



Title	深層学習を活用したガラスの構造緩和を決定する特徴量を抽出する技術の開発
Author(s)	金, 鋼
Citation	サイバーメディアHPCジャーナル. 2024, 14, p. 63-65
Version Type	VoR
URL	https://doi.org/10.18910/96526
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

深層学習を活用したガラスの構造緩和を決定する

特徴量を抽出する技術の開発

金 鋼

大阪大学 大学院基礎工学研究科

1. はじめに

多数の分子が集合した複雑分子系、柔らかな分子系、ソフトマターなどと称される物質は高分子・液晶・ガラスなどを含みさらに水そのものや生体膜まで対象を拡大したものを総称しており物理、化学、生物、工学を幅広く跨る学際領域を形成している。特に、分子が凝縮することによってガラス転移で見られる遅いダイナミクスを示す共通性質に着目し、遅いダイナミクスの本質を理解することを基礎的問題として捉え、ソフトマターにおける協調運動や自己組織化の統合的な解明に注力している。そこで、ガラスの遅いダイナミクスに対する分子動力学(MD)シミュレーションによる研究をおこない深層学習の融合研究を実施した。

2. 研究内容

2.1 説明可能な深層学習による反応座標探索

タンパク質の構造変化などの複雑分子系において、安定状態をつなぐ遷移経路を特徴づけることは遅いダイナミクスの時定数を特定する上で重要課題である。そのため原子位置に関する高次元の配置座標から決まる集団変数 Q に関する確率分布関数 $P(Q)$ をサンプリングし、その対数を取ることで平均力ポテンシャル(potential of mean force; PMF) $F(Q) = -k_B T \ln P(Q)$ が計算される。PMF は自由エネルギー地形とも呼ばれ、安定状態は鞍点により区別され、さらに実際の経路が鞍点を通過するとき、変数 Q は構造変化を特徴付ける反応座標といえる。本研究ではアラニンジペプチドの立体配座が変化する異性化過程に着目する。真空中ではエネルギー的に安定な2状態であ

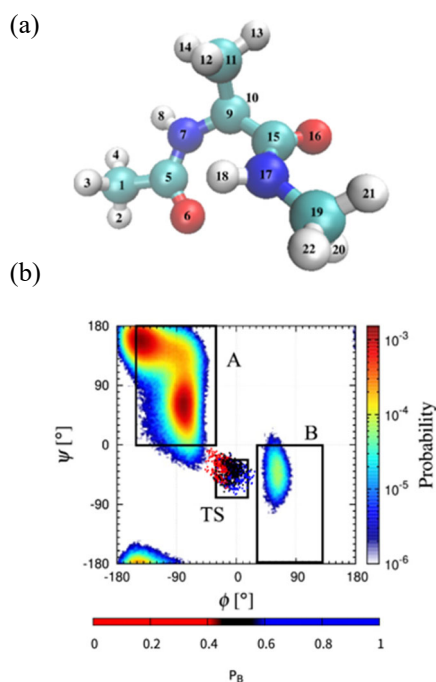


Fig 1. (a) Alanine dipeptide molecule and number of each atom. (b) PMF with dihedral angles ϕ, ψ as variables. Committor p_B is widely distributed and the dividing line is not clear.

る β シート構造(状態 A)と α ヘリックス構造(状態 B)が存在することが知られ、遷移状態(TS)を特徴付ける候補変数として二面角(ϕ, ψ)が重要視されてきた(Fig. 1)。先行研究では MD 計算と機械学習を用いた二面角の系統だった探索方法が提案されている。さらに原子間距離も集団変数となり得ることを考慮すると、異性化過程のより精緻な記述が期待できる。そこで、原子間距離を候補変数として説明可能な深層学習を行うことでどの原子間距離が適切な反応座標となり得るかを明らかにすることを目的とした。

2.2 深層学習によるガラス形成液体の温度変化に伴う構造変化の解明

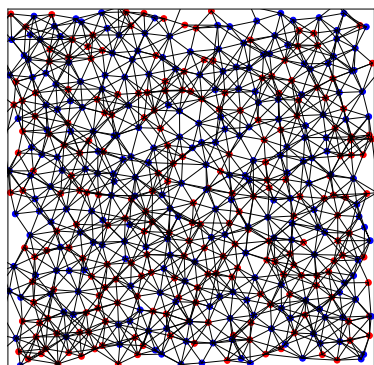


Fig 2. Particle configurations and corresponding graph structure.

液体を冷却しても結晶化が阻害されると、融点以下の液体状態が維持されガラスになる。ガラス形成液体の構造緩和時間は温度の低下とともに急激に遅くなりガラス転移温度で非晶質構造が凍結する。このガラス転移現象をもたらす何らかの特徴的な静的構造があるかという点については多くの研究がされている。近年、機械学習を用いてガラス形成液体の構造における特徴量を抽出する研究が多数報告されている。注目すべきは、畳み込みニューラルネットワーク(CNN)やグラフニューラルネットワーク(GNN)といった最新の深層学習技術が応用されていることである。例えば、GNN を用いればガラス形成液体の構造と動力学を学習し、従来の機械学習手法より高い性能予測を示すことが報告されている。さらにCNN による画像認識を応用すると、ガラスと液体の構造分類とその分類根拠となる特徴的構造が抽出されることも報告された。本研究ではGNN によって異なる2つの温度のガラス形成液体の構造を分類することを目指した。

