

タンパク質-化合物複合体の構造解析を加速させる  
スクリーニング系の構築

生命科学研究科 生体高分子動的構造解析学分野

伊藤 翔

指導教員：山本 雅貴 教授

2020年9月

兵庫県立大学

## 略号一覧

CNN	Convolutional neural network 畳み込みニューラルネットワーク
CryoEM	Cryo-electron microscopy クライオ電子顕微鏡法
FBDD	fragment-based drug discovery フラグメント化合物を利用した創薬研究
HSA	human serum albumin ヒト血清アルブミン
Kd	dissociation constant 解離定数
MAD	multi-wavelength anomalous dispersion method 多波長異常分散法
NMR	nuclear magnetic resonance 核磁気共鳴
PDB	Protein Data Bank プロテインデータバンク
PIL	Python Image Library パイソンイメージライブラリー
SAD	single-wavelength anomalous dispersion method 単波長異常分散法
SBDD	structure-based drug design タンパク質立体構造に基づく薬剤設計
SPR	surface plasmon resonance 表面プラズモン共鳴
SSD	single shot detector

シングルショットディテクター

TSA

thermal shift assay

熱シフトアッセイ

# 目次

## 第1章 序論

- 1-1 タンパク質 X線結晶構造解析
- 1-2 X線結晶構造解析における測定のハイスループット化
- 1-3 構造情報に基づいた創薬研究
- 1-4 本研究の目的と意義
- 1-5 参考文献

## 第2章 深層学習を利用した自動結晶センタリングシステム DeepCentering および 全自動構造解析パイプライン imaproc の開発

- 2-1 深層学習と X線回折実験への応用
- 2-2 材料と方法
  - 2-2-1 画像処理を用いた自動センタリングの流れ
  - 2-2-2 LoopDetecotr 用の教師データの収集
  - 2-2-3 CrystalDetector 用の教師データの作成
  - 2-2-4 独自のニューラルネットワークの作成と検出テスト
  - 2-2-5 自動構造解析パイプライン imaproc の開発
  - 2-2-6 DeepCentering と imaproc を用いたリゾチームの自動構造解析
- 2-3 結果と考察
  - 2-3-1 LoopDetector および CrystalDetector の検出テスト
  - 2-3-2 自動構造解析パイプラインによるリゾチームの構造解析
- 2-4 結論
- 2-5 参考文献

## 第3章 DeepCentering を用いたヒト血清アルブミンと環状ペプチドの複合体構造

## スクリーニング

### 3-1 ヒト血清アルブミンと環状ペプチドの創薬への応用

### 3-2 材料と方法

#### 3-2-1 ヒト血清アルブミンのアポ結晶の調製

#### 3-2-2 ヒト血清アルブミン-環状ペプチド複合体結晶の調製

##### 3-2-3-1 DeepCentering を用いた HSA アポ結晶のスクリーニングおよび構造解析

##### 3-2-3-2 DeepCentering を用いた HSA-環状ペプチドソーキング結晶のスクリーニング

#### 3-2-4 HSA-Dalbavancin 複合体結晶の構造解析

### 3-3 結果と考察

### 3-4 結論

### 3-5 参考文献

## 第4章 マイクロ流路を用いた化合物ソーキングシステムの開発

### 4-1 マイクロ流路を利用したX線結晶構造解析への応用

### 4-2 材料と方法

#### 4-2-1 結晶捕捉のための流路の設計および固定治具の作製

#### 4-2-2 リゾチームとソーマチンを用いた検証実験

#### 4-2-3 トリプシンと種々の化合物を用いた複合体スクリーニング

### 4-3 結果と考察

### 4-4 結論

### 4-5 参考文献

## 第5章 結論

### 5-1 本研究の考察

### 5-2 総括

### 5-3 今後の展望

謝辭

## 第 1 章

### 序論

## 1-1 タンパク質 X線結晶構造解析

タンパク質は、20 種類の L-アミノ酸がペプチド結合で重合し、アミノ酸配列に依存した固有の 3 次元構造を持つことで、生体内において様々な機能を担う生体高分子である。タンパク質の機能は多種多様であり、多くの場合、タンパク質は基質もしくは他のタンパク質と相互作用することによってその機能を発揮する。従って、タンパク質の機能を理解する為に、構造情報を基盤として、相手となる基質との相互作用解析を行うことは非常に重要である。現在、原子分解能でのタンパク質の構造決定方法は、大別して、① X線結晶構造解析、②CryoEM、③NMR、④中性子結晶構造解析、⑤MicroED、の 5 種類があり、それぞれの手法で長所と短所が存在する。X線結晶構造解析は、理論上解析可能な対象の分子量に制限がなく、十分な質の結晶さえ得ることができれば構造解析ができる。さらに、最も汎用的に用いられて来たため高度化も進んでおり、現在では解析に必要な結晶の体積も  $1 \mu\text{m}^3$  まで小さくなり結晶化の負担も軽減されている。測定の高速度化も図られ、短い場合では 1 分以内に必要なデータ測定を完了できる。一方、構造解析における結晶化の負担は軽減されているとは言え、必ずしも結晶が出来るわけではなく、また、結晶の質が構造解析の質に直結することから、結晶化が必要な点が欠点と言える。CryoEM では、結晶化が不要であり、分子量の大きいタンパク質や超複合体等の構造解析において非常に有効な手段である一方、一般に分解能が X線結晶構造解析よりも低いこと、実験のスループットが低いこと等が欠点として挙げられる。NMR は、溶液中におけるタンパク質の構造解析が可能である点、緩和時間解析によるタンパク質-基質相互作用等の動態解析が可能である点が長所としてあげられる。しかし、分子量が大きくなると、シグナルの数と線幅の増加により、解析が困難になる点が問題となる。中性子結晶構造解析は、中性子線と原子核が相互作用することで散乱が生じるため、散



乱能は電子数に依存しない。このため、水素をはじめとする軽元素の観測が容易である点、また、結晶への損傷が非常に小さい点が長所である。しかし、現在利用可能な中性子線の強度が弱いため、少なくとも 1 mm 以上、時には 10 mm 程度の巨大単結晶が測定に必要となること、測定時間が非常に長い（約 1-2 週間）ことが短所である。MicroED は、nm オーダーの結晶でも構造解析が可能であり、微小結晶の構造解析に有効であることが示されているが、手法として未だ確立されているとは言い難く、ルーチン的に実験を行うことが難しい。また、電子線回折で問題となる、多重散乱の影響を考慮した精密化プログラムが現状で存在しないことも問題として挙げられる。

タンパク質 X 線結晶構造解析は、原子分解能でタンパク質構造を簡便かつ汎用的に解析し、視覚化出来る最も有力な手法である。X 線結晶構造解析を用いた研究は、基礎研究のみならず、創薬研究に代表される産業利用においても広く利用されている。

タンパク質 X 線結晶構造解析の歴史は古く、1926 年にタンパク質では初めて、加水分解酵素であるウレアーゼが結晶になることを Summer が明らかにした[1]。また、1930 年に、Northrop によりアスパラギン酸プロテアーゼであるペプシンの結晶化がなされ[2]、その 4 年後の 1934 年に Bernal と Hodgkin によってペプシン結晶から初めてタンパク質の X 線回折が観測された[3]。更にその約 20 年後となる 1958 年に Kendrew らによって、世界初のタンパク質の結晶構造解析となる、ミオグロビンの結晶構造が報告された[4]。タンパク質 X 線結晶構造解析における放射光 X 線の利用も、数多くの結晶構造解析に貢献している。SPring-8[5]や Photon Factory[6]に代表される放射光施設は、X 線結晶構造解析のパラダイムシフトをもたらした。実験室系で利用していた X 線と比較して、遥かに輝度の高い X 線が利用可能になったことで、それまで不可能であった微小結晶や、結晶性の悪い試料の構造解析が可能となった。実験に使用する X 線波長についても、一般的なタンパク質 X 線結晶構造解析で使用する範囲において、任意の波長が利用可能となった。また、X 線集光技

術、種々の検出器[7]の開発もタンパク質結晶構造解析の発展に大きく寄与している。これらの技術革新によって、膜タンパク質に代表されるような、高難度のサンプル由来のタンパク質X線結晶構造解析が可能となった。さらに、前述の通り任意のX線波長が使用可能になったことにより、X線回折強度の波長依存性を利用したSAD法[8]やMAD[9]法などに代表される異常分散法が成熟し、構造解析の成功例が増加した。これまでX線結晶構造解析がなされたタンパク質の例を挙げると、光合成細菌の反応中心複合体[10]、放射菌由来のカリウムチャンネルKcsA[11]、ATP合成酵素[12]、リボソーム[13]、 $\beta_2$ アドレナリン受容体[14]など枚挙に暇がない。2020年6月1日の時点で、タンパク質構造データベースである、PDB(Protein Data Bank)に登録されている構造の数は、164,391であり、その内の約8割がX線結晶構造解析によって決定された。

これらの構造情報を基盤として、タンパク質の機能解析、タンパク質の高機能化、計算科学的手法による動的構造のシミュレーション[15-18]、創薬研究におけるFBDD (Fragment-based Drug Design) [19]に代表されるリード化合物の探索と創製およびSBDD (Structure-based Drug Design) によるリード化合物の最適化など、多くの応用がなされている。このように、タンパク質X線結晶構造解析は現代の生命科学研究において、タンパク質の機能を理解するために必須の手法となっている。

## 1-2 X線結晶構造解析における測定の高スループット化

タンパク質研究において、X線結晶構造解析がタンパク質の機能理解のために大きく貢献していることは前述のとおりである。しかしながら、結晶構造が得られるまでにはいくつかのボトルネックが存在する。一般的なX線結晶構造解析の流れをFigure 1-1に示す。

タンパク質X線結晶構造解析では、タンパク質の大量発現系を構築し、精製、結晶化を行う。結晶が得られたら、X線を結晶に照射し、X線回折強度測定を行い、それに次いで種々の方法による位相決定、構造モデルの構築、構造精密化を行い構造解析が完了となる。結晶化までの工程では、現在でも多くの条件検討を要するが、近年では、自動結晶化装置[20]等の開発により、より少ないサンプル量かつ短時間で、広範囲のケミカルスペース[21]をカバーした結晶化条件の探索が可能となり、高品質の結晶を得られる可能性が高まっている。

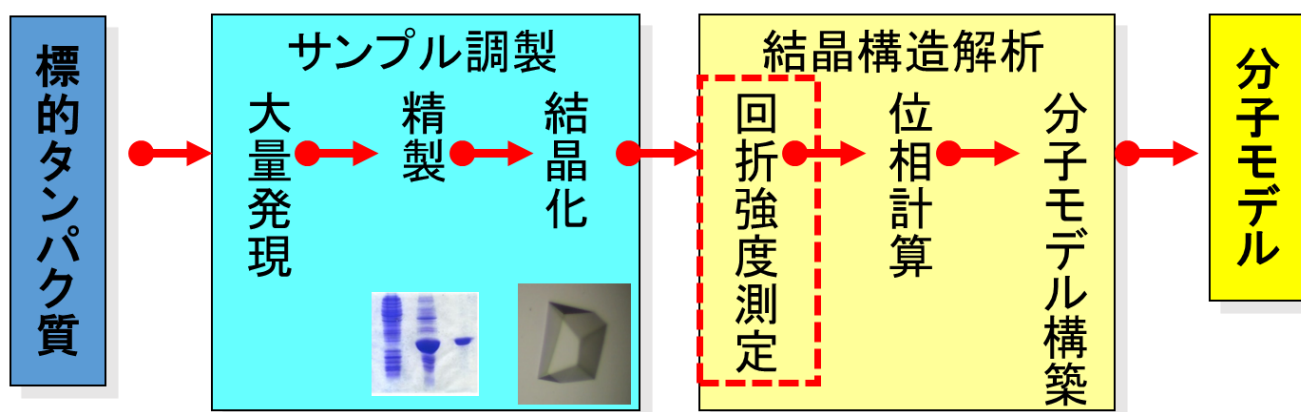


Figure 1-1 タンパク質結晶構造解析の流れ

サンプル調製の工程のみならず、X線回折強度測定以降の工程においても技術革新が進んでいる。その最たる例が放射光施設の登場である。

放射光施設では、輝度の高いX線を用いた回折強度測定を効率的に行えるよう、測定の自動化のための開発が行われてきた。前述の通り、X線結晶構造解析の利用拡大や試料の高難度化によって、数多くの結晶に対して、X線回折強度測定を効率よく行う必要性がますます高まっている。多くの結晶の測定に対応するため、自動サンプルチェンジャー[22-24]や自動データ収集、自動データ処理プログラム[25, 26]の開発など、世界中の放射光施設ではハイスループット化が進んでいる。

一方、次節で述べるように、FBDDでは、結晶サイズの比較的大きい、多量の複合体結晶の調製を試み、そのすべてについて、X線回折強度測定を行う必要がある。そのような多数の大型結晶を自動測定する場合には、センタリングと呼ばれる、測定の間、結晶がX線の光路上から外れないようにする位置決め工程において、長時間を要することがX線回折強度測定における高速化のボトルネックとなっており、実験の自動化に改善の余地が残されている。

