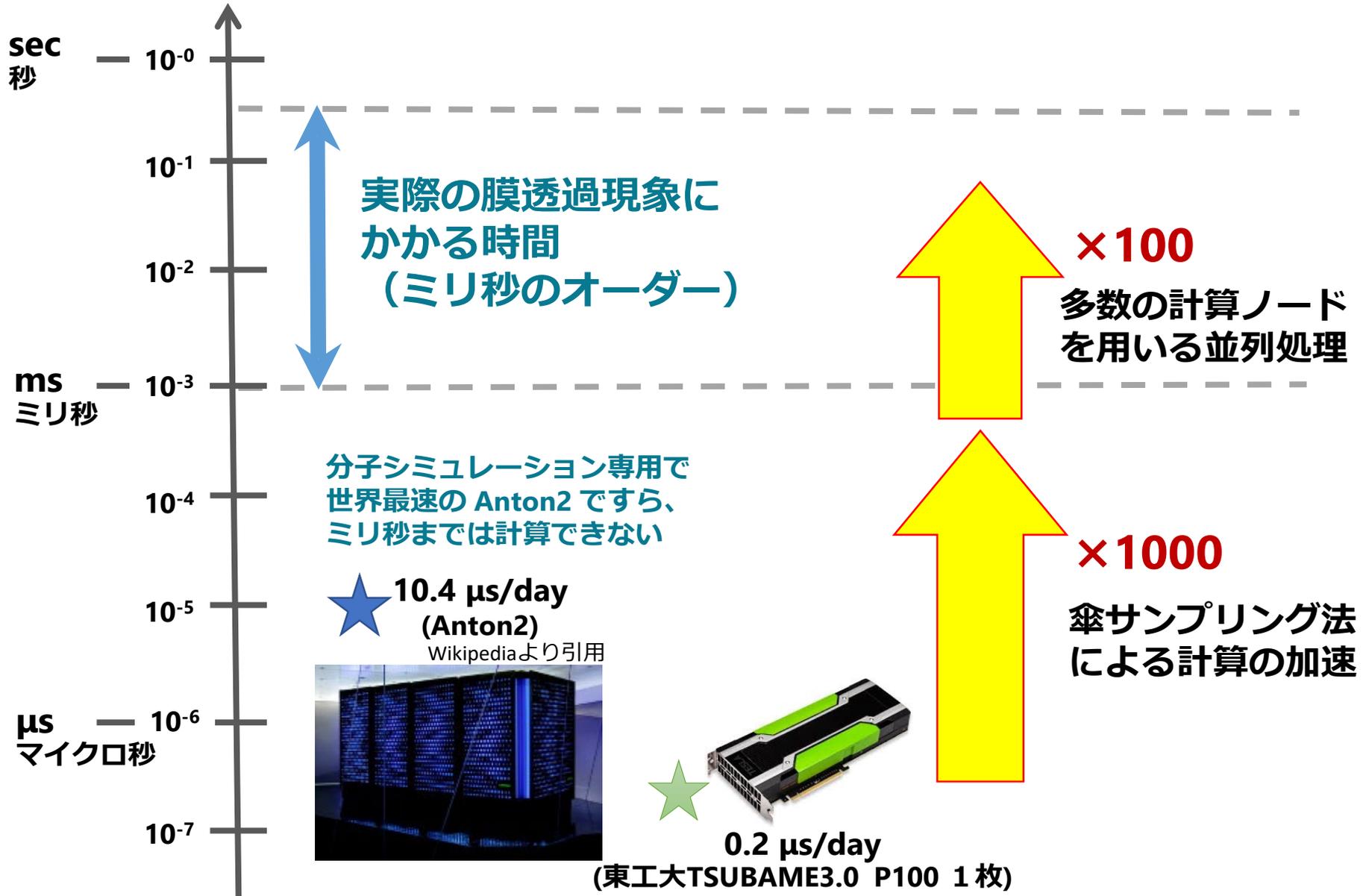
しかも、薬効の高い大型ペプチドほど膜透過性の良いものは見つけにくい



Sugita, et al.  
(2021) JCIM



# 膜透過シミュレーションは人類に可能か？

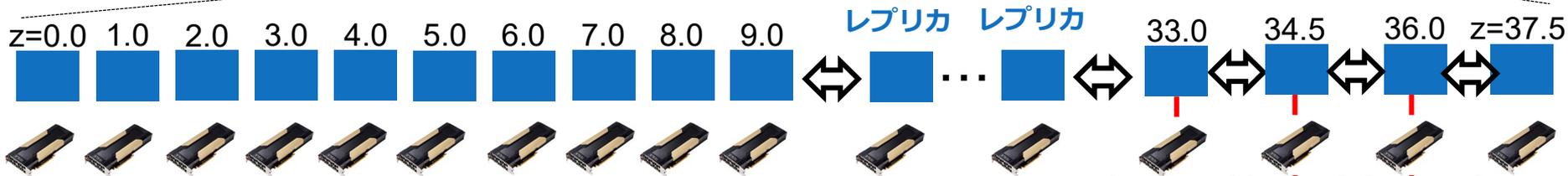
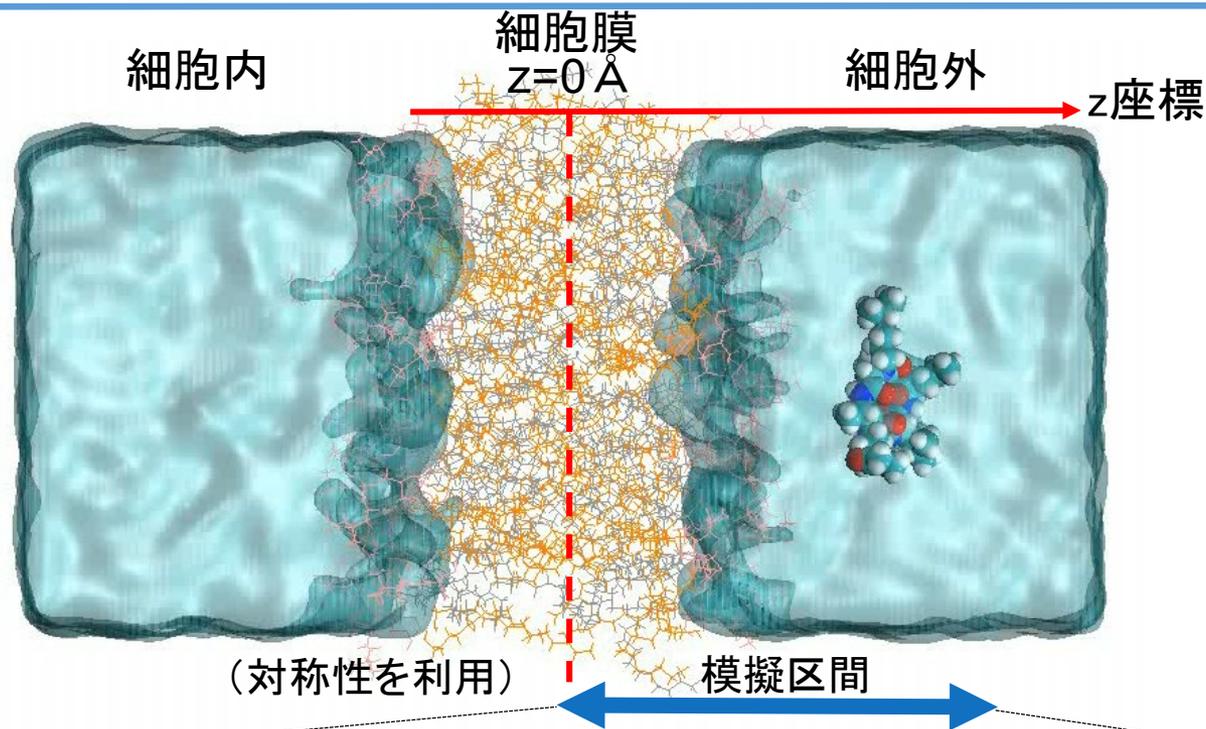


