

データ収集・処理の自動化による 多数の微小結晶を用いた構造解析の効率化

山下恵太郎

理化学研究所 放射光科学総合研究センター

シンクロトロン放射光を用いた X 線結晶構造解析は、生体高分子の構造情報の蓄積に多大な貢献をしてきた。さらに X 線の高輝度化に加えて検出器の高感度化、タンパク質調製技術の進歩などにより、1つのタンパク質の構造決定に必要な時間やコストは大幅に削減されつつある。しかしながら、生命機能の解明に重要な膜タンパク質等の解析難度は高く、現状簡便に解析が可能な数 10 μm 以上の大きさの結晶を得ることは難しい。マイクロビームビームラインである SPring-8 BL32XU は、最小 1 μm ×1 μm のサイズの高強度ビームが利用可能であり、微小結晶からのデータ収集を現実的なものにした。10 μm 程度の大きさの結晶ならば、単一の結晶からの完全なデータ収集が十分に可能である。しかしながら、より小さな結晶、あるいは回折能の低い結晶では、放射線損傷の問題から、単一の結晶から完全なデータを得ることができないため、複数の結晶が必要になる。近年の解析ターゲットの高難度化により、多数の結晶を用いる戦略の重要性が増してきている。

多数の結晶を用いた構造解析をより簡単に行えるようにするため、私たちはデータ収集およびデータ処理の自動化をすすめている。BL32XU で開発された自動データ収集システム Zoo は、ロボット(SPACE)による試料交換、クライオループのセンタリング(INOCC)に続き、低線量回折スキャンとその結果に基づく結晶位置の検出(SHIKA)、放射線損傷を考慮したストラテジ作成(KUMA)を自動化することで実現した。これにより、クライオループにマウントされた多数の微小結晶から各 5-10°程度の振動範囲のデータセットが順次収集される。

多数の結晶に由来する回折データを処理する際は、個々の回折写真の積分処理だけでなく、別々の結晶に由来するデータをマージする際の取捨選択も重要になる。この作業を効率化するため、自動データ処理システム KAMO を開発した。自動収集された多数のデータセットは、それぞれ XDS (Kabsch, 2010)を用いて指数付けおよび積分処理が行われる。その後、同型なデータに分類するため、階層的クラスタリングが行われる。BLEND (Foadi, 2013)による格子定数に基づいた分類のほか、データ間の強度相関係数によるクラスタリングも可能である。階層的クラスタリングの結果から、完全度・多重度の大きな集団を見出し、それぞれマージを行う。マージには XSCALE を用いるが、この際に相関係数や結晶間の系統誤差を反映する統計値などを参照して回折写真ごとまたは結晶ごとに異常値となるものを除外する。発表では Zoo システムの詳細を実際の適用例をまじえて紹介したい。