### 1-3 構造情報に基づいた創薬研究

タンパク質の構造情報に基づく創薬研究 (Structure Based Drug Design; SBDD) [27-29]は1980年代から利用されてきた。本論文では、疾患関連タンパク質を標的タンパク質と定義する。SBDDでは、まず、標的タンパク質とその結合相手となりうる化合物の複合体結晶構造解析を行う。得られた複合体結晶構造を利用してタンパク質と化合物の相互作用解析を行うことにより、新薬の候補となる化合物の合理的設計、合成を行う。次いで、新たに合成した化合物と標的タンパク質との複合体構造解析を行い、得られた複合体構造から化合物の更なる改良を行う。この工程を繰り返すことで薬剤の最適化を行う。

SBDDに代わって、2000年代初頭から、フラグメント化合物を利用した創薬研究

(FBDD) が新規薬剤候補化合物を開発する効果的な手法として急速に発展してきた。FBDD の概要を Figure 1-2 に示す。FBDD の工程は以下の通りである。① フラグメントライブラリーと呼ばれる、凡そ分子量が 500 Da 以下かつ化合物の化学的特徴の多様性が広い化合物候補群を用いて、表面プラズモン共鳴 (SPR) [30] や熱シフトアッセイ (TSA) [31] 等の物理化学的手法によって、標的タンパク質に結合しうるリード化合物を選出する。② ① で選出されたリード化合物と標的タンパク質の複合体構造解析を行い、リード化合物と標的タンパク質の相互作用部位の結合様式を詳細に解析する。③ リード化合物が標的タンパク質により強く結合するように、リード化合物に置換基を導入するなどして、構造情報に基づいた薬剤設計を行う。ここで重要となるのは、FBDD では、スクリーニングに比較的分子量の小さい化合物を使用していることである。低分子化合物を使用することで、一般的に FBDD で用いられるライブラリーの化合物数は 1000~10000 化合物程度と、従来のハイスループットスクリーニング (HTS) [32] に代表されるような手法と比較して、化合物数が少ないにも関わらず、化合物の多様性 (ケミカルスペース) を広くカバーすることが可能となる。更に、低分子化合物は一般的に水溶性が高いため、フラグメントライブラリーを利用することで、必然的に水溶性の高い化合物をスクリーニングに用いることが可能となる点も重要である。化合物スクリーニングの初期段階では、化合物のタンパク質に対する結合は弱く、 $K_d$  にして mM から  $\mu\text{M}$  程度である。そのような結合の弱い場合に複合体結晶構造を得るためには、高濃度の化合物溶液を標的タンパク質に添加させる必要がある。溶解度の高いフラグメントライブラリーを利用することで、比較的高濃度の化合物溶液が複合体構造解析に利用可能となり、ヒット化合物の検出が可能になる。これらの長所をもつ FBDD によって、近年、多くの薬剤候補化合物が開発されており、創薬研究において、リード化合物の最適化に非常に有用な手法として認識されている [33, 34]。

FBDD は創薬研究のための強力な手法であるが、複合体結晶の調製がボトルネックとなっている。ソーキング法の概要を Figure 1-3 に示す。一般的なソーキング法の

手順は、①結晶化プレートからアポ結晶を取り出し浸漬を行うための他の容器へ移す。②化合物をアポ結晶に浸漬する。③抗凍結剤を添加する。④液体窒素等で結晶を凍結する。これらの作業をすべて実験者が手作業で行う必要がある。さらに、適切な化合物濃度、浸漬時間、抗凍結剤濃度など、数多くの最適化すべきパラメーターが存在する。FBDDでは、アポ結晶に対して、TSAやSPRなどの物理化学的手法により、タンパク質との結合が確認された様々なヒット化合物とタンパク質のソーキング実験を行う。一般的に物理化学的手法を用いて選出されるヒット化合物は複数個であることが多いため、ソーキング法によって多くの複合体結晶調製を要求される場合が多い。このように、自動化が進んでいるX線結晶構造解析において、ソーキングによるタンパク質-化合物複合体結晶調製は未だに時間と労力を要する工程であり、タンパク質-化合物複合体構造決定のボトルネックとなっている。

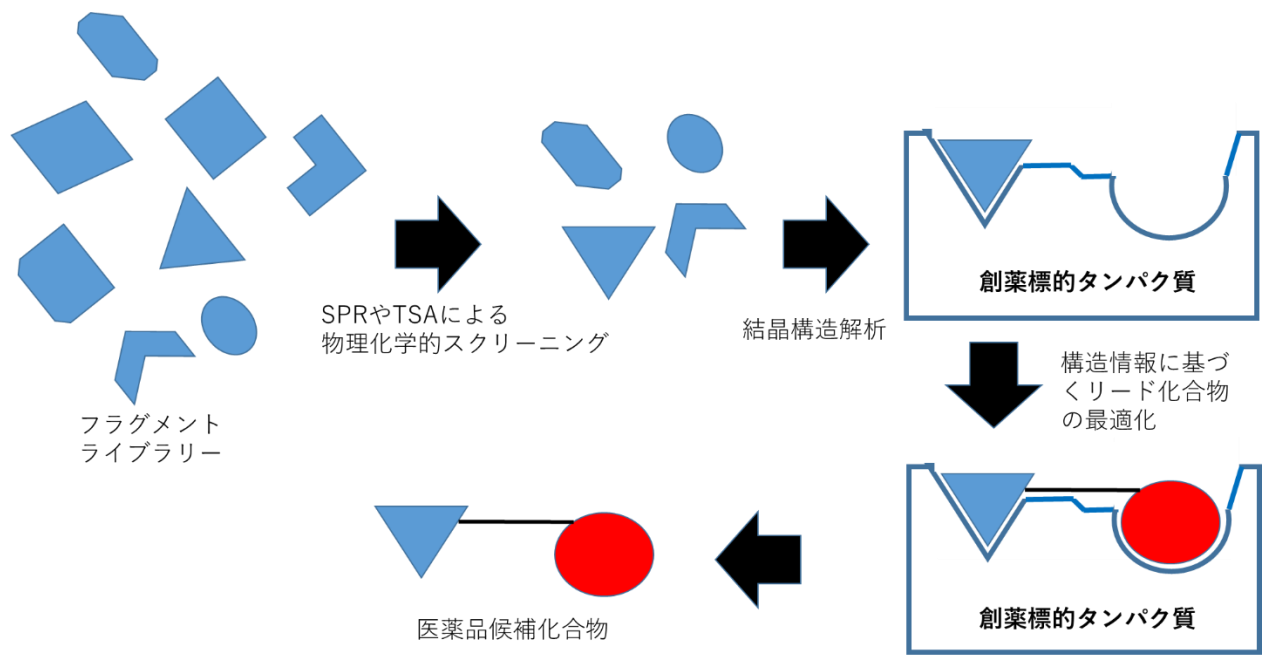


Figure 1-2 FBDD の概要

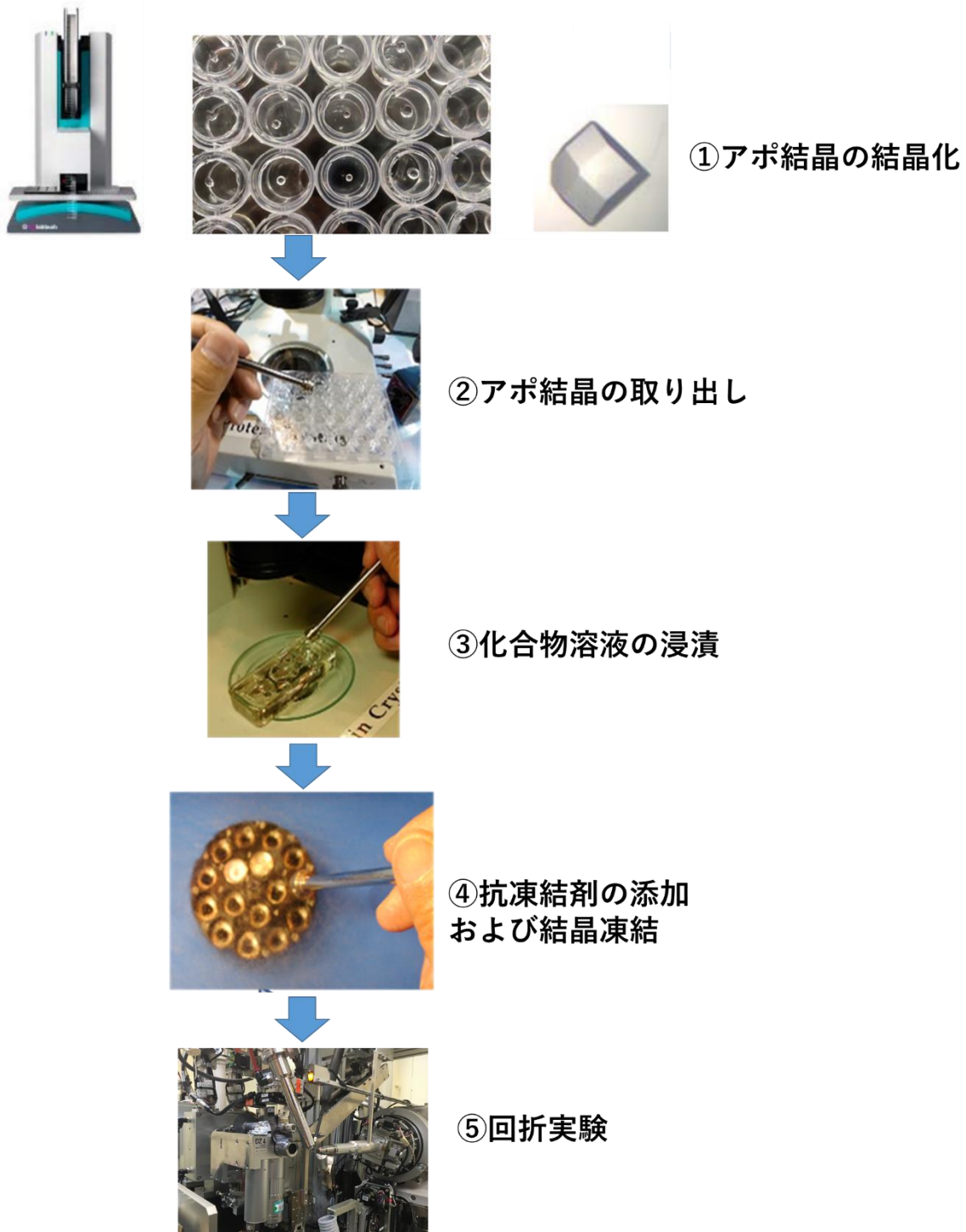


Figure 1-3 ソーキング法の実験手順



#### 1-4 本研究の目的と意義

これまで述べてきたように、X線結晶構造解析による、タンパク質-化合物複合体構造解析は基礎研究のみならず、創薬研究をはじめとする多くの産業利用もなされており、今後も構造未知のタンパク質や、創薬研究において重要なタンパク質-化合物複合体の構造決定のために、多くのX線回折強度測定が必要不可欠であることは疑いの余地がない。しかし、前述の通り、複合体結晶のX線結晶構造解析のハイスループット化においては、タンパク質-化合物複合体調製の重要な工程であるソーキング法の省力化および放射光施設における大型結晶の自動測定の迅速化が求められている。そこで、本研究では、①大型結晶を対象とする深層学習を用いた自動回折強度測定システム、②マイクロ流路を用いた化合物ソーキングシステムを開発し、上記の問題解決に取り組んだ。これら実験系の構築により、測定の迅速化、省力化が見込まれ、これから構造解析が行われるであろう、多くの標的タンパク質の構造解析に貢献することが期待される。

本論文の構成は下記の通りである。

第2章では、深層学習を用いた画像処理技術を大型結晶の結晶センタリングに応用したプログラム、DeepCenteringの開発とその結晶検出精度テスト結果について述べる。また、放射光施設の全自動構造解析パイプラインとして開発したプログラムである imaproc と、DeepCentering を組み合わせることで、完全自動構造解析システムを構築した。このシステムのテストのためにリゾチームを用いた自動測定、自動構造解析についても述べる。

第3章では、DeepCenteringの応用例として、ヒト血清アルブミンのアポ構造、およびヒト血清アルブミン-環状ペプチド複合体構造のスクリーニング実験とその構造解析結果について述べる。

第4章では、タンパク質-化合物複合体結晶の調製における、ソーキング実験の簡

便化、省力化のために開発を行った、常温測定でのマイクロ流路ソーキングシステムおよびその応用例として行ったトリプシンの化合物スクリーニングについて述べる。

第5章では、本研究を総括し、今後の展望について述べる。

1-5 参考文献

- [1] Summer, J. B. (1926). *J. Biol. Chem* **69**, 435-441.
- [2] Northrop, J.H. (1930). *J. Gen. Physiol* **13**, 739-766.
- [3] Bernal, J. D., Crowfoot, D. (1934). *Nature* **133**, 794-795.
- [4] Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H. & Phillips, D. C. (1958). *Nature* **181**, 662.
- [5] <http://www.spring8.or.jp/ja/>
- [6] <https://www2.kek.jp/imss/pf/>
- [7] D. Renker, E. Lorenz. (2009). *Journal of Instrumentation* **4**, 4004.
- [8] Hendrickson, W. A. and Teeter, M. M. (1981). *Nature* **290**, 107-113.
- [9] Hendrickson, W. A. (2014). *Q Rev Biophys* **47**, 49-93.
- [10] Deisenhofer, J., Epp, O., Miki, P., Huber, R. & Michel, H. (1985). *Nature* **318**, 618-624.
- [11] Declan, A. Doyle., Jao Morais Carbral., Richard, A. Pfuetzner., Anling, Kuo., Jacqueline, M. Gulbis., Steven, L. Cohen., Brian, T. Chait., Roderick, MacKinnon. (1998). *Science* **280**, 69-77.
- [12] Abrahams, J. P., Leslie, A. G. W., Lutter, R., Walker, J. E. (1994). *Nature* **370**, 621-628.
- [13] D.E. Brodersen, W.M. Clemons, A.P. Carter, R.J. Morgan-Warren, B.T. Wimberly, V. Ramakrishnan. (2000). *Cell* **103**, 1143-1154.
- [14] J. Harms, F. Schluenzen, R. Zarivach, A. Bashan, S. Gat, I. Agmon, H. Bartels, F. Franceschi, A. Yonath. (2001). *Cell* **107**, 679-688.
- [15] N. Ban, P. Nissen, J. Hansen, P.B. Moore, T.A. Steitz. (2000). *Science*, **289**, 905-920