	<b>Steered MD (StMD)</b> (Izrailev, <i>LNCSE</i> 1999)	<b>Metadynamics (MTD)</b> (Laio, <i>PNAS</i> 2002)	<b>Supervised MD (SuMD)</b> (Sabbadin, <i>JCIM</i> 2014) <b>PaCS-MD</b> (Harada, <i>JCP</i> 2013)
<b>特徴</b> Descr.	人工的な力を加えて、観測したい <b>現象を誘発</b> するMD	エネルギー曲面にエネルギーを加えていき、観測したい <b>現象を誘発</b> するMD	標準的なMD条件下で、観測したい現象に寄与する変化のみを <b>優先選択</b> するMD
<b>長所</b> Merit	短時間(~数十ns)で計算可能	短時間(~数十ns)で計算可能、エネルギーの見積もりが可能	StMDやMTDと異なり、人工的なエネルギー操作を行わない
<b>短所</b> Issue	シミュレーション結果が、力の加え方に依存する	Collective variableの選択を誤ると、観測したい現象を誘発できない	出現しにくい現象の場合、計算時間が増大する



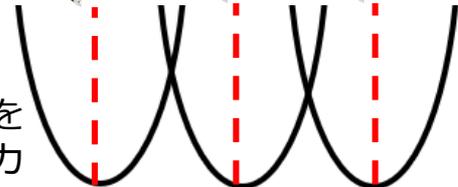
**StMD + REUSで対応することとした**

# 膜の法線方向に沿った レプリカ交換 (RE)

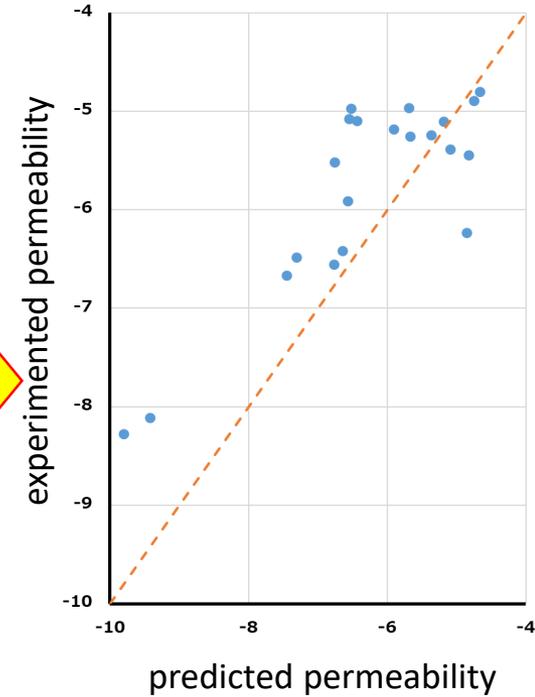
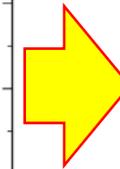
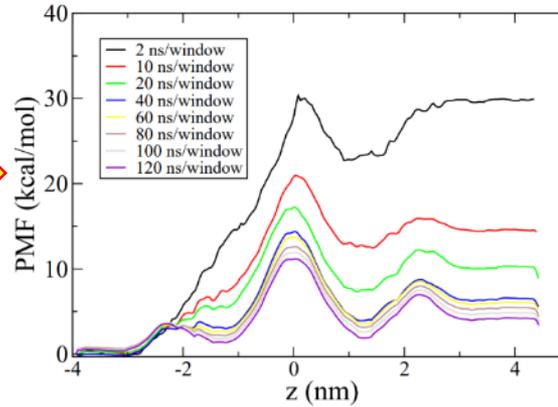
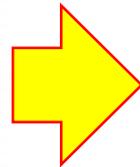
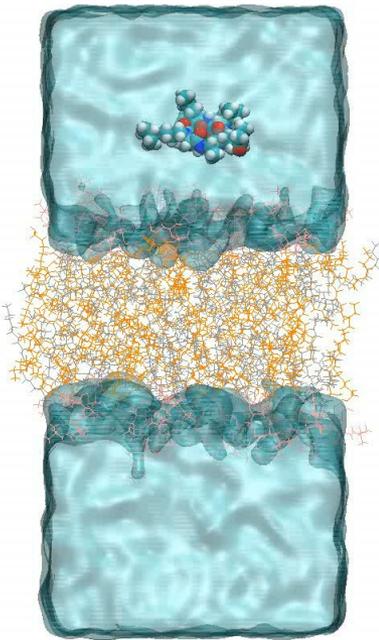


