平成 27 年度　全体研究開発報告書

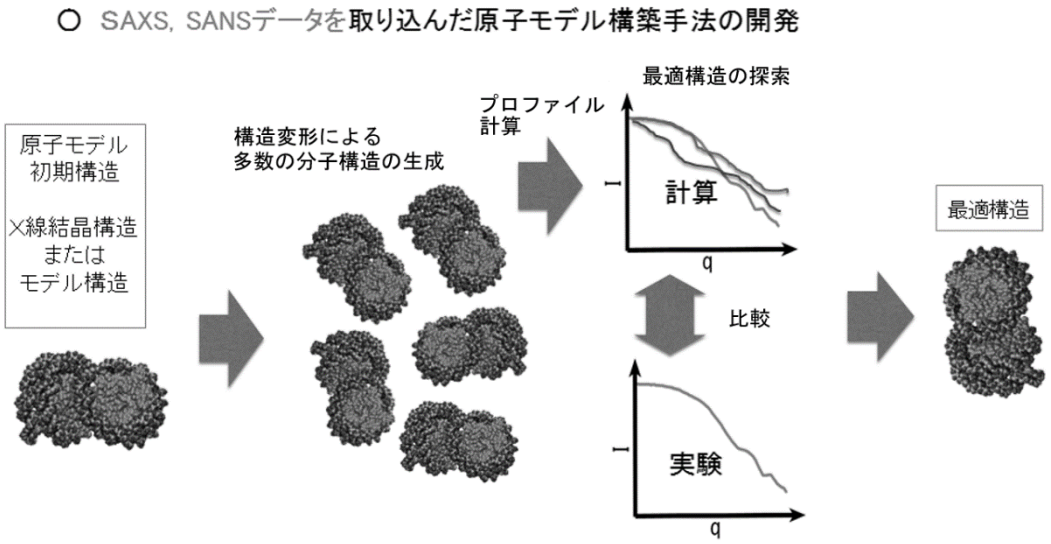
１．補助事業名：創薬等ライフサイエンス研究支援基盤事業
（創薬等支援技術基盤プラットフォーム事業）
２．補助事業課題名：タンパク質の複合体構造・相互作用に関する総合的な予測・解析の実施と高度化
（超分子モデリングと重要アミノ酸残基の推定）
３．研究開発代表者：国立研究開発法人量子科学技術研究開発機構
量子ビーム科学研究部門関西光科学研究所
量子生命科学研究部生体分子シミュレーショングループ
グループリーダー　河野　秀俊
４．研究開発の成果

本研究は、タンパク質の複合体構造・相互作用に関する総合的な予測・解析の 実施と高度化（代表機関：東京大学、清水謙多郎）の分担として、「高分解能立体構造情報と低分解能立体構造情報を統合し、異なる分子状態での原子モデル構築や超分子のモデリングとアミノ酸置換による分子の構造や機能の変化予測について支援とその方法論の高度化を実施」している。今年度は、以下の支援と高度化の研究を実施した。

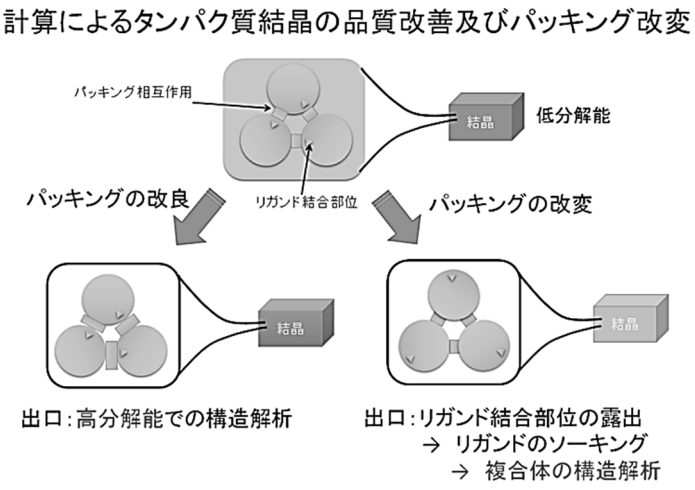
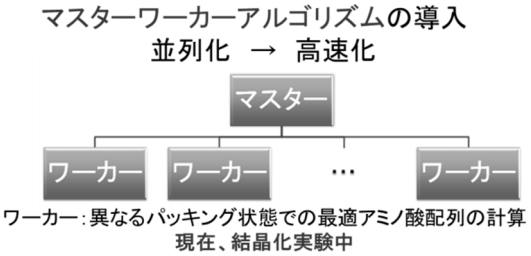
支援と高度化