- [16] Soren, G. F. Rasmussen., Hee-Junc Choi, Daniel, M. Rosenbaum., Tong Sun Kobilka, Foon Sun Thian, Paricia, C. Edwards., Manfred Burghammer, Venkata, R. P. Ratnala., Ruslan Sanishvili, Robert, F. Fischetti., Gebhard, F. X. Shertler., William, I. Weis., & Brian, K. Kobilka. (2007). *Nature* **450**, 383–387.
- [17] Egan, W. J., Merz, K. M., & Baldwin, J. J. (2000). *Journal of medicinal chemistry* **43**, 3867–3877.
- [18] Jorgensen, W. L., & Duffy, E. M. (2002). *Advanced drug delivery reviews* **54**, 355–366.
- [19] Sliwoski, G., Kothiwale, S., Meiler, J., & Lowe, E. W. (2014). *Pharmacological reviews* **66**, 334–395.
- [20] Korczynska, J., Hu, T.C., Smith, D.K., Jenkins, J., Lewis, R., Edwards, T., Brzozowski, A.M. (2007). *Acta Crystallogr. D* **63**, 1009–1015.
- [21] Kirkpatrick, P., Ellis, C. (2004). *Nature* **432**, 823.
- [22] Ueno, G., Hirose, R., Ida, K., Kumasaka, T. & Yamamoto, M. (2004). *J. Appl. Cryst.* **37**, 867–873.
- [23] H. Murakami, G. Ueno, N. Shimizu, T. Kumasaka, M. Yamamoto (2012). *J. Appl. Cryst.* **45** 234–238.
- [24] Murakami, H., Hasegawa, K., Ueno, G., Yagi, N., Yamamoto, M. & Kumasaka, T. (2020). *Acta Cryst.* **D76**, 155–165.
- [25] K. Hirata, K. Yamashita, G. Ueno, Y. Kawano, K. Hasegawa, T. Kumasaka, and M. Yamamoto. *Acta Cryst* (2019). **D75**, 1–13.
- [26] K. Yamashita, K. Hirata, and M. Yamamoto. *Acta Cryst* (2018). **D74**, 441–449.

- [27] Baker, M. (2013). *Nat Rev Drug Discov* **12**, 5–7.
- [28] Hubbard, R. E. & Murray, J. B. (2011). *Methods Enzymol* **493**, 509–531.
- [29] Blundell, T. L., Jhoti, H. & Abell, C. (2002). *Nat Rev Drug Discov* **1**, 45–54.
- [30] Klebe, G. (2000). *J Mol Med (Berl)* **78**, 269–281.
- [31] Gane, P. J. & Dean, P. M. (2000). *Curr Opin Struct Biol* **10**, 401–404.
- [32] Shepherd, C. A., Hopkins, A. L. & Navratilova, I. (2014). *Prog Biophys Mol Biol* **116**, 113–123.
- [33] Jhoti, H., Cleasby, A., Verdonk, M. & Williams, G. (2007). *Curr Opin Chem Biol* **11**, 485–493.
- [34] Hajduk, P. J., Greer, J. (2007). *Rev. Drug Discovery*, **6**, 211–219.

## 第2章

深層学習を利用した自動結晶センタリングシステム

DeepCentering および全自動構造解析パイプライン

imaproc の開発

## 2-1 深層学習と X線回折実験への応用

近年、音声認識、言語処理、画像処理などの多くの分野において、深層学習を用いた技術開発が進んでいる[1-12]。深層学習[13, 14]は、機械学習[15]の一種であり、ニューラルネットワークと呼ばれる脳の神経回路の一部を模した数理モデルを用いて、課題解決を行う手法である。とりわけ、画像処理の分野においては、自動運転技術への応用など、活発に利用されている。画像処理分野の深層学習では、畳み込みニューラルネットワーク (Convolutional neural network、CNN) と呼ばれる方法が広く用いられており、物体検出の正解率を競う画像処理コンペティションでは、深層学習による手法のほとんど全てが CNN をベースとしている。タンパク質 X 線結晶構造解析の分野においても、シリアル結晶学における回折イメージの質の判定や、タンパク質結晶化後の結晶化プレート観察において、結晶が析出したか否かの判定に深層学習が応用されている[16, 17]。

深層学習を用いた画像検出は 3 つのステップに大別される。① ニューラルネットワークの作成のための教師データ収集、② 学習による最適なニューラルネットワークの重みの調整、③ 作成したニューラルネットワークを用いた物体検出といった手順で行われる。ここで重要となるのは、目的となる物体検出のために、適切な重みづけがされたニューラルネットワークを作成することである。従来の機械学習（非深層学習）を用いた学習では、「人」の考えた画像中のエッジ等の特徴量を抽出し、機械学習によってその規則性を見つけ出す。規則性の探索も、そのアルゴリズム（識別器とよばれる）は「人」が決定する。一方深層学習では、機械学習で「人」が決定したこれらの工程をすべて自動で行うことが可能である (Figure 2-1)。これにより、入力データと適切に重みづけされたニューラルネットワークさえあれば、物体検出は一段階の処理で完結する。

前述の通り、深層学習において、ニューラルネットワークの重みを適切に調整することを「学習」とよぶ。Figure 2-2 に単純なニューラルネットワークの例を示

す。ニューラルネットワークは、入力層に処理の対象となるデータ（画像処理の場合は画像データ）を入力する。Figure 2-2 の  $X_1$  が画像中の1つのピクセルの値に対応する。従って、色の情報を持たない二値化された画像データでは、入力層のニューロン数は画像中の総ピクセル数に等しい。また、Figure 2-2 中の円で示したそれぞれをニューロンもしくはノードと呼ぶ。

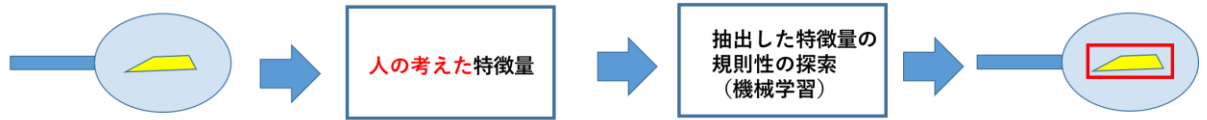
ニューラルネットワークでは、重みを係数として、入力層から中間層、中間層から出力層へと値を伝送する。例として、 $Y_1$  の計算方法を示すと、 $Y_1 = h(w_{11}X_1 + w_{21}X_2)$  となる。ここで、 $h()$  は活性化関数と呼ばれ、シグモイド関数などが用いられる。一般に、4層以上の層を持つニューラルネットワークを用いた系を深層学習と呼ぶ場合が多い。また、学習の工程では、損失関数と呼ばれる、ニューラルネットワークの性能の「悪さ」を示す指標（ニューラルネットワークが教師データに対してどれほど適合していないかを表す）を手がかりに、最適な重みの探索を行う。損失関数には任意の関数を用いることができるが、一般に  $1/2 \sum_k (y_k - t_k)^2$  で表される二乗和誤差や、 $-\sum_k t_k \log y_k$  で表される交差エントロピー誤差を用いる。ここで、 $y_k$  は出力層のニューロンの値、 $t_k$  は教師データの正解ラベルである。出力層の値と教師データの正解ラベルの差が小さいほど、損失関数の値は小さくなり、重みづけが適切にされていることを意味する。

ニューラルネットワークの学習において重要なことは、あらゆるデータに対応可能な、汎化性の高いニューラルネットワークを構築することである。これとは逆に、教師データだけに過度に対応した状態を過学習とよぶ。本研究を例にとると、結晶の保持に使用されるクライオループの検出において、複数の種類が存在するクライオループの内、特定のタイプのクライオループだけを教師データとして学習を行って作成されたニューラルネットワークが、そのタイプのクライオループでのみ検出精度が高くなる（他のタイプのクライオループは検出できない）ような場合が過学習に該当する。一般的に、ニューラルネットワークの学習では、過学習が起きないように、教師データを訓練セットとテストセットに分割し、それぞれに対して



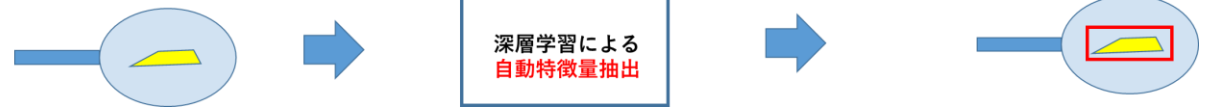
学習を行い、両方で損失関数が減少するかを確認しながら行う。更に、訓練セットとテストセットは任意の数に分割され、それぞれに対して学習を行う。このグループの数をバッチ数と呼ぶ。また、エポック数と呼ばれる値を次のように定義している。①、②を一回行うことをエポック数 = 1 とする；①教師データセットをバッチ数に従ってN個のサブセットに分ける。②各サブセットを学習に回す。つまり、N回学習を繰り返す。既に述べた、また、各層のニューロン数や学習の際のバッチサイズなどは、ハイパーパラメーターとよばれ、学習の際に開発者が適宜調整することもある。

機械学習（従来法）



入力画像

深層学習



検出結果

Figure 2-1 従来の機械学習と深層学習の比較

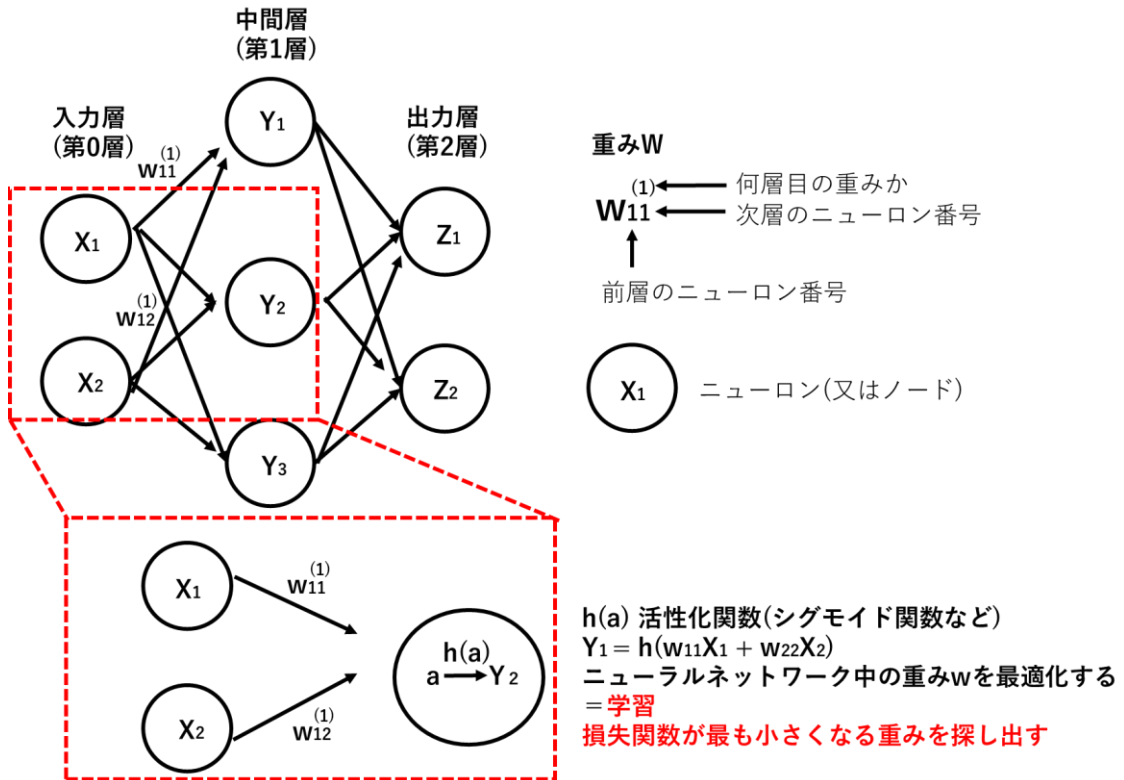


Figure 2-2 ニューラルネットワークの概要

本研究では、これまで述べたように、昨今技術進歩が目覚ましい深層学習による画像検出技術をX線回折強度測定における試料の物体検出に応用することを試みた。一般にX線回折強度測定では、センタリングと呼ばれる、結晶を回転させながらX線を照射している間、結晶位置がX線から外れないようにするための位置合わせが必要となる (Figure 2-3)。センタリングの具体的な手順は次の通りである。