膜の法線方向 (z軸) に対して28地点を選択。  
膜中心の近くはz座標の間隔を細かく刻んでいる。  
**28レプリカを 28GPUを使って同時実行**

各位置に拘束をかけたレプリカ間で状態交換



# シミュレーションからの細胞膜透過率予測



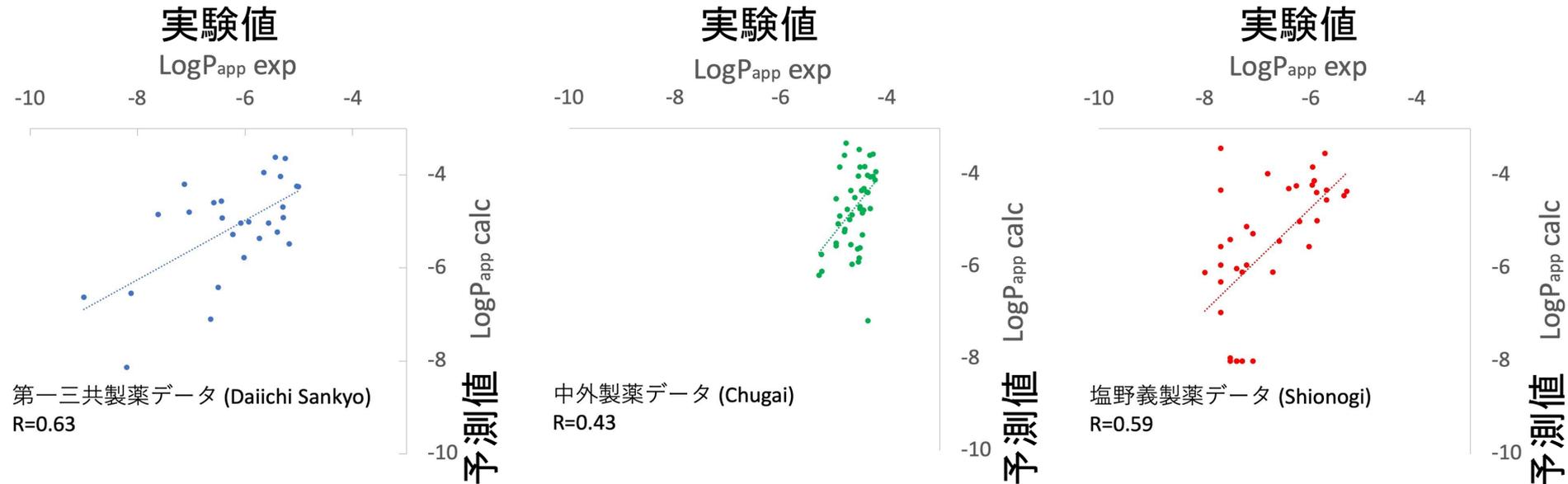
細胞膜透過現象を再現した  
分子動力学シミュレーション  
をスパコン上で実施する

稀にしか到達しない状態をつなげる  
傘サンプリング法とWHAMにより、  
平均カポテンシャルPMFを高速推定

$$\frac{1}{P} = \int_{z_1}^{z_2} \frac{\exp[\beta w(z)]}{D(z)} dz$$

PMFの推定値を利用しながら  
上記の方程式の右辺を計算し  
膜透過率 ( $P$ ) が求められる

# 実験値と計算値との相関（第一次プロトコル）



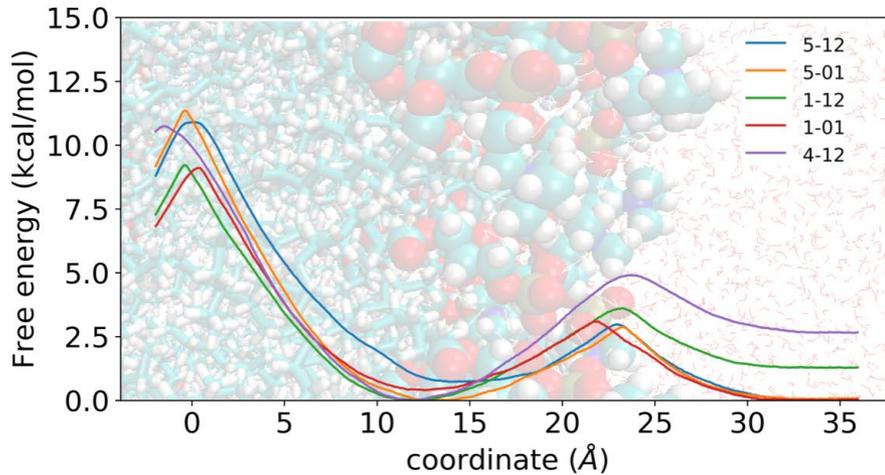
R=0.63 (D-S), R=0.43 (Chugai), R=0.58 (Shionogi)