3. 計算手法

3.1 説明可能な深層学習による反応座標探索

真空中のアラニンジペプチドの異性化過程についてMD 計算を行い、TS 近傍の分子構造2000点をサンプリングした。それぞれを初期状態とし1 ps のMD 計算を100回実施し、状態Aに辿り着く前に状態Bに遷移する確率をコミッター p_B として定量化した。入力変数として初期状態の化

学結合を除いた原子間距離を用い、コミッター p_B を出力変数として、シグモイド関数 $p_B(q) = [1 + \tanh(q)]/2$ に回帰することで適切な反応座標 q の学習を行った。なお用いたニューラルネットワークは5つの隠れ層から構成され、奇数層、偶数層それぞれが400ノード、200ノードを持っている。

3.2 深層学習によるガラス形成液体の温度変化に伴う構造変化の解明

ガラス形成液体のモデルとして3次元2成分soft-sphere モデルを採用し、MD シミュレーションをおこなった。GNN はグラフを扱うための深層学習手法であり、グラフとは、要素をノード、要素間の関係をエッジとして構成されたデータ構造である。GNN はグラフデータを入力とすることを活かし、入力グラフのノードとエッジの情報の更新を繰り返し、自ら特徴量を作り出すことができる。得られた粒子配置について、粒子をノード、第一配位圏以下の粒子間のつながりをエッジとしてモデル化したグラフ構造を構築した(Fig. 2)。粒子の種類と粒子間の相対座標をグラフに符号化することで、異なる2つの温度の構造分類(二値分類)をおこなった。

4. 結果

4.1 説明可能な深層学習による反応座標探索

深層学習による p_B のシグモイド関数への回帰結果をFig. 3(a)に示す。相関係数0.903となり、反応座標 q を獲得できたことを意味する。ただし深層学習モデルにおいて複数回の非線形結合が行われていることから、入力変数の q への寄与を直接的に求めることは困難である。そこで深層学習に対して解釈性を与えることのできる「説明可能なAI」(XAI)の一手法であるLIME(local interpretable model-agnostic explanations)を適用した。その結果、最も寄与の大きい入力変数として原子6-10間の距離 r_{6-10} 、二番目に寄与が大きい入力変数として r_{6-11} が得られた。特定された2つの変数を用いてPMFを描画すると、遷移状態

をあらわす $p_B \sim 0.5$ が状態 A と状態 B を適切に区別する分割線をなすことがわかった(Fig. 3(b)) [1]。

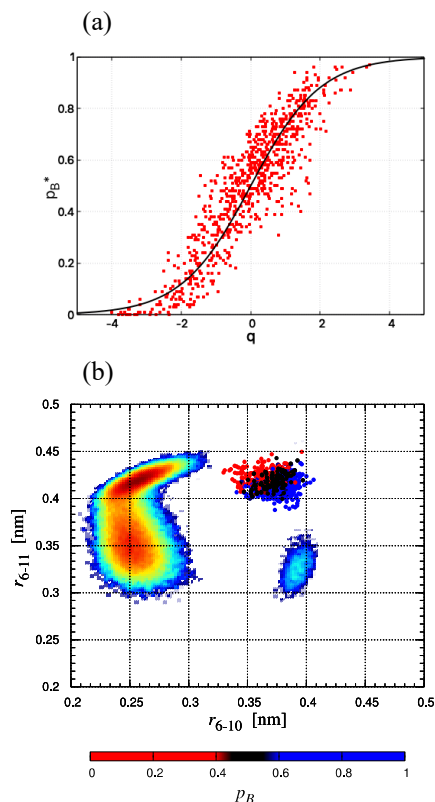


Fig 3. (a) Regression results from deep learning (b) PMF of interatomic distance r_{6-10} , r_{6-11} indicated by XAI. We can distinguish between states A and B and see that the saddle point is bounded by a structure in the transition state $p_B \sim 0.5$.

4.2 深層学習によるガラス形成液体の温度変化に伴う構造変化の解明

GNN は、ガラス転移温度近傍の構造と温度の高い液体構造の違いを 100 %識別できることがわかった。ただし、一方で高温の液体構造同士は識別できないことがわかった。このことは温度の変化にしたがって、ガラスと液体とで微細な構造変化を GNN は学習していることを示唆する。また Attention により算出された粒子間の重要度について高温では粒子間の重要度にばらつきが見られ、低温になるにつれてその重要度が概ね均一になることがわかった。さらに Attention により抽出された寄与度が高い粒子は、ボンド秩序変数 Q_6 が小さい傾向にあることを見出した。以上のことから、GNN はガラス形成液体の温度の違いを

Q_6 が小さく局所的に結晶秩序の度合いが小さい粒子を根拠に分類していると考えられる [2]。

5. おわりに

本研究課題では、ソフトマターに対する MD シミュレーションに深層学習を応用する融合研究を展開した。それぞれの研究テーマは個別的であるが、MD シミュレーションによる時間変化する原子配置データセットから次元削減による物性の説明を目指す高度化された研究になりえるものと考えている。

参考文献

- (1) K. Okada, *et al.* Unveiling interatomic distances influencing the reaction coordinates in alanine dipeptide isomerization: An explainable deep learning. <http://arxiv.org/abs/2402.08448> (2024).
- (2) K. Yano, *et al.*, in preparation.