- ① ある方位での結晶位置のX線照射位置へ並進の位置合わせ
- ② クライオロープ/結晶を 90° 回転
- ③ 回転後の位置で再度クライオロープ/結晶をX線照射位置へ並進

一般的に、この工程は、実験者が回折計に取り付けられた試料観察カメラで結晶を目視で確認しながら手作業で行う方法か、以下に述べるその他の自動結晶検出方法を用いて行う。

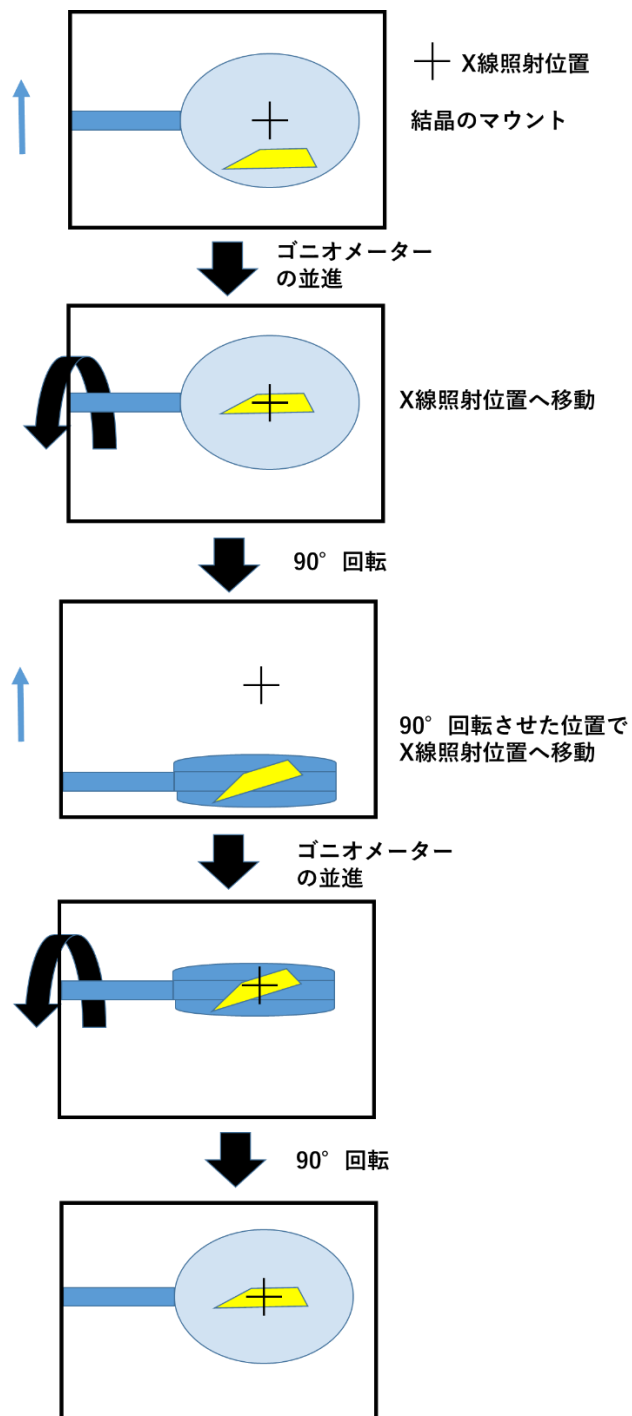


Figure 2-3 センタリングの手順

図中の黄色多角形が結晶、楕円形シアンでクライオループを示す。

Table 2-1 に既に関連されている自動センタリングの方法をまとめた。現在多くのタンパク質構造解析ビームラインにおいて、そのほとんどがラスタースキャン [18] と呼ばれる、強度を弱めた X 線を用いて、クライオループ上を走査しながら結晶位置を探索する方法が採用されている。その他、画像処理技術を用いた自動センタリングプログラムや、タンパク質結晶中のトリプトファンの蛍光を検出し結晶位置を可視化する装置に代表されるような追加装置を回折計に導入し結晶位置を検出する方法 [19-24] があるが、前者については、検出精度が低いこと、後者については、試料周りに追加装置を導入するための空間的余裕がないビームラインが多く、装置の導入が困難であることからあまり利用が進んでいない。

これら 2 つの自動センタリングシステムの利用が進まない一方で、画像処理を用いた自動センタリングは、X 線損傷を完全に回避できるという点でメリットがある。これは、金属タンパク質や常温測定などの X 線損傷に敏感な結晶に対して有効であると考えられる。更に、画像処理分野、とりわけ物体検出において従来法よりも、検出精度が非常に高いことで注目を集めている深層学習を結晶センタリングに利用することで、画像処理による結晶検出の精度の改善が期待できる。

また、ラスタースキャンで自動センタリングを行う場合において、とりわけ大型結晶のセンタリングには長時間を要することがあり、大型結晶の自動センタリングの高速化が望まれている。

そこで、本研究では、深層学習を用いた画像検出技術を、X 線回折強度測定時の結晶及びクライオループの検出に応用し、自動センタリングシステム、DeepCentering [25] を開発した。とりわけ、ラスタースキャンを用いるとセンタリングに時間を要する、50  $\mu\text{m}$  以上の比較的大きい結晶サイズを対象とした。

Table 2-1 種々の自動センタリング方法

それぞれのセンタリング方法とその長所、短所を示す。微小ビームが利用可能なアンジュレータービームラインでは、ラスタースキャンに代表されるX線回折を用いた方法が最もよく利用されている。

センタリングの種類	手法/プログラム名	長所	短所
回折の利用	ラスタースキャン	高精度	X線損傷、要長時間
画像処理	C3Dその他多数	無損傷	低精度、視認可能サイズのみ
追加装置の導入	蛍光やSHG (Second Harmonic Generation)の利用	高精度	要追加装置

## 2-2 材料と方法

### 2-2-1 画像処理を用いた自動センタリングの流れ

画像処理を用いた自動センタリングにおいては、クライオループの検出および結晶の検出のためのニューラルネットワークの2点が必要となる。本研究では、それぞれの検出のための検出プログラムを LoopDetector および CrystalDetector と名付け、検出に必要なクライオループと結晶それぞれのニューラルネットワークを作成した。自動センタリング全体の工程を Figure 2-4 に示す。DeepCentering では、サンプルチェンジャーによる、結晶をマウントしたクライオループのゴニオメーターへの取り付け後に、LoopDetector を用いたクライオループのセンタリングを行う。その後、回折計に取り付けられた結晶観察カメラからクライオループの面が最もよく見えるようにクライオループの面出しを行う。結晶検出精度を上げるため、自動液体窒素滴下装置を用いて、クライオループ上に付着している霜やごみを取り除き、最後に CrystalDetector による結晶検出を行いセンタリングが完了となる。



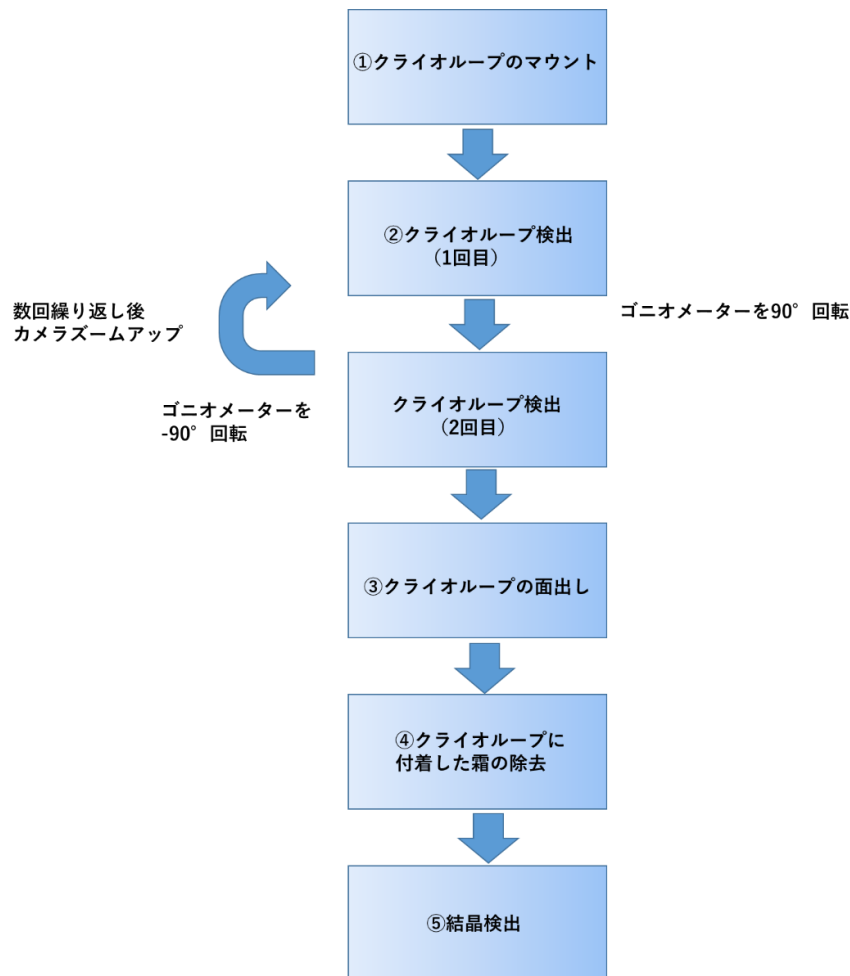


Figure 2-4 DeepCentering によるセンタリングの流れ

DeepCentering は以下の手順でセンタリングを実行する。

- ① 自動サンプルチェンジャーSPACE による回折計へのクライオループのマウント
- ② LoopDetector によるクライオクライオループ位置の検出および中心への移動
- ③ 結晶検出の前段階としてのクライオクライオループの面が結晶観察カメラの視野方向に対して垂直になるような位置決め
- ④ 自動液体窒素滴下装置によるクライオクライオループ状に付着した霜の除去
- ⑤ CrystalDetector による結晶検出

結晶観察カメラの視野から SPACE によるサンプルマウント後のクライオループが外れないように、LoopDetector によるクライオループのセンタリングは、第1段階では低倍率、低倍率でクライオループの位置が試料観察カメラの中央に位置した後は高倍率の2種類の倍率で行われる。

### 2-2-2 LoopDetector用の教師データの収集

LoopDetectorのための教師データは、SPring-8 BL26B2のメイン測定[26, 27]の際に、結晶観察カメラの倍率、輝度を様々な値に変更しながら6031枚の画像を取得し、LoopDetectorのための教師データとした。

### 2-2-3 CrystalDetector用の教師データの作成

クライオループ自体の物体検出と比較して、結晶の検出は、結晶化溶媒が周囲に存在するため、検出難易度が高くなることが予想された。予備実験として、約5000枚の結晶画像を用いて学習を行い、検出テストを行ったが、多くの場合において結晶の検出に失敗した。適切なニューラルネットワークの重みづけのためには更に多くの教師データを用意し、学習を行う必要がある。別のアプローチとして、画像として見た時の結晶は、多角形であるから、より多角形の境界が明瞭な画像を使えば、結晶としての特徴量を効率よく抽出可能なニューラルネットワークが作成できると考えた。そこで、効率的に学習を行うことを目的とし、CrystalDetectorのための教師データを、Pythonのライブラリーである、PIL (Python Image Library) [28]を用いて、ランダムに418枚の多角形画像を作成しこれをCrystalDetectorの教師データとした (Figure 2-5)。

### 2-2-4 独自のニューラルネットワークの作成と検出テスト

Tensorflow、Chainer、Caffe、Keras[29-32]など、深層学習のフレームワークは数多く存在する。本研究では、比較的簡単にニューラルネットワークの作成が可能かつ活発に開発が行われている、TensorFlow Object Detection API[33]を利用してLoopDetectorとCrystalDetectorのためのニューラルネットワークを作成した。いずれのニューラルネットワークについても、学習のベースモデルには、You Only Look Once (YOLO) [34]、Faster R-CNN[35]、Single Shot Detector[36]などがよく利用されるが、本研究では、検出速度が速いSingle Shot Detector (SSD)を用いた。

LoopDetector のニューラルネットワークについては、過学習を防ぐために、6031 枚の教師データ画像を 2 分割し、訓練セットとテストセットの損失関数をモニターすることで学習が適切に進んでいるかを確認した。LoopDetector の学習におけるバッチ数は 24 とした。学習が適切に進んでいるかの指標となる損失関数（値が小さいほど学習が適切に進んでいることを意味する）が十分に減少した時点で学習を終了させた。LoopDetector では、エポック数 50 で学習を終了し、これを LoopDetector のニューラルネットワークとした。その他のハイパーパラメーターはデフォルトの設定で学習を行った。CrystalDetector のニューラルネットワークについては、過学習に陥った場合は、検出精度が著しく低下したため、過学習か否かの判断が非常に容易であった。従って、教師データの分割を行わず、全画像データを学習に用いた。バッチ数は 24 とした。様々なエポック数でニューラルネットワークを構築し、実際に結晶検出を行い最も検出精度の高いものを CrystalDetector のニューラルネットワークとした。その時のエポック数は 8 であった。