膜を通りそうもない設計と、透過性のあるペプチドを、  
**化学合成や、評価実験なしに、計算機上でおおよそ弁別可**

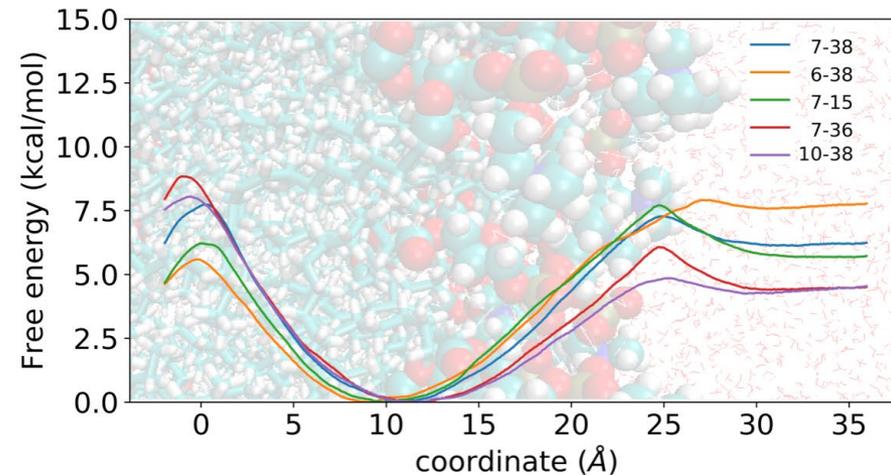
# 親水的 vs. 疎水的なペプチドのPMF形状の差異

25

(a) 親水的 Hydrophilic peptides



(b) 疎水的 Hydrophobic peptides



- 親水的 Hydrophilic peptides

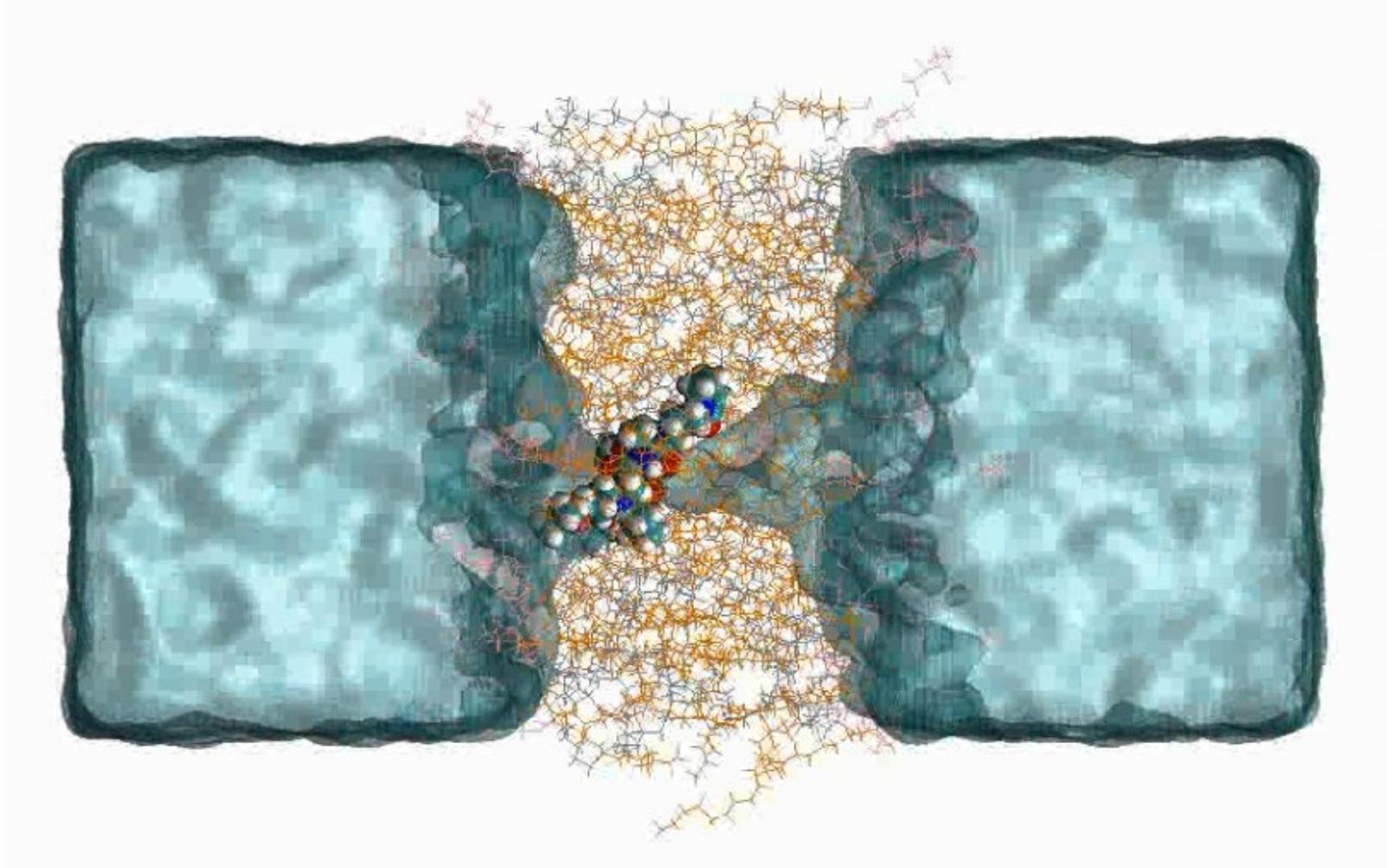
- high free energy barrier near the center of the membrane
- some peptides are most stable near the reaction coordinate  $z=10\text{\AA}$
- some peptides are most stable in bulk solution

- 疎水的 Hydrophobic peptides

- the barrier near the center of the membrane is no longer rate-limiting for several peptides
- all peptides are most stable near the reaction coordinate  $z=10\text{\AA}$

# シミュレーションから見えてくること

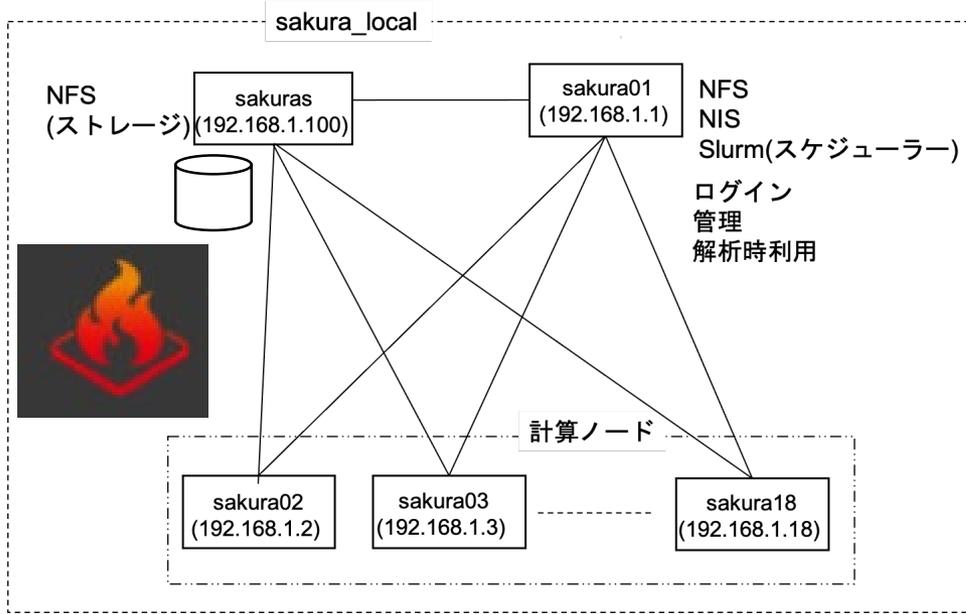
26



# 3つの環境を駆使してシミュレーション実施

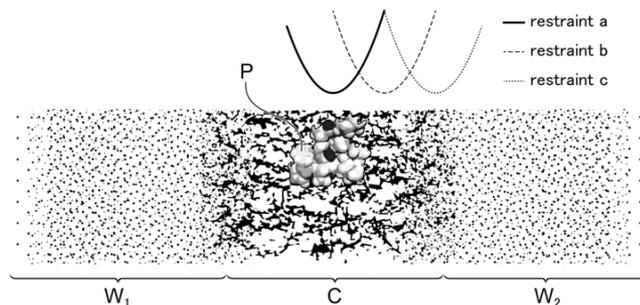
	REUS本体 28レプリカ	US 後処理 ×28	速度比
<b>東工大 T3.0</b> <b>P100 GPU</b> 	17.3 h × 4回 (400 ns) =69.2 h (3日弱)	3.5 h ×28個 /6GPU ≈16.3 h (1日弱)	×1.00
<b>さくら高火力</b> <b>V100 GPU</b> 	11.2 h × 4回 =44.8 h (2日弱) <i>big</i> キューで実施	2.3 h ×28個 /6GPU ≈10.7 h (半日弱) <i>small</i> キューで実施	×1.55
<b>産総研 ABCI</b> <b>V100 GPU</b> 	10.5 h × 4回 =42.0 h (2日弱) PCIeとSXM2の性能差?	2.2h ×28個 /6GPU ≈10.3 h (半日弱)	×1.65

