

SACLA XFEL を用いた高効率タンパク質結晶構造解析を目指したデータ処理手法の開発

Development of the Data Processing Method for Protein Crystallography at SACLA

平田 邦生
Kunio Hirata

(国研) 理化学研究所
RIKEN

アブストラクト:

我々はタンパク質の結晶構造解析においてエネルギーバンド幅の異なる X 線を利用した回折実験を SACLA で行った。その結果、得られた回折データをうまく処理するためにはエネルギーバンド幅が回折データに与える影響を調査する必要があると考え、BL40XU において凍結タンパク質結晶のデータ収集とその評価を行った。

キーワード：タンパク質結晶構造解析、ピンクビーム、シリアル結晶構造解析

背景と研究目的:

タンパク質の結晶構造解析では、電子密度図の三次元再構成を実現するために結晶を回転させながら X 線を照射して回折像を得るいわゆる振動写真法を用いたデータ収集を行うことが一般的である。近年、タンパク質結晶構造解析に利用できるようになった X 線自由電子レーザーはフェムト秒オーダーの極短パルス光であるため、回折像を取得する際には結晶は静止しており、さらに一回の照射によりその体積は再利用不可能な状態に損壊もしくは放射線損傷を受ける。また利用する X 線のエネルギーバンド幅も SPRING-8 のシリコン二結晶分光器に比較して 50~100 倍も大きいため、得られる回折像、特に検出器上の反射点の形状を半径方向へ広げる効果を持つ。関連して、結晶に照射する X 線のエネルギーバンド幅が広いことはエワルド球の厚みが増すことに相当するため、結晶を回転させずとも 1 枚の回折像から得られる各反射強度の部分度は高くなる傾向があり、データ収集の効率はエネルギーバンド幅の狭い光に比較して高いとも言える。まとめると、バンド幅の大きな光を利用したタンパク結晶からのデータ収集では、振動をさせずにより高速なデータ収集が実現できるというメリットがある反面で、結晶格子が大きい試料、モザイク角が大きい試料を用いた場合に、1 枚の回折像上で各ブラッグ反射の伸びた反射パターンどうしが重なってしまうなどのデメリットもあるため従来利用してきたデータ処理プログラムだけでは適切な対応が難しくなりつつある。

申請者らは世界にさきがけて XFEL 光を用いた膜タンパク質の“無損傷”かつ“高分解能”結晶構造解析に成功した[2]。SACLA の XFEL をタンパク質結晶に照射して得られる静止回折像のデータ処理手法開発を目指した基礎データ収集を行った。しかしながら、現状 SACLA で利用可能なフェムト秒パルス X 線は各パルスに含まれる光子数、中心エネルギー、エネルギースペクトルにばらつきがあり、シンクロトロンとのデータと比較して得られる回折イメージ上の反射強度の積分が非常に困難であった。

このため、有効に積分できる回折データを選定する手順が必要となり、高精度構造解析に多大な時間を要した。申請者らはこの問題を解決するためのソフトウェア開発を行っており、既存データ処理プログラムで採用されている反射強度の回折幅のモデルのうち、特にエネルギー分解能の項がうまくXFELで得られた回折データを説明できないという問題に直面した。さらに図1に示されるように動径方向への回折プロファイルのノビをフォローして積分するタンパク結晶回折用ソフトウェアも当時はなかった。このため、SACLAを用いた無損傷・高分解能構造解析を高効率化することを最終目的に見据え、SACLA-XFEL光のエネルギー分解能($\Delta E/E \sim$ 数%)と同等で、強度、エネルギースペクトルが安定しているBL40XUのピンクビームを用い、データ積分(以下データ処理)手法開発に必要な基礎データを取得することを目的として実験を行った。

実験

試料名:チトクロムc酸化酵素結晶(空間群 $P2_12_12_1$, 格子定数 $a,b,c = 178.5, 182.6, 204.8$)

方法:振動写真法および静止写真法

使用装置:BL40XU高精度回折計および低温吹き付け装置、持ち込み浜松ホトニクスCMOS検出器

測定条件:X線エネルギー 10 keV, 16 keV ビームサイズ 100 μm 角、フラックス $10^{10} \sim 10^{11}$ photons/sec

結果および考察

既存の回折計および持ち込んだ3 Hz読みだし可能なフラットパネルCMOS検出器を用いて種々の条件で100K凍結状態のチトクロム酸化酵素の回折データ測定を行った。比較のために、エネルギー 16 keV(波長 0.78223 Å)の単色ビームとピンクビーム(バンド幅 $\sim 3\%$)で、それぞれ 1 つ の結晶を用い、振動幅 0.02° ステップで 5000 枚ずつの静止写真を撮影した(10 keVのデータについては示していない)。フレームあたりの入射光子数は同一になるようにした。データ処理プログラムXDS を用いて連続振動写真として処理を行い、その際に間引くフレーム数がデータの質に与える影響を調べた(図1)。単色ビーム(エネルギー分解能は 0.01% 程度)を利用した回折実験データでは、データを間引いていくことによりデータ「内」での整合性(等価な反射強度どうしの相関($CC(1/2)$))が崩れる度合いが、ピンクビームに比較して顕著であるという実験前に想定した結果が示された。また同様にフレームを徐々に間引いた強度と、通常データ測定により得たチトクロム酸化酵素のデータに対する相関(データ間)を調査した結果(図1:右下)についても同様にピンクビームによるデータの劣化度合いのほうが高いことがわかった。