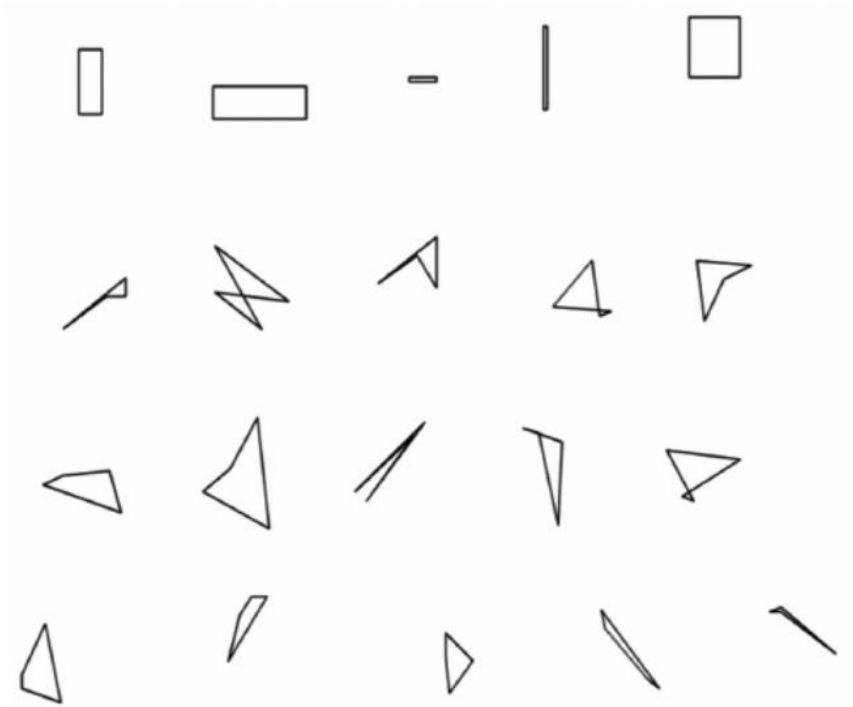


Figure 2-5 CrystalDetector の教師データとして用いた画像の例

Python Image Library (PIL) を用いて 418 枚の多角形画像を生成し、これを教師データとした。

作成した LoopDetector および CrystalDetector を用いて、検出テストを行った。テスト画像には、BL26B2 で撮影した 100 枚のクライオループ画像を、結晶のテスト画像は、BL26B2 でのメールイン測定の際に撮影した 960 枚の結晶画像を用いた。LoopDetector および CrystalDetector のいずれの検出テストについても、ビームセンターから 30  $\mu\text{m}$  以内に結晶が位置している場合を、自動センタリングの成功と定義した。

#### 2-2-5 自動構造解析パイプライン imaproc の開発

回折強度測定の自動化だけでなく、位相付けや構造精密化の自動化も各放射光施設において進められてきた。特に、イギリスの Diamond Light Source[37]を中心に開発された dimple[38]は、Phaser[39]による分子置換法を用いた位相付けと refmac[40]による構造精密化をすべて自動で行うことが可能なプログラムである。しかし、回折強度測定と位相付けの間では、ファイル変換等を手作業で行う必要があり、これらの操作は一般的に解析者自身がそれぞれの回折強度データに対して行う場合が多い。そこで、本研究では DeepCentering による自動センタリングと、既に開発がされている、自動回折イメージ処理プログラム KAMO と、自動位相付け兼精密化パイプラインである dimple を組み合わせることで、X線回折強度測定から構造精密化までをシームレスに繋ぐためのパイプライン、imaproc を開発した。プログラミング言語である Python2.7 を用いて開発を行った。imaproc の処理スキームを Figure 2-6 に示す。imaproc の主要な役割は、KAMO が出力する回折イメージの各回折点の強度の積分結果をアスキー形式で記録した XDS\_ASCII.HKL ファイルを自動で変換し、dimple が input とファイルとして指定するバイナリ形式のタンパク質 X線結晶構造解析における標準回折強度データファイルフォーマットである MTZ ファイルに変換することである。X線回折強度測定の測定条件については、ビームライン装置制御プログラムである BSS によって、測定サンプルの数だけ任意の条件を一括で登録する。

具体的な imaproc の動作は次の通りである。

①KAMO が出力する反射データである XDS\_ASCII.HKL のディスクへの書き込みを常に監視しつづける。

②XDS\_ASCII.HKL が生成されると、データ変換プログラムである xdsconv[41]を自動で実行し、MTZ ファイルに変換する。

③作成された MTZ ファイルに FreeR flag[42]と呼ばれる、構造精密化の際に必要な処理を加える。

④ ③で作成された MTZ ファイルに対して、自動で dimple を実行する。

以上の工程を、X線回折強度測定の間、常に実行しつづけることで自動処理を行う。

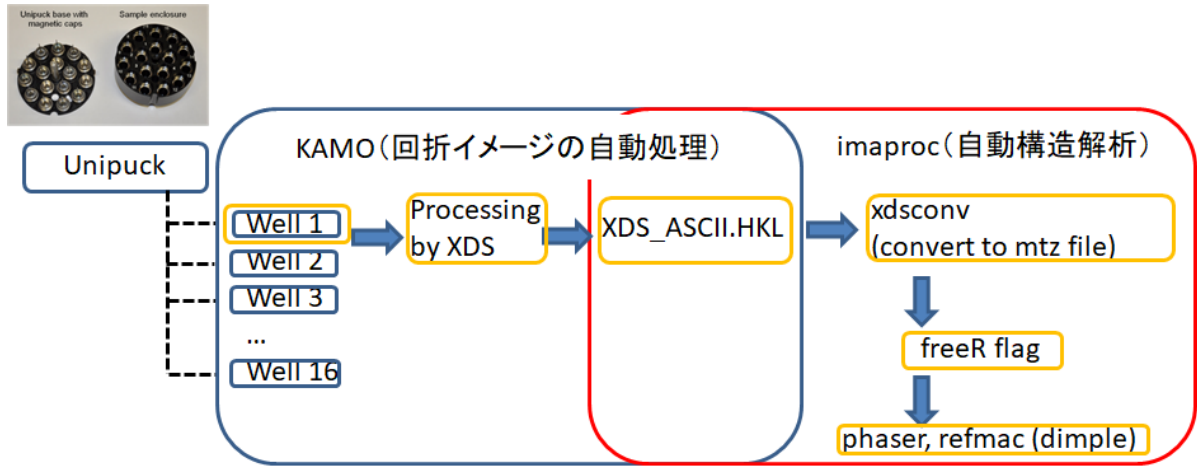


Figure 2-6 自動構造解析パイプライン imaproc の処理スキーム

DeepCentering による自動測定に次いで、imaproc は以下のスキームで処理を行う。

- ① 回折強度測定が終了すると、自動的に KAMO が回折イメージデータを処理する。
- ② imaproc が KAMO の出力である XDS\_ASCII.HKL を検出し、dimple を使用するためのファイル変換等を行う。
- ③ dimple による位相付けおよび構造精密化を行う。

## 2-2-6 DeepCentering と imaproc を用いたニワトリ卵白由来リゾチームの自動構造解析

DeepCentering と imaproc を用いた自動構造解析パイプラインの検証実験のため、ニワトリ卵白由来リゾチーム結晶の自動構造解析を行った。比較のために、手で結晶をセンタリングした構造解析結果と比較した。リゾチームはナカライテスクから購入し、購入品をそのまま結晶化実験に使用した。リゾチームを 50 mg/mL になるように 50 mM 酢酸緩衝液 (pH 4.6) に溶解した。結晶化はバッチ法を用いて行った。100  $\mu$ L タンパク質溶液と 100  $\mu$ L 1.0 M 塩化ナトリウムをマイクロチューブで混合し、20°C で静置した。得られた結晶に 20% Glycerol を抗凍結剤として添加し、液体窒素で凍結処理を行った。DeepCentering と手動センタリングの比較のために、リゾチーム結晶を 15 個調製し、各結晶に対して、DeepCentering と手動センタリングで X 線回折強度測定を行った。すべての結晶について、DeepCentering を用いた測定を先に行い、結晶学的統計値を比較した。次に示す測定条件で X 線回折強度測定を行った；振動範囲 180°、振動角 0.5°、露光時間 0.5 秒、波長 1 Å、カメラ長 150 mm この時の 1 測定あたりの放射線量は 1.17 MGy であった。ここで用いた Gy は、放射線の吸収エネルギーであり、1 Gy は放射線によって 1 kg の物質に 1 J のエネルギーが吸収されたときの吸収線量と定義される。本実験における 1 測定あたりの放射線量は、Henderson Limit[43]として知られる、凍結結晶において回折能が半減するおおよその値である 20 MGy よりも非常に小さい値であるから、測定順が結晶学的統計値に与える影響は小さいと考えられた。



## 2-3 結果と考察

### 2-3-1 LoopDetector および CrystalDetector の検出テスト

検出テストの結果、検出精度は、LoopDetector では 100%であった。Figure 2-7 に検出テスト結果の例を示す。興味深いことに、LoopDetector では、結晶観察カメラの視野中にクライオループの一部分のみしか存在していない場合においても、問題なく検出することが可能であった。また、従来用いられていたクライオループ検出プログラムでは、一日毎など、一定期間毎にバックグラウンド画像を撮影する必要があったが、LoopDetector では、バックグラウンド画像を用いなくても高精度のクライオループの検出が可能であった。同様に、画像処理技術を用いた従来法では、結晶観察カメラの倍率を変更すると、カメラの輝度やコントラストが変わり、クライオループの検出が失敗する場合があったが、これについても DeepCentering を用いたクライオループ検出では高精度の検出結果を残した。このように、DeepCentering を用いたクライオループの検出では、プログラム使用時に、結晶観察カメラ等の細かなパラメータ設定が不要であった。これはプログラム使用環境の維持という点で重要であり、その省力化に貢献したといえる。

結晶検出のための CrystalDetector においては、検出精度が 90.8 %であった。さらに、従来法では検出が非常に困難であった、クライオループ上に霜が付着している場合や、結晶の外形が曖昧である場合においても検出が可能であった (Figure 2-8)。一方で、結晶検出精度に関しては、約 10 %の結晶画像において検出を失敗した。検出が不可能であった理由として、①霜の付着や結晶化溶媒によって結晶が全く結晶観察カメラで確認できなかった ② 結晶の境界が極端にあいまいであったこと等が考えられる (Figure 2-8(b))。本研究では、効率的なニューラルネットワークの作成のために実際の結晶画像を用いずに多角形の画像を作成し学習データとしたため、上記②のような場合の結晶の検出精度が低かったと考えられる。今後、さらに結晶検出の精度を向上させる方法として、一般的な深層学習の方法と同じよう

に、膨大な数の結晶画像を集めて、新たにニューラルネットワークを作成することや、今回開発した CrystalDetector を基に、ニューラルネットワークを改良すること等が考えられる。