# 24時間計算体制 (さくらインターネットとの共同研究)



計18ノード (V100×36)

- ・サーバー sakura01
  - ・クライアント sakura02-18
- Slurm 20.11.3 でのジョブ管理  
NFSマウントによるストレージ共有

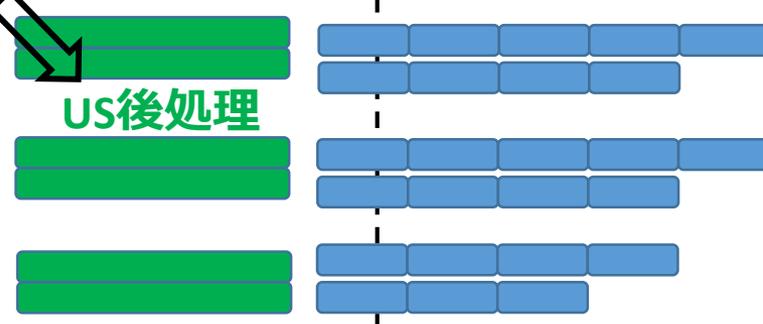
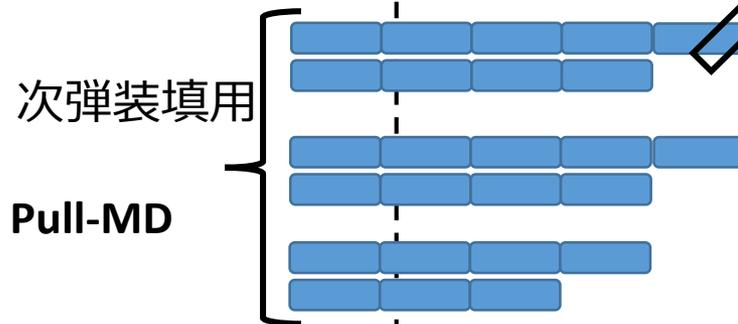


**big キュー**  
sakura02-15  
(28GPU)

**small キュー**  
sakura16-18  
(6GPU)

**28レプリカでのREUS計算**  
100ns x 4~5回

**28レプリカでのREUS計算**  
100ns x 4~5回



## 課題名：大規模REUSシミュレーションと機械学習の融合による 創薬向け環状ペプチド100件以上の細胞膜透過性予測

実施日： 2020年8月26日12:30 ~ 27日12:00 (23:30)

利用資源： **ABCI 1088ノード** (Large採択)



(a)環状ペプチドの膜透過率予測

**910ノード利用 (v100×3640枚)**、出力データ量 10.3TB

**製薬企業 (第一三共、中外、塩野義) 由来の100件以上の環状ペプチドの計算**

**→ 目標値の100件を上回る、130件超の膜透過シミュレーションを実施**

$$910 \times 4 \div 28 = 130$$



(b)性能計測 (strong/weak-scaling性能、ベアメタルとSingularityの比較、他)

**178ノード利用 (v100×712枚)**

**→ Singularity利用で、対ベアメタルで約90%の速度性能を達成**

## Large-Scale Membrane Permeability Prediction of Cyclic Peptides Crossing a Lipid Bilayer Based on Enhanced Sampling Molecular Dynamics Simulations

Masatake Sugita, Satoshi Sugiyama, Takuya Fujie, Yasushi Yoshikawa, Keisuke Yanagisawa, Masahito Ohue, and Yutaka Akiyama\*