しかしながら、図2に見られるように「単色光」を前提とした通常データ処理パッケージ(ここではXDS)では不適切な積分マスクが適用されたり、入射光スペクトルによるスポットの重なりに対処できない場合が数多く見られた。このため、図1右図のグラフにあるようにピンクビームデータでは正解の構造因子(同じ回折計で単色X線で得たデータを正解とした)との相関が低く、分解能が単色光に比較して非常に低くなってしまいうため構造解析に利用できるデータとして利用することが困難であることがわかった。

BL40XUで利用可能なバンド幅～3%のX線では、チトクロム酸化酵素のように格子定数が比較的長く(最長軸 ～205Å)、結晶のモザイク角が0.1～0.3°ほどの場合に、スポットの重なりが大きくなり従来のデータ処理プログラムでは指数付けや積分計算そのものが困難化することがわかった。これが原因となり、正しい構造因子が取得困難であること、分解能が単色ビームと比較して悪化することもわかった。0.1～0.3°のモザイク角のタンパク結晶から高効率にデータ収集を行うためのX線バンド幅は3%より小さいことが肝要であることはわかった。

本実験結果から事前にビームラインで利用できると想定していたバンド幅(1～2%)よりも実測したバンド幅～3%が大きかったこと、エネルギースペクトルが低エネルギー側に非常に長い裾を引く形状であったことも回折スポットの周囲のバックグラウンドの不均一化、スポットの形状変化などを著しくし、データ処理を困難化している原因であることが明らかとなった。回折像シミュレータSIM_MXを利用した検討の結果(図3)、BL40XUのバンド幅(～3%)で実験を行う場合には、結晶格子の短い試料(長軸<150 Å)を選定するか、モザイク角の小さな(< 0.2°)試料を選定するなど、試料で制限がかかる可能性も示唆された。

今後の課題

データ処理の結果から判断して回折計算モデルの検討については、もっと結晶格子の短い標準試料タンパク質が適切であったと考えられる。エネルギーバンド幅に関するシミュレーションを入念に行うとともに、エネルギーバンド幅が幅広く利用できるビームラインを選定するか、異なるエネルギー分解能が利用可能なビームラインを利用して比較検討を行うことが必須であろう。