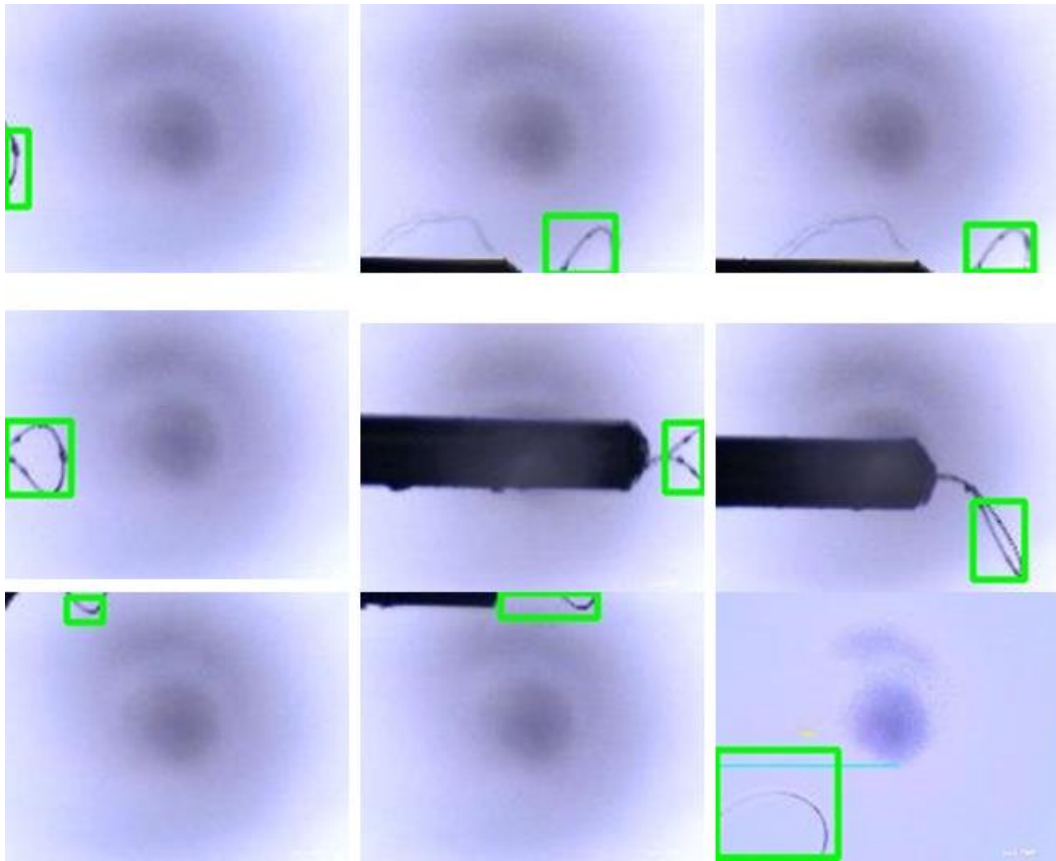


Figure 2-7 LoopDetector による検出結果

LoopDetector によって検出されたクライオループ位置を緑色矩形で示した。

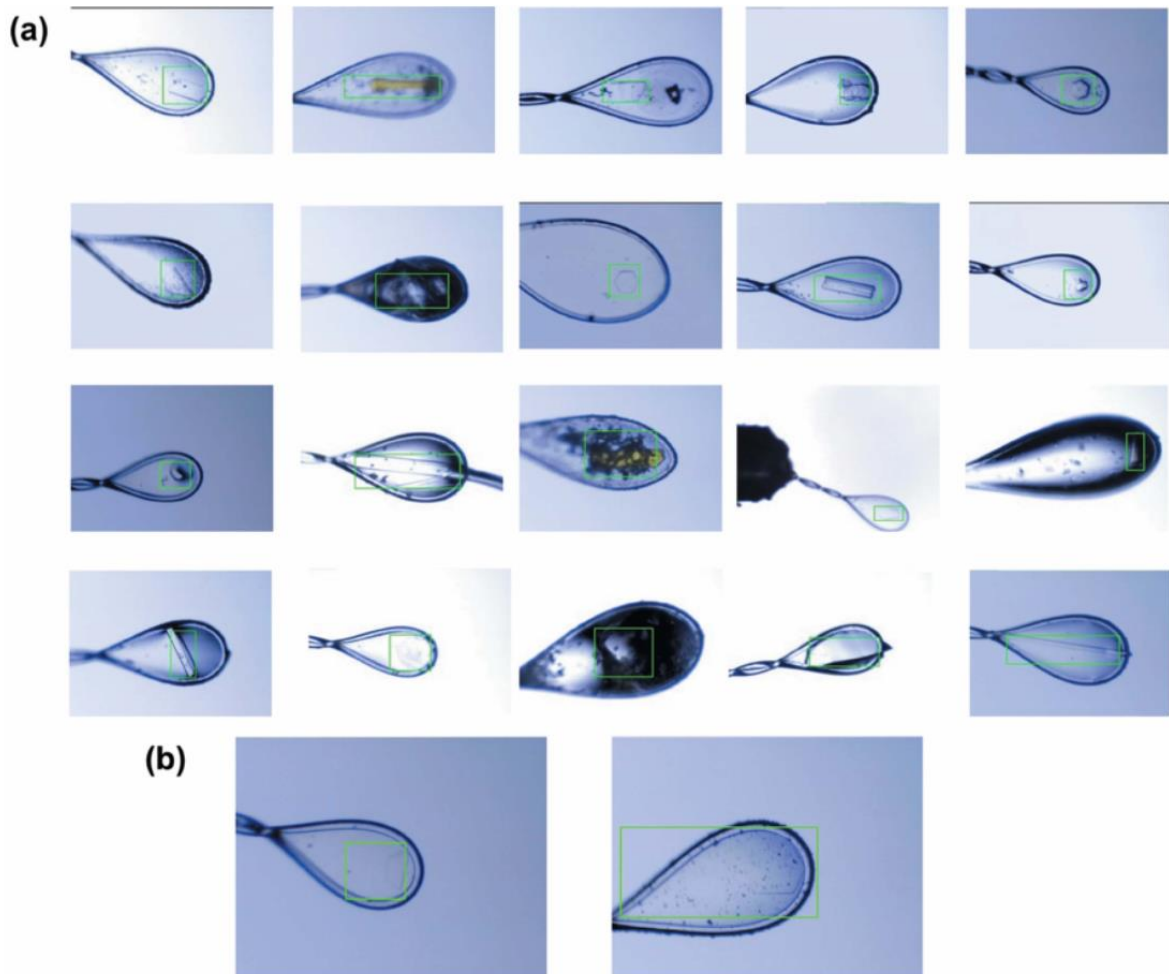


Figure 2-8 CrystalDetector による検出結果

検出の成功例(a)および失敗例(b)を示す。CrystalDetector によって検出された結晶位置を緑色矩形で示した。

### 2-3-2 自動構造解析パイプラインによるリゾチームの構造解析

DeepCentering および imaproc を用いた自動構造解析の統計値を Table 2-2 に示す。15 個のリゾチーム結晶すべてについて、問題なくセンタリングがされた。測定を行った多くのデータセットの最外殻のデータ完全性が低かった。これは受光面が正方形の X 線検出器のコーナーを、データ処理の最大分解能としたため、高分解能の領域において、特定のブラッグ反射が測定できなかったことによる。手作業で結晶センタリングを行った場合での構造解析結果と比較すると、多くの場合で自動センタリングと手動センタリングで構造解析の統計値に大きな差はみられなかった。

一部結晶では、自動センタリングで測定を行った場合に、 $R_{\text{meas}}$  の値の増加していた。これらの場合では、結晶の中心位置が他の結晶の場合よりも、DeepCentering の場合でわずかにずれ、手動センタリングではより結晶の厚みのある部分が回折に寄与したため、結果として  $I/\text{sig}(I)$  の値が手動センタリングの場合の方が大きくなり、 $R_{\text{meas}}$  の値に差が生じた考えられる (Figure 2-9)。一方で、構造精密化の指標である  $R_{\text{factor}}$  および  $R_{\text{free}}$  の平均値については、DeepCentering ( $R_{\text{factor}}/R_{\text{free}} = 27.9/30.0$ ) と手動センタリング ( $R_{\text{factor}}/R_{\text{free}} = 28.0/29.7$ ) で大きな差は見られなかった。以上の結果から、DeepCentering によるセンタリングは、手動センタリングよりも精度が劣る場合があるが、その誤差は構造解析の結果に大きな影響を及ぼすレベルではないと考えられた。

Table 2-2 自動センタリング(a)と手動センタリング(b)での構造解析統計値の比較  
 表中の括弧内の数字は最外殻の値を示す。測定は以下の条件で行った。

振動範囲 180°、振動角 0.5°、露光時間 0.5 秒、波長 1 Å、カメラ長 150 mm

(a) DeepCentering と imaproc による自動構造解析結果の結晶学的統計値

Crystal No	Resolution (Å)	Completeness (%)	Redundancy	I / sig(I)	Rmeas (%)	Space group	Unit Cell (Å)	Rfactor (%)	Rfree (%)
1	360 x 420	50.0 - 1.27	80.7 (32.2)	21.07 (2.64)	5.7 (23.8)	$P 4_3 2_1 2$	a = b = 78.81 c = 37.15	28.9	29.9
2	480 x 520	50.0 - 1.27	80.3 (26.6)	15.39 (2.14)	7.1 (25.2)	$P 4_3 2_1 2$	a = b = 79.01 c = 37.02	29.5	31.6
3	360 x 350	50.0 - 1.27	87.5 (37.9)	17.05 (2.44)	7.1 (19.4)	$P 4_3 2_1 2$	a = b = 78.78 c = 37.02	27.9	28.8
4	280 x 400	50.0 - 1.27	88.4 (42.1)	20.19 (1.92)	5.6 (31.9)	$P 4_3 2_1 2$	a = b = 78.55 c = 36.98	27.4	28.3
5	240 x 240	50.0 - 1.28	85.5 (41.8)	15.12 (1.79)	15.12 (1.79)	$P 4_3 2_1 2$	a = b = 79.06 c = 36.92	27.3	29.5
6	120 x 200	50.0 - 1.36	96.3 (80.8)	19.00 (1.95)	6.5 (61.7)	$P 4_3 2_1 2$	a = b = 78.50 c = 36.90	29.1	31.6
7	240 x 250	50.0 - 1.38	92.5 (65.5)	20.28 (1.60)	5.1 (75.8)	$P 4_3 2_1 2$	a = b = 78.03 c = 37.23	27.9	28.8
8	55 x 255	50.0 - 1.60	99.9 (99.6)	12.65 (1.23)	12.1 (192.5)	$P 4_3 2_1 2$	a = b = 78.72 c = 36.94	27.9	33.4
9	80 x 240	50.0 - 1.80	99.7 (98.3)	16.77 (1.91)	8.8 (116.2)	$P 4_3 2_1 2$	a = b = 78.56 c = 37.03	26.8	28.8
10	40 x 80	50.0 - 1.38	98.4 (81.8)	18.84 (1.41)	5.6 (86.0)	$P 4_3 2_1 2$	a = b = 78.72 c = 36.94	27.1	29.7
11	70 x 280	50.0 - 1.49	96.0 (79.3)	16.76 (1.54)	6.9 (107.9)	$P 4_3 2_1 2$	a = b = 78.56 c = 37.03	28.1	29.5
12	60 x 270	50.0 - 1.54	99.9 (98.7)	15.40 (1.53)	9.5 (123.9)	$P 4_3 2_1 2$	a = b = 78.78 c = 36.84	27.5	31.1
13	50 x 120	50.0 - 1.42	99.7 (98.2)	16.32 (1.68)	7.9 (80.0)	$P 4_3 2_1 2$	a = b = 78.93 c = 36.38	28.2	30.1
14	180 x 150	50.0 - 1.38	98.3 (89.6)	24.77 (1.62)	3.9 (63.9)	$P 4_3 2_1 2$	a = b = 79.07 c = 36.20	27.6	29.1
15	120 x 240	50.0 - 1.28	88.0 (43.9)	18.31 (2.30)	6.4 (30.4)	$P 4_3 2_1 2$	a = b = 78.65 c = 37.03	27.3	29.2

## (b) 手動センタリングと imaproc による構造解析結果の結晶学的統計値

Crystal No	Resolution (Å)	Completeness (%)	Redundancy	I / sig(I)	Rmeas (%)	Space group	Unit Cell (Å)	Rfactor (%)	Rfree (%)
1	360 x 420	50.0 - 1.27	81.9 (35.6)	23.09 (3.24)	5.4 (22.0)	$P 4_3 2_1 2$	a = b = 78.81 c = 37.15	28.9	29.9
2	480 x 520	50.0 - 1.27	80.7 (27.9)	25.14 (4.28)	4.4 (15.6)	$P 4_3 2_1 2$	a = b = 79.01 c = 37.02	29.5	31.6
3	360 x 350	50.0 - 1.27	80.8 (28.1)	24.95 (4.08)	4.4 (16.3)	$P 4_3 2_1 2$	a = b = 78.78 c = 37.02	27.9	28.8
4	280 x 400	50.0 - 1.27	89.4 (43.7)	21.82 (2.08)	5.2 (31.4)	$P 4_3 2_1 2$	a = b = 78.55 c = 36.98	27.4	28.3
5	240 x 240	50.0 - 1.27	82.5 (31.3)	15.59 (2.20)	8.8 (29.5)	$P 4_3 2_1 2$	a = b = 79.06 c = 36.92	27.3	29.5
6	120 x 200	50.0 - 1.35	85.2 (33.6)	31.14 (3.28)	3.6 (21.7)	$P 4_3 2_1 2$	a = b = 78.50 c = 36.90	29.1	31.6
7	240 x 250	50.0 - 1.38	90.4 (60.1)	13.00 (1.05)	7.7 (96.4)	$P 4_3 2_1 2$	a = b = 78.03 c = 37.23	27.9	28.8
8	55 x 255	50.0 - 1.68	99.9 (99.7)	15.20 (1.46)	9.7 (157.4)	$P 4_3 2_1 2$	a = b = 78.72 c = 36.94	27.9	33.4
9	80 x 240	50.0 - 1.57	99.7 (98.1)	17.86 (1.92)	7.0 (105.5)	$P 4_3 2_1 2$	a = b = 78.56 c = 37.03	26.8	28.8
10	40 x 80	50.0 - 1.41	97.6 (87.3)	18.62 (1.73)	6.0 (76.1)	$P 4_3 2_1 2$	a = b = 78.78 c = 36.84	27.1	29.7
11	70 x 280	50.0 - 1.44	93.1 (66.3)	16.81 (1.26)	6.7 (120.6)	$P 4_3 2_1 2$	a = b = 78.56 c = 37.03	28.1	29.5
12	60 x 270	50.0 - 1.52	100.0 (99.9)	14.87 (1.42)	8.1 (130.3)	$P 4_3 2_1 2$	a = b = 78.78 c = 36.84	27.5	31.1
13	50 x 120	50.0 - 1.38	98.3 (90.2)	17.32 (1.58)	6.1 (61.4)	$P 4_3 2_1 2$	a = b = 79.07 c = 36.20	28.2	30.1
14	180 x 150	50.0 - 1.28	90.0 (47.8)	24.50 (1.62)	4.1 (51.4)	$P 4_3 2_1 2$	a = b = 78.56 c = 37.08	27.6	29.1
15	120 x 240	50.0 - 1.27	85.4 (33.6)	17.35 (1.66)	6.9 (42.0)	$P 4_3 2_1 2$	a = b = 78.65 c = 37.03	27.3	29.2