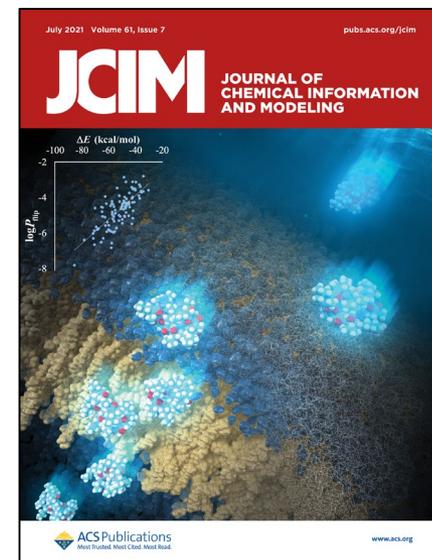
6残基×100ペプチド

8残基×56ペプチド

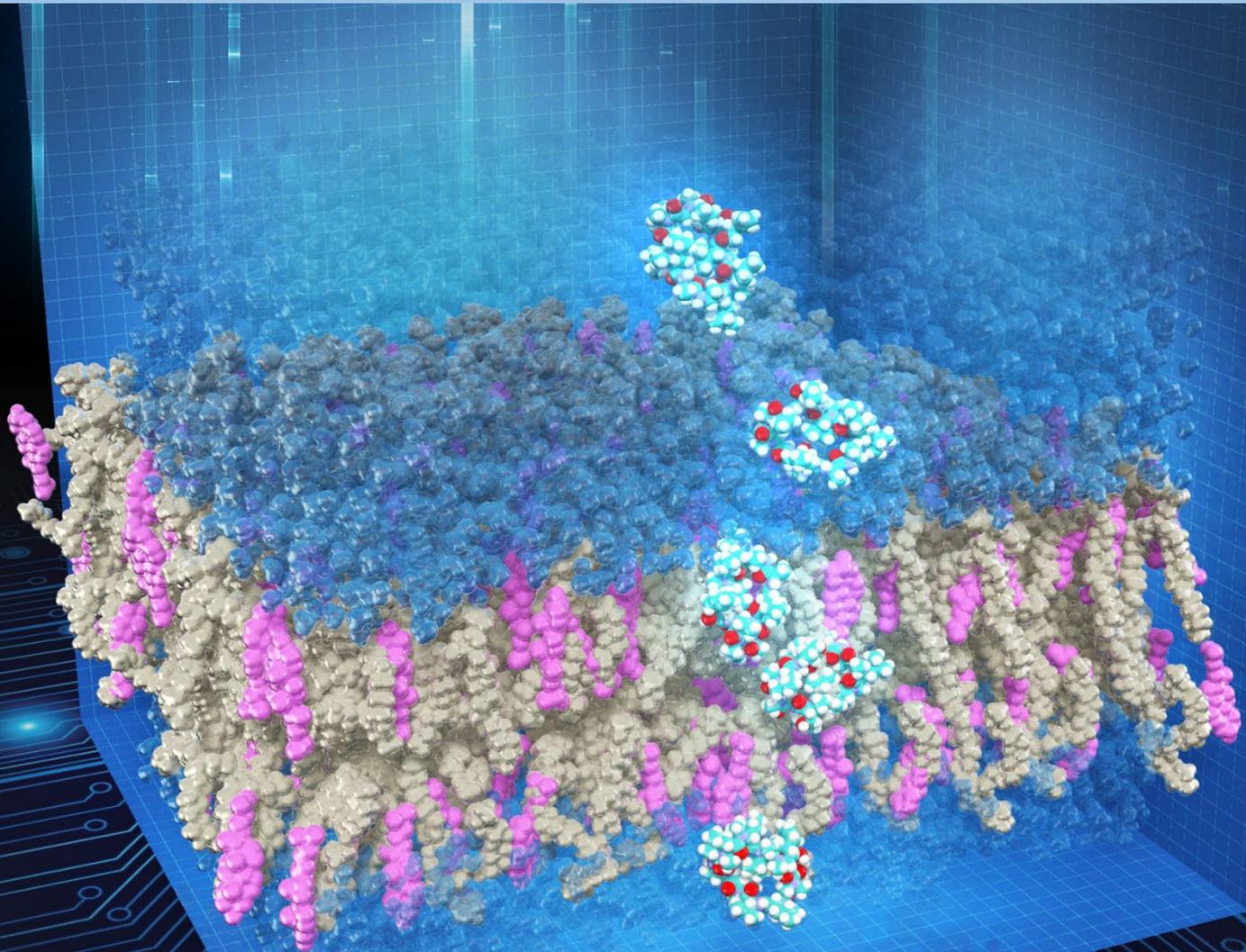
約30万GPU時間 (P100)

2021年7月電子公開

8507 view (2022年12月5日現在)