１．独自に開発した分子構造変形技術と、電子顕微鏡像などの低分解能解析データへの原子モデル構造あてはめ技術を用いた、分子のさまざまな状態での原子モデルの構築支援と高度化

これまでの、電子顕微鏡データに適合する原子モデルの構築方法を高度化し、X 線小角散乱データ(SAXS)および中性子線小角散乱データ(SANS)に適合する原子モデルの構築に成功した（下図)。この方法では、まず、ある一つの構造から出発して、コンピュータにより、物理化学的に無理のない小さな構造変形を繰り返すことで、多数の候補構造を作る。次に、生成した各構造に対して、実験データと等価な物理量を計算して、最も実験データに適合した原子モデルを選択する。本方法を、ヌクレオソーム(DNA とタンパク質の複合体で、細胞核内での DNA 収納の基本構造体）が連結した分子の構造解析に適用し、2 量体及び 3 量体のヌクレオソームの原子構造モデルを構築した。これらの構造は、細胞核内での DNA 構造状態やその変化さらにはタンパク質の DNA アクセスを考える上で重要な基盤情報となる。



２．DNA 物性の観点からの転写因子の ChIP-seq データ解析及び結晶内分子パッキング改変（薬剤など低分子のソーキングによって低分子とタンパク質の結合様式を調べるため）の支援と高度化

転写因子の認識モチーフ（結合する DNA 配列パターン）は、ゲノム配列中に複数存在するものの、実際に転写因子が結合する場所は、その一部である。我々は、結合位置前後の配列まで考えることで、STAT3 転写因子が結合する部位を高精度で予測することができることを示した。本研究は、STAT3 の ChIP-seq データの解析支援として実施した。結合位置前後の配列情報、DNA 構造、DNA 動特性の 3 つを考慮して、ChIP-seq データ結合部位を機械学習し、結合部位が予測できるかどうかを調べた。DNA の構造及び動特性を調べるため、すべての 4 塩基対配列パターン 136 種類をそれぞれ持つ 12 塩基対の DNA 配列について 100 ナノ秒の分子動力学計算を実行し、配列ごとに平均構造及び動特性解析を行った。それらの結果を用いて、転写因子の結合部位周辺の配列を解析したところ、DNA 動特性を考慮したモデルで機械学習したときが最もよく実験データを再現した。これは、DNA の動特性をタンパク質が認識して結合していることを示唆している。

結晶内分子パッキングの改変（下図左）では、あるタンパク質をモデルケースとして、解析拠点の研究者とともに研究を進めている。このタンパク質は、結晶化条件によって、モノマーもしくはダイマーとして結晶化することが知られている。そこで、アミノ酸残基の改変により、ダイマーの結晶を不安定化しモノマー結晶を安定化することができるかどうかを試みた。アルゴリズム上は、複数のアミノ酸残基を改変することができるが、実験による検証を考慮して、１残基置換に限定して計算を行った。すべての表面残基について、19 種類の他のアミノ酸残基に置換し、相対的な分子の配向及び分子間距離を変えながら分子構造の最適化計算を繰り返し実行し、もともとの結晶で見られる分子パッキングのエネルギー順位の変動を調べた。結果、ダイマーを不安定化しモノマーを安定化する置換体が複数見つかった。現在、それらのうち、エネルギー順位の変動が大きかった変異体について、実験的な検証を行っている。分子構造の最適化計算は、分子の配向及び相対距離を変えて行うため、非常に大きな計算量となる。そのため、マスターワーカーアルゴリズム（下図右、各ワーカーが一つの相対配置ないし変異体を担当）を導入し、CPU 数に反比例して計算時間が短縮できるように計算アルゴリズムを高度化した。