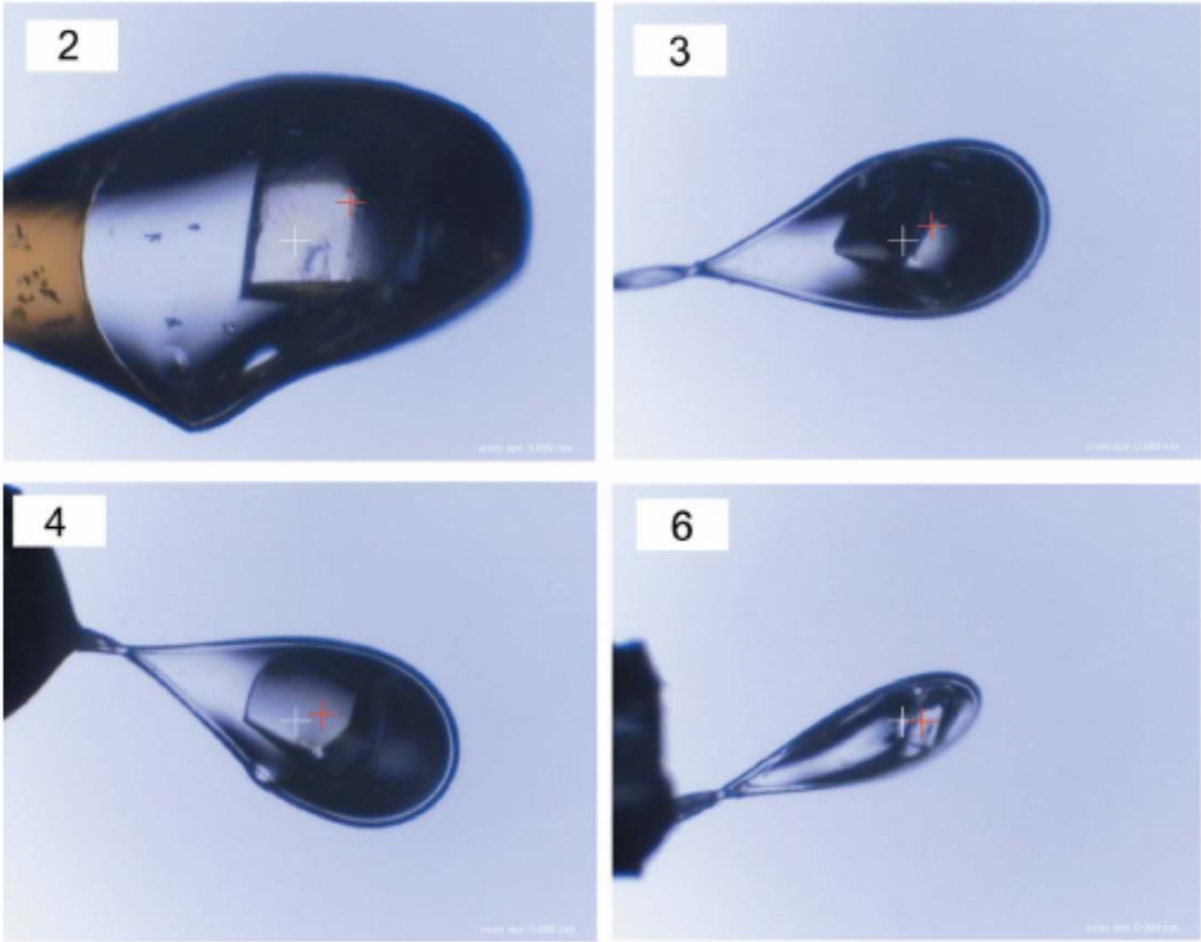


Figure 2-9 自動、手動センタリングで、結晶学的統計値に差の出た結晶センタリング位置の比較

手動センタリング位置を赤十字、自動センタリング位置を白十字で示す。各画像の左上の数字が結晶番号を表す。手動センタリングの場合の方が、結晶のより厚みのある部分にセンタリングされていた。



## 2-4 結論

深層学習を用いた自動センタリングプログラム DeepCentering を開発し、画像処理技術による高精度での自動センタリングに成功した。また自動構造解析パイプライン imaproc を開発し、標準結晶であるリゾチーム結晶を用いて、自動構造解析が問題なく実施できることを示した。これらのシステムの開発により、大型結晶を対象とする、回折強度測定実験の自動化、加速化、省力化に貢献できると期待される。また、DeepCentering はX線を用いない自動センタリング方法であることから、X線損傷による影響がより問題になる、常温測定での結晶位置の自動検出にも応用が期待される。

## 2-5 参考文献

- [1] Li Wan, Matthew Zeiler, Sixin Zhang, Yann L. Cun, & Rob Fergus. *JMLR Workshop and Conference Proceedings*. (2013). 1058–1066.
- [2] Visual Object Classes Challenge 2012 (VOC2012)  
<http://host.robots.ox.ac.uk/pascal/VOC/voc2012/>
- [3] Ross Girshick, Jeff Donahue, Trevor Darrell, & Jitendra Malik. *Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation*. 580–587.
- [4] SHaoqing Ren, Kaiming He, Ross Grishick, & Jian Sun. *Advances in Neural Information Processing Systems* **28**, 91–99.
- [5] Jonathan Long, Evan Shelhamer, & Trevor Darrell. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. (2015).
- [6] Oriol Vinyals, Alexander Toshev, Samy Bengio & Dumitru Erhan. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [7] Leon A. Gatys, Alexander S. Ecker, & Matthias Bethge. *arXiv:1508.06576* (2015).
- [8] Alec Radford, Luke Metz, & Soumith Chintala. *arXiv:1511.06434* (2015).
- [9] Vijay Badrinarayanan, Kendall & Roberto Cipolla *arXiv:1511.00561* (2015).
- [10] <http://mi.eng.cam.ac.uk/projects/segnet/>
- [11] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Ptersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg & Demis Hassabis. *Nature*. (2015). **518**, 529–533.

- [12] David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis antonoglou, Veda Panneershelvam, Marc Lanctot, SNder Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel & Demis Hassabis. *Nature*. (2016). **529**, 484-489.
- [13] F. ROSENBLATT. *Psychological Review*. (1958). **65**, 386-408.
- [14] S. Amari. *Proceedings of the IEEE*. **3**, 299-307.
- [15] Auther L. Samuel *IBM Journal of Research and Development*. (1959). **3**, 221-229.
- [16] Ke, T.-W., Brewster, A. S., Yu, S. X., Ushizima, D., Yang, C. & Sauter, N. K. *J. Synchrotron Rad.* **25**, (2018). 655-670.
- [17] Bruno, A. E., Charbonneau, P., Newman, J., Snell, E. H., So, D. R., Vanhoucke, V., Watkins, C. J., Williams, S. & Wilson, J. *PLoS One*. (2018). **13**, e0198883.
- [18] Song, J., Mathew, D., Jacob, S. A., Corbett, L., Moorhead, P. & Soltis, S. M. (2007). *J. Synchrotron Rad.* **14**, 191-195.
- [19] Asanov, A. N., McDonald, H. M., Oldham, P. B., Jedrzejewski, M. J. & Wilson, W. (2001). *J. Cryst. Growth*. **232**, 603-609.
- [20] Bourgeois, D., Vernede, X., Adam, V., Fioravanti, E. & Ursby, T. (2002). *J. Appl. Cryst.* **35**, 319-326.
- [21] Pohl, E., Ristau, U., Gehrman, T., Jahn, D., Robrahn, B., Malthan, D., Dobler, H. & Hermes, C. (2004). *J. Synchrotron Rad.* **11**, 372-377.
- [22] Vernede, X., Lavault, B., Ohana, J., Nurizzo, D., Joly, J., Jacquamet, L., Felisaz, F., Cipriani, F. & Bourgeois, D. (2006). *Acta Cryst. D*. **62**, 253-261.

- [23] Chavas, L. M. G., Yamada, Y., Hiraki, M., Igarashi, N., Matsugaki, N. & Wakatsuki, S. (2011). *J. Synchrotron Rad.* **18**, 11-15.
- [24] Madden, J. T., DeWalt, E. L. & Simpson, G. J. (2011). *Acta Cryst. D.* **67**, 839-846.
- [25] Ito, S., Ueno, Go., & Yamamoto, M. (2019). *J. Synchrotron Rad.* **26**, 1361-1366.
- [26] Okazaki, N., Hasegawa, K., Ueno, G., Murakami, H., Kumasaka, T. & Yamamoto, M. (2008). *J. Synchrotron Rad.* **15**, 288-291.
- [27] Murakami, H., Ueno, G., Shimizu, N., Kumasaka, T. & Yamamoto, M. (2012). *J. Appl. Cryst.* **45**, 234-238.
- [28] <https://pillow.readthedocs.io/en/stable/>
- [29] <https://github.com/tensorflow/tensorflow>
- [30] <https://docs.chainer.org/en/stable/index.html>
- [31] <https://caffe.berkeleyvision.org/>
- [32] <https://keras.io/about/>
- [33] Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., Fischer, I., Wojna, Z., Song, Y., Guadarrama, S. & Murphy, K. (2017). 2017 Conference on Computer Vision and Pattern Recognition (CVPR 2017), 21-26 July 2017, Honolulu, HI, USA.
- [34] J. Redmon and A. Farhadi. *arXiv*. (2018). **1804.02767**.
- [35] Ren, S., He, K., Girshick, R., Sun, J. *NIPS* (2015).
- [36] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y. & Berg, A. C. (2016). Proceedings of the 14th European Conference on Computer Vision (ECCV'16), 8-16 October 2016, Amsterdam, The Netherlands.
- [37] <https://www.diamond.ac.uk>
- [38] Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P.,

Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Cryst. D.* **67**, 235-242.

[39] McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. *J. Appl. Cryst.* (2007). **40**, 658-674.

[40] Vagin, AA, Steiner, RS, Lebedev, AA, Potterton, L, McNicholas, S, Long, F and Murshudov, GN. *Acta Cryst D.* (2004). D. **60**, 2284-2295.

[41] W. Kabsch *Acta Crystallogr. D.* (2010). **66**, 125-132.

[42] Brünger, A. T. *Nature.* (1992). **355**, 472-475.

[43] Henderson, R. *Proc. R. Soc. Lond. B.* (1990). 241, 6-8.

## 第3章

DeepCentering を用いたヒト血清アルブミンと  
環状ペプチドの複合体構造スクリーニング

### 3-1 ヒト血清アルブミンと環状ペプチドの創薬への応用

前章において、大型結晶を対象とした自動センタリングシステム DeepCentering の開発について述べた。本章では、DeepCentering を用いてヒト血清アルブミン（以下 HSA）と種々の環状ペプチドとの複合体結晶スクリーニング、およびアポ結晶のスクリーニング実験を行った応用例について述べる。

近年、医薬品の新たなモダリティとして環状ペプチドが注目されている[1, 2]。環状ペプチドが注目されている理由として、①抗体医薬品[3]のように標的タンパク質に非常に強く結合できること、②タンパク質間相互作用面に代表されるような平坦で浅い相互作用面に結合できること、③低分子化合物のように経口投与可能であることが挙げられる。

環状ペプチド創製のための高速スクリーニングシステムの確立により、環状ペプチドは創薬研究において、特に魅力的なものとなっている。これまでに開発された高速スクリーニングシステムには、*in vivo* 産生に基づく *de novo* 創薬アプローチとして、ファージディスプレイ[4, 5]や SICLOPPS[6, 7]などがある。さらに、Passioura らは、mRNA ディスプレイに基づく RaPID[8]と呼ばれる別の方法を開発した。このシステムでは、非常に多くのユニークな配列（最大  $10^{14}$ ）をスクリーニングすることが可能であり、従来用いられていた、ファージディスプレイや *in vivo* 生産に基づく他のシステムでは、 $10^8$ - $10^9$  しかスクリーニングができないことを考慮すると、RaPID により膨大な数のスクリーニングが可能となった。また、RaPID は、D-アミノ酸、 $\beta$ -アミノ酸、主鎖修飾、非天然アミノ酸を含む環状ペプチドのスクリーニングも可能であり、天然および非天然アミノ酸を含む環状ペプチドを多数の組み合わせからスクリーニングすることで、様々なタンパク質に対して高い結合能を示すペプチドを同定できることが報告されている[9, 10]。さらに、従来のペプチド医薬品はタンパク質分解の問題を抱えていたが、環状ペプチドに含まれる主鎖構造の修飾や非天然なアミノ酸により、分解の影響を受けにくくなっている点も注目されている[11-13]。現在、40 種類以上の環状ペプチドが上市されており、さらに 20

種類以上の環状ペプチドが臨床研究の段階にある[14]。

環状ペプチドを用いた創薬が新たなモダリティとして非常に注目を集めている一方で、吸収、分布、代謝、排泄、毒性 (ADMET) の観点からいくつかの課題に直面している。例えば、環状ペプチドを治療薬として開発する場合、腎排泄が問題となる。一般に、外因性ペプチドは腎臓で迅速に濾過される。腎臓の糸球体は多孔質構造を有しており、分子量 5-10 kDa 未満の分子量の血漿分子を効率的にろ過する[15]。ヒト血清アルブミン (HSA) は血液中に最も多く存在する輸送タンパク質であり、生体内の多くの物質輸送を担っている。HSA は糸球体濾過の影響を受けず、血中半減期が非常に長い (約 20 日) ため、外因性物質が腎臓でろ過される腎クリアランスを回避するためには HSA との結合が重要である。これまでに、アルブミンへの結合を増加させることを目的とした方法が開発され、ペプチド医薬品に使用されてきた。例えば、exendin-4-albumin conjugate (CJC-1143-PC) や GLP-1 では、ケミカルリンカーを用いたアルブミンとの共役化が行われた[16-19]。他の研究では、アルブミン結合剤として知られている分子との結合により、ペプチド医薬品の腎クリアランスを回避することに成功している。