# 4) 分子動力学法による予測への挑戦 (第二次)

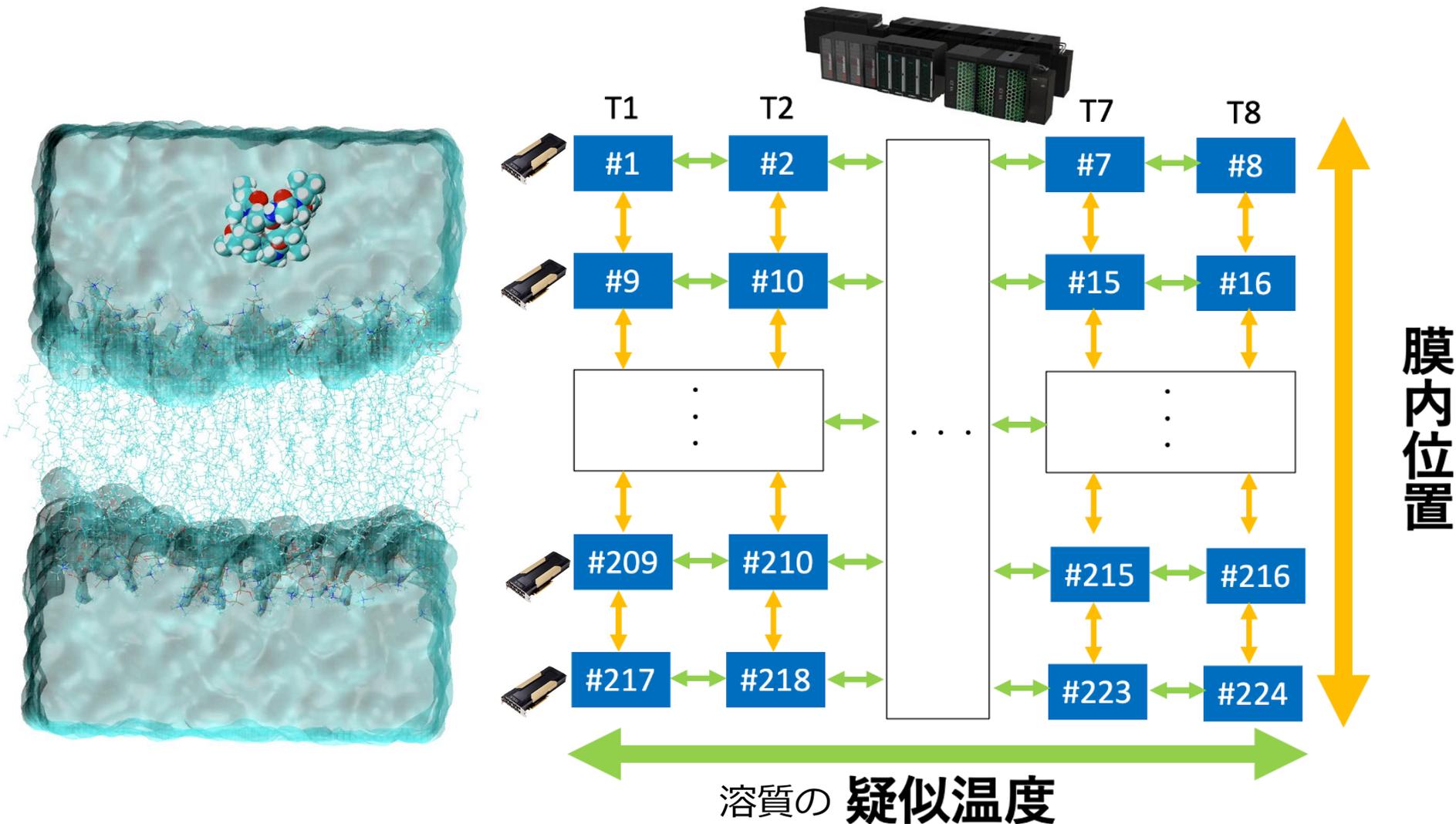


Sugita, et al.  
(2022) JCIM



# REST/REUS法の採用、二次元レプリカ交換へ

32

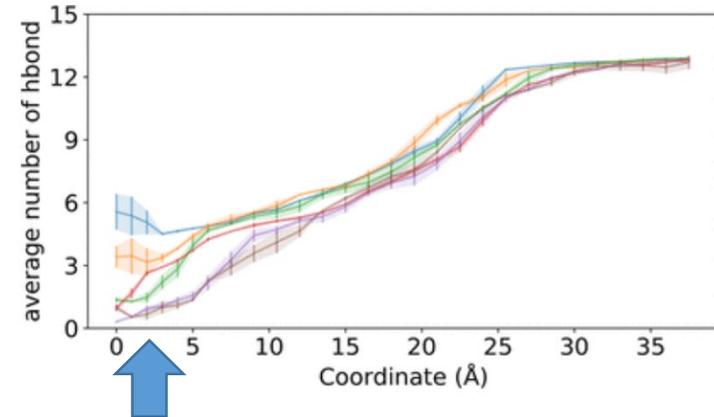
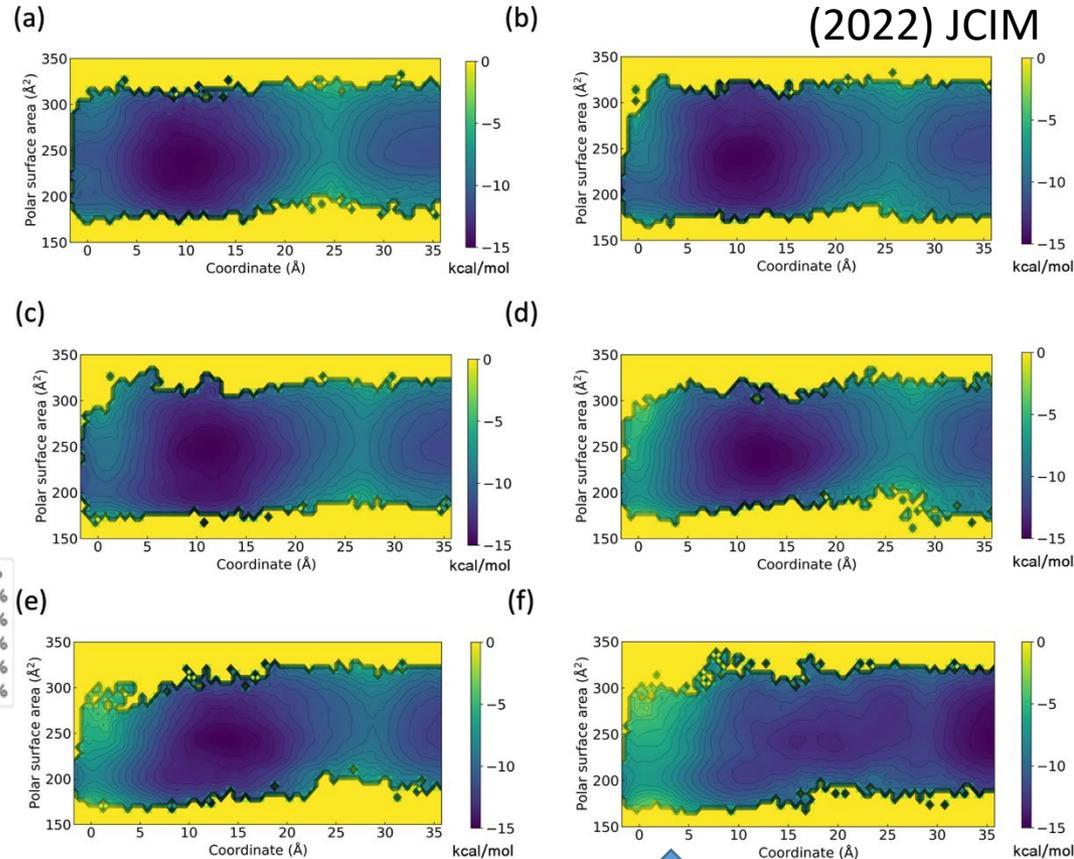
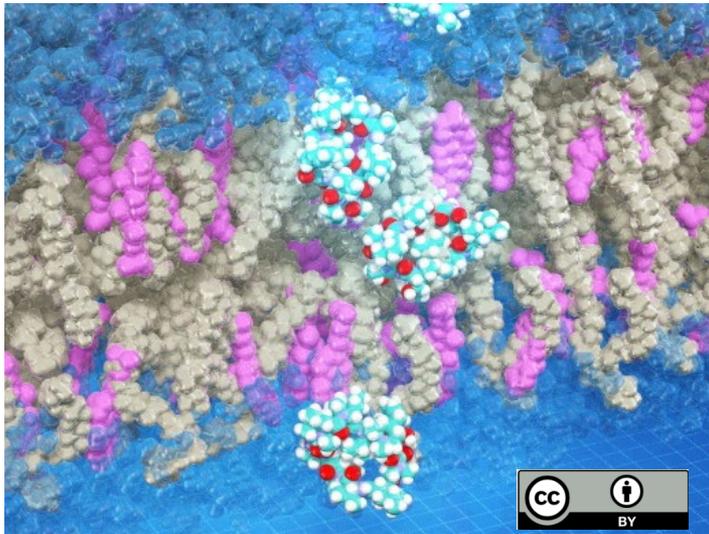


各ペプチド毎に、

224GPU × 約90時間 (P100の場合) で 500ns 程度のシミュレーションを実施

# 脂質二重膜にコレステロールを添加

Sugita, et al.  
(2022) JCI



40 ~ 50 mol %の組成では  
膜中心 ( $z=0 \text{ \AA}$ ) 付近まで  
水分子が浸入しにくくなる

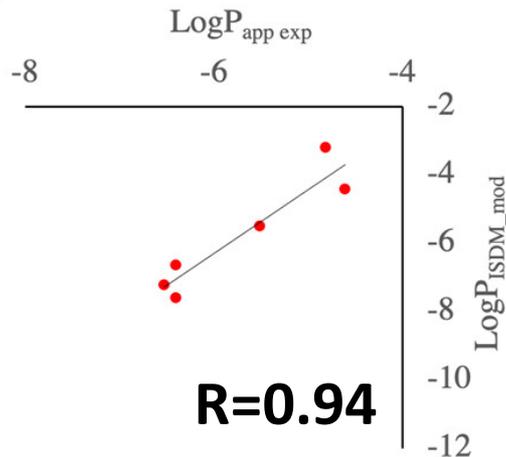
(a) 0 mol %, (b) 10 mol %,  
(c) 20 mol %, (d) 30 mol %,  
(e) 40 mol %, (f) 50 mol %