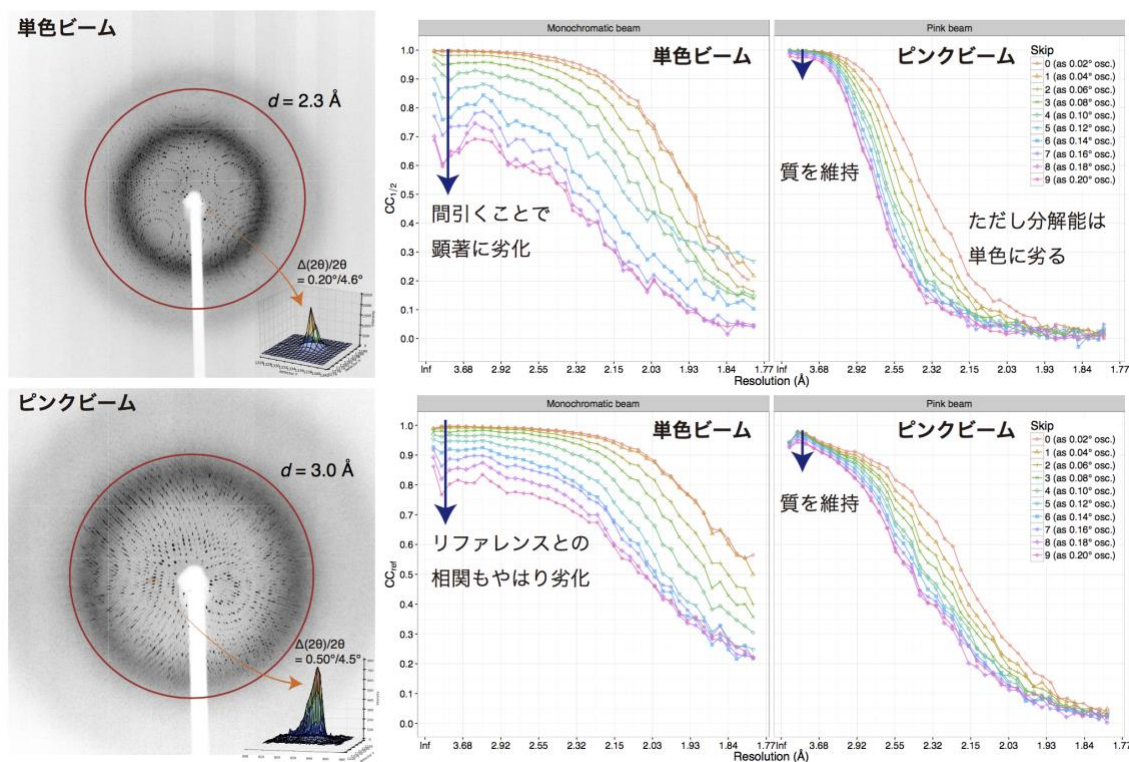


図1: 単色ビームとピンクビームを利用したデータの比較

左上図:単色ビームで測定したチトクロム酸化酵素の回折像(振動写真法)、左下図:ピンクビーム(バンド幅~3%)で測定した同酵素の回折像、ピンクビームでは動径方向に回折像がのびていることがわかる。右図上段:測定データを間引いた場合の分解能(d)に対する CC(1/2)の劣化速度のプロット(左:単色ビーム、右:ピンクビーム)、右図下段:チトクロム酸化酵素のリファレンスデータに対する相関係数を(左)単色ビーム(右)ピンクビームで比較。

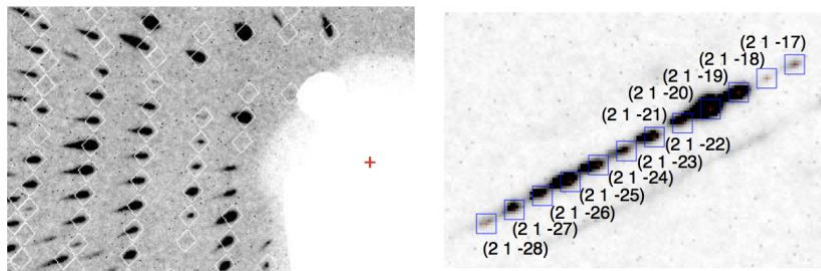


図2:ピンクビーム実験で得られたチトクロム酸化酵素の回折パターンの例

(左)ピンクビームの低分解能領域の回折像、(右)入射光スペクトルが広いことにより回折斑点同士が分離不可能な例

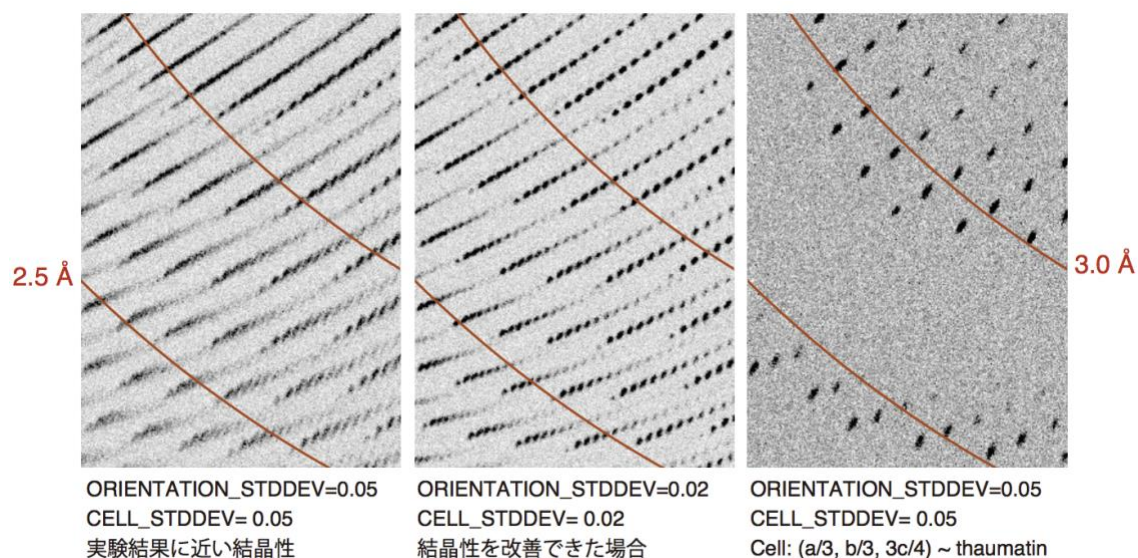


図3: 回折写真シミュレーションの結果(SIM_MX(Diederichs, 2009))

(左)今回用いたチトクロム酸化酵素の結晶に近い結晶性(モザイク角や格子定数の偏差)を反映したシミュレーション回折像、(中)さらに実際の結晶性を改善した場合のシミュレーション回折像、(右)格子定数の小さな結晶かつ結晶性が同程度の試料を利用した場合のシミュレーション回折像
図中のパラメタの定義は参考文献参照のこと[1]。

参考文献:

- [1]: K. Diederichs, *Acta Cryst. D*, **65**, 535-542 (2009).
[2]: K. Hirata et al., *Nat. Meth.*, **11**, 734-736 (2014).

(Received: March 30, 2020; Accepted: July 6, 2020; Published: August 21)