

## 質疑応答

Q. アミノ酸を8次元で表現をされていたりしましたが、もう少し多かったり少なかったりも可能かと思います。そのあたりは今後ご検討されるのでしょうか。

Ans. 今回の発表では触れませんでした。全てのペプチドに対して予測を行う前にいくつかの表現方法を検討しております。次元数は最も少ないもので3次元、多いものでは10次元です。検討方法は表現方法ごとにMDデータ319種類の一部で学習を行い、残りのデータの中の最も自己組織能の高い配列をどの方法が最も良く予測できるか確認しました。

Q. 予測ができたものを最後、配列レベルで考察をされていたりしましたが、ご使用された指標から、何か言えることはあったのでしょうか。

Ans. 今回の指標として例えばAPスコアはシミュレーション前後の表面積の比をベースとして計算しています。この指標では自己組織化・凝集すると表面積が小さくなるような配列群は探索できますが、大きく隙間が空いた状態で自己組織化するような配列は探してこれません。そのため今回高い自己組織化能と予測された配列は今回の指標に適合するもののみで、他にも別の指標を用いることで見つかる高い自己組織化能を持つ配列もあるかと考えております。

Q. 今回の計算で発見された配列は天然では利用されているのでしょうか？

Ans. 今回の計算で発見された配列は天然での利用は報告されていないようです。6残基以上の長さになってくると報告例が増えてきますが、6残基を網羅的に探索するには配列の種類が20の6乗=64,000,000通りとなり今回の400倍であるため、また新たな工夫が必要となってきます。

Q. 今回は10nsの計算で凝集するものが見つっていますが、講演の中でも触れられているように、タンパク質が一晩かけて凝集するなどのゆっくりした凝集現象も起こります。このように、凝集速度に非常に大きな差があるのはなぜなのでしょう？

Ans. 凝集速度の違いはペプチド・タンパク質を構成しているアミノ酸の物性や凝集後の構造の安定性などの複数の要素が関わってきます。また、温度やペプチド・タンパク質の濃度も影響してきます。今回の計算は短時間のシミュレーションで凝集をみるために、低温(280K)かつ高濃度(200 mM)でシミュレーションを行っております。

▶ [【北日本支部】2020年度オンラインシンポジウム「情報科学を駆使して生命分子を見る・知る・使う」](#)