

課題名(タイトル): 深層学習などによる NMR データおよび生体高分子の解析

利用者氏名: ○小林直宏

理研における所属研究室名: 生命機能科学研究センター先端 NMR 開発・応用研究チーム

1. 本課題の研究の背景、目的、関係する課題との関係
HOKUSAI のような大規模 PC クラスタの利用は、NMR のデータ解析の自動化および高度化、深層学習ツールの開発などへの応用を目的とした。現状では NMR 解析法は他の構造解析法として X 線結晶解析やクライオ電顕解析などに比べて大きく遅れを取っている。研究代表者は HOKUSAI に搭載されている 40~200 Intel-Core を利用することで大規模な分子動力学計算を試験的に実行し、その一方で分子運動と関連する NMR データ解析との整合性解析を行った。

2. 具体的な利用内容、計算方法
分子動力学計算ツールである Amber (20-40 core) によるレプリカ交換法による大規模な並列計算に関して小型蛋白質を対象とする計算を実行した。このような大規模計算には十分なストレージサイズとコア数を必要するため、HOKUSAI による計算実行は可能であった。

3. 結果
レプリカ数を増やした計算は効果的であったものの、ノード間通信速度に限界があるため実質的な計算時間はローカルで構築された GPU 搭載計算機に及ばなかった。当初の目的では小型蛋白質がミリ秒ほどの時間スケールで運動する様子を原子レベルで考察し、溶液 NMR で得られる動的情報との整合性を解析する予定であったが、十分な議論を得るデータになり得なかった。並列化で多数のレプリカを計算させる際にパラメータ交換を実行した後の緩和処理に問題があったか、あるいは他のレプリカ交換法に類した方法 (REST, REST2 など) を試す必要性が在ったのかもしれない。

4. まとめ
今年度の初期において標準的なレプリカ交換法を用いた分子動力学計算を実行したが十分に実験データを議論できるデータを得る事が出来なかった。しかしながら、レプリカ交換法における問題点や REST 法あるいは加速型 MD 計算といった他の手法との比較は計算手法確立に大いに役立ったと言える。

5. 今後の計画・展望

今後は Amber 以外の分子動力学計算ツールとして Gromacs, Genesis などを用いてより先進的な並列型分子動力学計算 (REST2, gREST など) を用いて、 μ 秒~ミリ秒に及ぶより広範な時間スケールでの運動性解析を行い、溶液 NMR データから得られる動的情報との関連性に関して有意義な解析結果が得られる事を期待している。更に、得られた情報をよりコンパクトな計算機でも再現可能にすべく機械学習の訓練データとして活用する事を計画している。

6. 利用がなかった場合の理由