

課題名(タイトル):

電子顕微鏡による単粒子解析

利用者氏名:

○竹本 千重(1), 横山 武司(1), 加茂 友美(1) 長内 隆(1), 赤坂 領吾(1)

理研における所属研究室名:

(1) 生命機能科学研究センター タンパク質機能・構造研究チーム

<p>1. 本課題の研究の背景、目的、関係するプロジェクトとの関係</p> <p>生命科学の分野において、生体高分子の複合体の高分解能構造情報を得ることは、それらの生理的機能や制御機構を知る上で重要である。クライオ電子顕微鏡解析は、NMR解析やX線結晶構造解析と並んで、そのための主要な手法のひとつである。近年、電子直接検出器の開発によって、著しく感度と分解能が向上し、試料調製からデータ収集において、飛躍的なハイスループット化が進んでいる。計算機自体の性能向上と共に、取得した二次元画像から三次元構造を得るためのプログラムも日進月歩で改良が進んでいるため、精度の高い構造を得るためのスキームの確立とアウトプットの迅速化が重要な課題となっている。また、実験的に得られた高分解能の立体構造情報は、分子シミュレーションに供して動的構造解析を行うことにより、細胞内の環境に近い状態での反応機構の予測や検証へと展開できるようになってきている。</p> <p>実際、クライオ電子顕微鏡による高分解能解析を用いる研究においては、ターゲット分子(複合体)の設計、調製方法の検討、取得画像の処理、構造モデルの構築、動的構造解析など様々な局面で計算機による予測や解析が必要となっている。本課題では、これらの諸段階で必要となる計算機環境を、研究室でのローカル環境と比較・検討を行いながら模索する。また、すでにスパコン上に整備されているプログラム(AMBERなど)の利用や、ローカルで計算資源の不足が生じた場合の利用を想定していた。</p>	<p>高いパッケージプログラムである Relion を Hokusai にインストールした。他に、ローカルで利用しているプログラムとしては、EMAN2, Gautomatch, CrYOLO, Scipion, Frealign, Cinderella, Spider, cryoSPARC, Phenix, CCP-EM などがある。</p> <p>3. 利用がなかった場合の理由</p> <p>画像処理プログラムの多くは、取得画像へのアクセスが必要なため、作業ディレクトリに置いてジョブを展開する。1枚の画像サイズは約 260 MB (@1.47 Å/pixel)で、テストデータの場合は 200 枚程度、高分解能解析のためには 1000-2000 枚(0.3-0.5 TB)の画像を使用する。例えば Relion を使って、100 MB の 3D 像(単粒子)の解析を行った場合、途中で生成するファイルを合計すると、最終的に約 1.7 TB 程度になる。ひとつの構造決定に至るまでに同様なデータ処理を複数並行して行うことになるため、5-10 TB のワーキングスペースが必要になる。また、現在主流となっている画像をクラス分けする手法や、三次元画像の再構成のステップにおいては、GPUの利用が計算時間の観点から必須である。さらに、ノード当たりのメモリが 64 GB では足りないステップもあるなど、一連の解析を行う環境を HOKUSAI 上で確保できないため、作業効率の観点から利用が難しい状況となっている。</p>
<p>2. 具体的な利用内容、計算方法</p> <p>公開されているプログラムを単独またはパッケージでサーバーにインストールし、主に GUI ベースで利用している。現時点では、ターゲットによって適するプログラムが異なることもあるが、もっとも使用頻度の</p>	<p>4. 今後の計画・展望</p> <p>スパコンの将来計画に GPU が含まれないことから、今後も扱うデータ量の増加が予想される単粒子解析のプラットフォームとしての利用は難しそうである。今後は、特定のステップを CPU で処理するプログラムについて、インタラクティブモードでの利用が現実的に可能かどうか検討したい。</p>