40~50 mol %は  
横軸  $z=0 \text{ \AA}$  付近で  
表面積が縮小

# 実験値と予測値の相関 (第二次プロトコル)

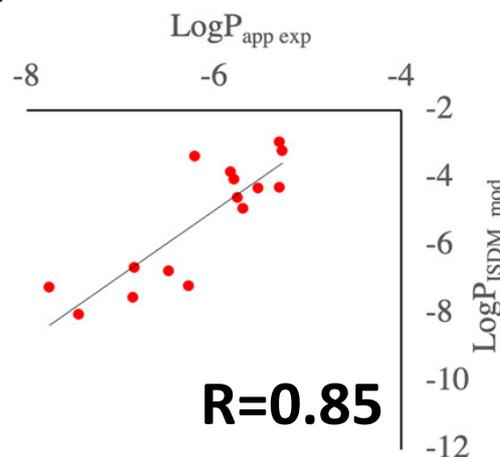
Sugita, et al.  
(2022) JCIM

(a) MDCK (イヌ腎臓由来)



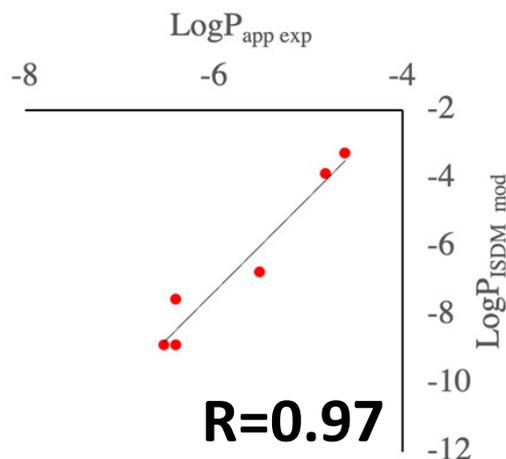
コレステロール  
40 mol %

(b) PAMPA (人工膜)



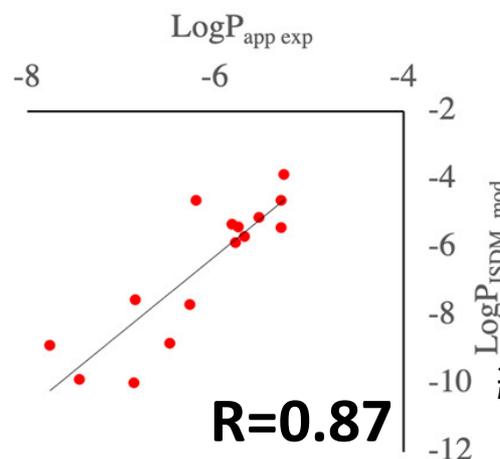
10残基ペプチド  
18個について  
膜透過率実験値  
との相関を示す。

(c)



コレステロール  
50 mol %

(d)



縦軸が予測値

## Lipid Composition Is Critical for Accurate Membrane Permeability Prediction of Cyclic Peptides by Molecular Dynamics Simulations

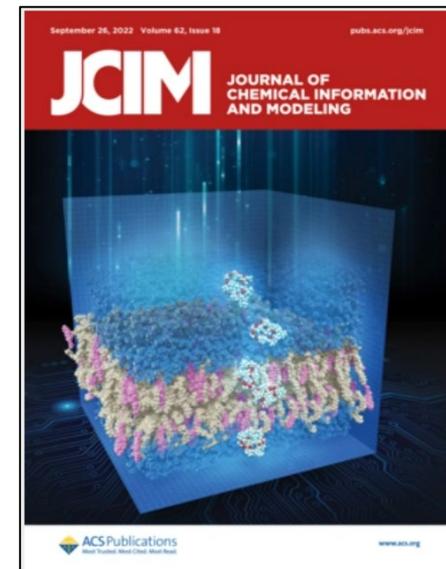
Masatake Sugita, Takuya Fujie, Keisuke Yanagisawa, Masahito Ohue, and Yutaka Akiyama\*

コレステロール濃度比較 (0,10,...50 mol%)  
及び10残基×18ペプチド (40 & 50 mol%)

約70万GPU時間 (P100)

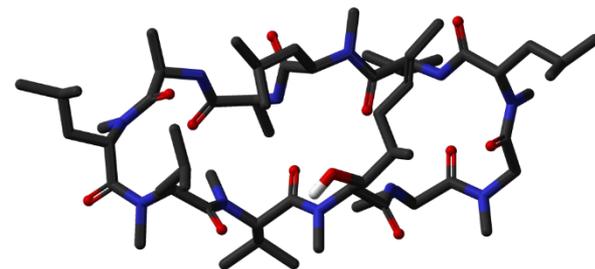
2022年9月電子公開  
1635 view (2022年12月5日現在)

特許7057004, 特許7125575



# 5) まとめ

- 環状ペプチドの細胞膜透過現象に対する直接的な分子動力学 (MD) シミュレーションを実施。
- 従来研究は「水中のふるまい」「油性環境中のふるまい」などを単独に調べるだけであったが、本研究では実際に**脂質二重膜と水の層を用いて計算**。ただし膜透過現象はミリ秒前後であるため、十分な状態サンプリングが致命的に重要。
- 現プロトコルでは、膜と鉛直方向 (z軸) に刻み方を変えた**28地点**を取り、さらに各点に**8種類の溶質温度**を設けた**REUS/REST法による二次元レプリカ交換**分子動力学シミュレーションを実施。 **$28 \times 8 = 224$ GPUを同時稼働**。シミュレーションは**約500ナノ秒**ずつ行い、PMF曲線の収束を確認している。
- 10残基の環状ペプチド18個について、コレステロール 40 mol% を加えた膜モデルを使ったとき、**実験値と予測値の相関係数は 0.94 (MDCK), 0.85 (PAMPA)**である。今後さらに多くの大型環状ペプチドで、実験値との相関を確認する必要がある。



## 6) 今後の展望

- **1ペプチド当たりのシミュレーションのコストが相当に高い**  
東工大TSUBAME産業利用（成果公開）  $100\text{円} \times 56\text{ノード} \times \text{約}90\text{時間} = \text{約}50\text{万円}$   
産総研ABCI（V100利用）  $220\text{円} \times 56\text{ノード} \times \text{約}60\text{時間} = \text{約}75\text{万円}$
- 28箇所×8溶質温度は丁寧すぎるかもしれない。  
サボれるところがないかを研究していく（ために、もっとたくさん計算する）。

**コストが現在の1/10（数万円）になれば、ペプチドを合成し、膜透過実験のために外注している現状に比べて大幅に安価になるため、利用価値が上がる。**

（もちろん合成や実験の分野においても日々コスト改善は続いているものの **HPC分野における指数関数的なコスト改善のペースをなめんなよ・・・**）。