しかしながら、現在のところ、環状ペプチドと HSA との複合体の構造は Protein Data Bank (PDB) には一つも登録されていない。利用可能な構造情報が存在しないため、HSA の環状ペプチド結合部位およびその結合の相互作用メカニズムは未解明である。

Dalbavancin[20-22]は、Telavancin、oritavancin、Teicoplanin、および Vancomycin のような他の類似の薬剤とともに、7 残基を有する共通のヘプタペプチドコアを含む環状リポグリコペプチドである。Dalbavancin は、高い血漿タンパク質結合能 (～93～98%) と長い半減期 (～6～11 日) を有しており、血漿タンパク質の約 60% を占める HSA と環状ペプチドの複合体構造を得るための有力な候補となっている。HSA-環状ペプチド複合体の構造が明らかになれば、環状ペプチドの創製における、高速スクリーニングの初期段階ですでに環状ペプチドの腎クリアランスの問題を解決すること



が可能になると期待される。そこで、本研究では HSA と Dalbavancin をはじめとする、種々の環状ペプチドの複合体構造解析を目指し、アポ結晶およびソーキング法により調製された HSA-環状ペプチド複合体結晶の DeepCentering を用いた回折強度測定スクリーニングおよび HSA と種々の化合物の微小共結晶のアンジュレータービームラインでのスクリーニングを行い、さらに構造解析を行った。本章では、これらのスクリーニングによって、①で得られた HSA アポ結晶構造および、②で得られた HSA-Dalbavancin 複合体構造を報告する。アポ構造と複合体構造の比較によって明らかになった環状ペプチドの結合様式の議論を行う。

## 3-2. 材料と方法

### 3-2-1. ヒト血清アルブミンのアポ結晶の調製

Sigma-Aldrich より凍結乾燥された HSA を購入した。アポ結晶の調製では、凍結乾燥 HSA を濃度が 100 mg/mL となるように、50 mM potassium phosphate (pH 5.25) に溶解した。凝集体を取り除くため、15,000 x *g*, 15 min、4°C で遠心し、上清を結晶化に用いた。リザーバーの溶液組成は、5 mM sodium azide、38% (v/v) polyethylene glycol 400 であった。シッティングドロップによる蒸気拡散法[23]で結晶化を行った。5  $\mu$ L のタンパク質溶液と 5  $\mu$ L の沈殿剤溶液を混合した。混合溶液に対して、予備実験で得られた微小結晶をシードとして用い、Seeding Tool (HAMPTON RESEARCH) を使ってシーディングを行った。結晶化は 20 °C で行った。結晶化後 1 日以内に結晶が析出した。析出した結晶サイズは、500  $\mu$ m 程度の大型結晶であった。

### 3-2-2 ヒト血清アルブミン-環状ペプチド複合体結晶の調製

HSA-Dalbavancin 複合体の結晶調製は、共結晶化法[24, 25]を用いた。HSA はバッファー A (50 mM potassium phosphate (pH 7.0) and 150 mM NaCl) に溶解した。凝集体を取り除くため、15,000 x *g*, 15min, 4°C で遠心した。その上清を、あらかじめバッファー A で平衡化した Superdex 200 10/300 カラムに添加し、精製した。結晶化のため、限外ろ過を繰り返して 50 mM NaCl (pH 7.0) のバッファーに置換するとともに、100 mg/mL に濃縮した。共結晶化には Figure 3-1 の環状ペプチドを用いた。各環状ペプチド溶液の組成が、50 mM potassium phosphate (pH 7.0)、150 mM NaCl、20% (v/v) ジメチルスルホキシド、50 mM Dalbavancin となるように調製した。共結晶化は環状ペプチドと HSA のモル比が 4:1 となるようにタンパク質溶液と Dalbavancin 溶液を混合し、4°C で 12 時間静置した。ただし、Ranreotide については、Ranreotide と HSA のモル比が 1.5 : 1 になるように溶液を混合した。リザーバーの組成は、0.1

M Sodium citrate tribasic dehydrate (pH 5.0) containing 30% (v/v) Jeffamine ED-2001 (pH 7.0)であった。結晶化は、Dalbavancin 添加後のタンパク質溶液とリザーバー溶液をそれぞれ 200 nL ずつ混合し、シッティングドロップによる蒸気拡散法を用いて 4°C で結晶化を行った。結晶化後、約 1 週間で結晶が析出した。析出した結晶サイズは 10-30  $\mu\text{m}$  の微結晶であった。

### 3-2-3-1 DeepCentering を用いた HSA アポ結晶のスクリーニングおよび構造解析

HSA アポ結晶を用いた予備的な回折強度測定から、HSA アポ結晶の溶媒含有率は 76% であった。一般的に、溶媒含有率の高い結晶は結晶性が悪く、本研究の場合においても、同一の結晶化条件で得られた結晶の品質差が大きいことが予想された。このような場合、X線回折能の高い結晶を探索するために多くの結晶に対してX線回折強度測定を行う必要がある。そこで、DeepCentering および imaproc を用いて、実験の省力化をしつつ HSA アポ結晶のスクリーニング実験を行った。本研究では、3 Å 以上での結晶構造解析を目的とし、128 個のアポ結晶について、DeepCentering および imaproc を用いて、SPring-8 の構造ゲノムビームライン BL26B2[26]で自動スクリーニングを行った。その結果、最も分解能の高い結晶は、2.6 Å 分解能まで回折した。得られた X線回折データは imaproc で自動構造解析を行った。imaproc 中で行われる位相付けのサーチモデルには、既に構造解析がされている HSA の構造 (PDB ID: 1A06[27]) を用いた。

### 3-2-3-2 DeepCentering を用いた HSA-環状ペプチドソーキング結晶のスクリーニング

3-2-1 で調製したアポ結晶と種々の環状ペプチドを用いてソーキング実験を行った。ソーキングに用いた環状ペプチドは、血漿タンパク結合 (PPB) 値が 80 %以上の環状ペプチドの中から選定した。ソーキングには共結晶化実験と同一の環状ペプチドを用い

た。25 mM potassium phosphate (pH 5.25)、2.5 mM sodium azide, and 20% (v/v) polyethylene glycol 400、20 % (v/v) dimethyl sulfoxide を母液とし、各環状ペプチドの濃度が 50 mM になるように、ソーキング溶液を調製した。ただし、Romidepsin については、溶解度が低いため、濃度を 10 mM としてソーキング溶液を調製した。ソーキング時間は 1、3、5、10、30、60、120、180 分で行った。アポ結晶では、ソーキング溶液に polyethylene glycol 400 が含まれているため、追加の抗凍結剤は添加せず、そのまま結晶を凍結した。X線回折強度測定は DeepCentering と imaproc を用いて、SPring-8 の構造ゲノムビームライン BL26B2 で自動スクリーニングを行った。imaproc による構造解析結果を用いて各環状ペプチドの結合の有無を確認した。その結果、化合物由来の電子密度が確認できなかったため、次項に示すように HSA-環状ペプチドの共結晶化実験を行った。

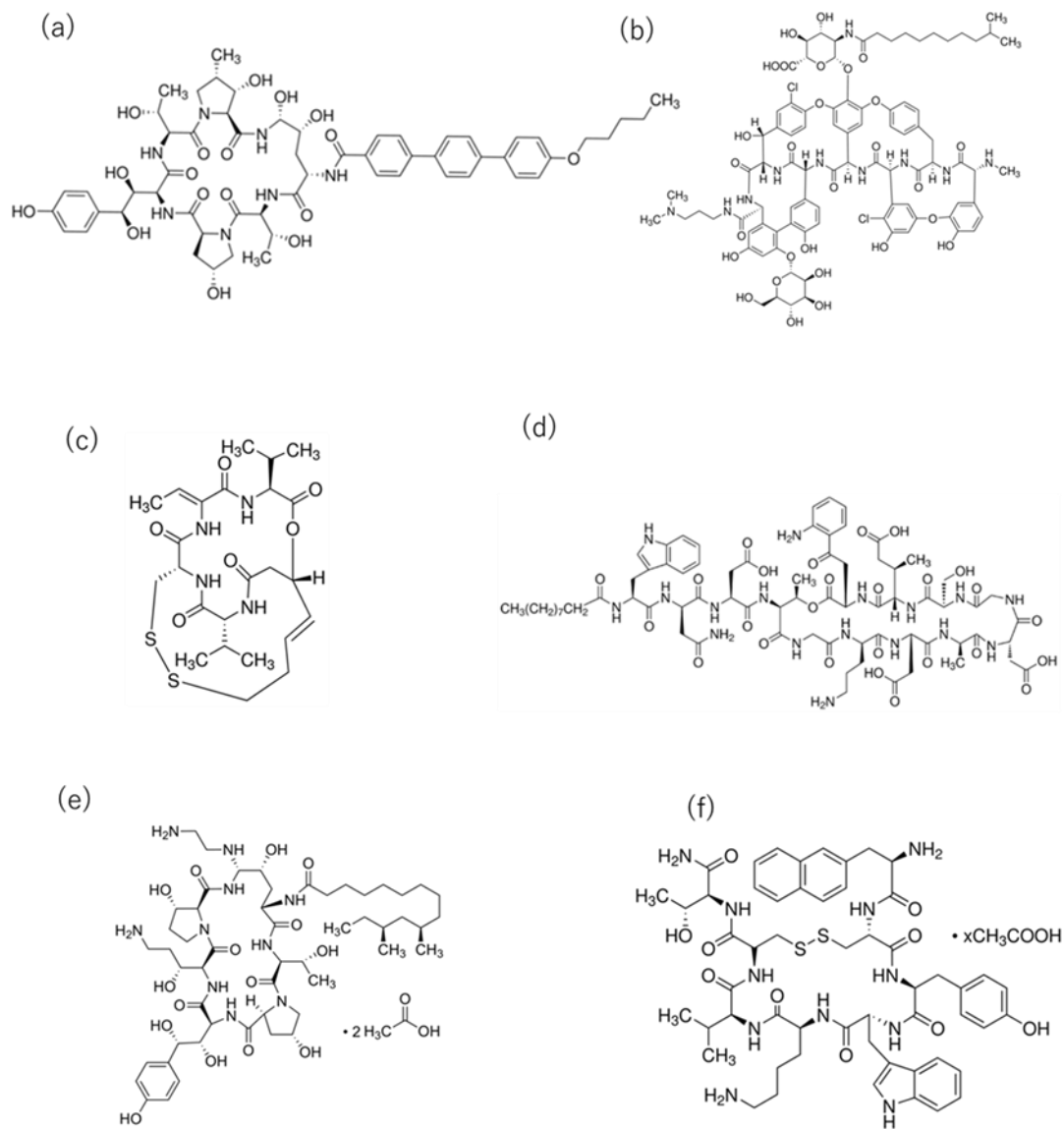


Figure 3-1 ソーキングおよび共結晶化に用いた環状ペプチド

- (a) Anidulafungin (b) Dalbavancin (c) Romidepcin (d) Daptomycin (e) Capsosfungin  
 (f) Lanreotide

### 3-2-4. HSA-Dalbavancin 複合体結晶の構造解析

共結晶化による種々の環状ペプチドと HSA との複合体結晶化スクリーニングの結果、Dalbavancin の場合でのみ結晶が析出した。HSA-Dalbavancin 複合体の結晶サイズは 10-30  $\mu\text{m}$  程度であった。①DeepCentering の教師データとして用いた多角形画像は、結晶サイズに換算するとそのすべてが 50  $\mu\text{m}$  以上であること、②微小結晶の検出のためには、検出に用いる画像解像度が低く、高い検出精度が見込めないことから、DeepCentering の適用が難しいと考えられた。そのため、SPring-8 BL32XU[28]において、自動データ収集システム ZOO を用いてデータ収集を行った。151 個の複合体結晶を、1 結晶当たり振動角 10° 分をそれぞれの結晶について測定した。収集した回折イメージは KAMO を用いて処理し、規格化構造因子を用いた相関係数に基づくクラスタリングによって 123 結晶の回折データが選出され、XSCALE[29]によってデータマージを行った。Phaser[30]を用いて分子置換法で位相決定を行った。サーチモデルには PDB ID: 1A06 を用いた。その後、Coot[31]を用いて手動でモデル改良を行い、phenix.refine[32]により精密化を行った。精密化後、Dalbavancin 由来の電子密度が明瞭に観測された。Coot を用いて電子密度に Dalbavancin をフィッティングし、再度 phenix.refine を用いて構造精密化を行った。データ処理および精密化の統計値を Table 3-1 に示す。

Table 3-1 データ収集及び精密化における統計値

表中の括弧内の数字は最外殻の値を示す。

PDB ID	Dalbavancin complex 6M5E	Apo 6M5D
<b>Data collection</b>		
Space group	C 2	<i>P</i> 4 2 <sub>1</sub> 2
Cell dimensions		
a, b, c (Å)	197.05, 125.68, 125.74	182.70, 182.70, 79.81
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90.32, 90	90, 90, 90
Resolution (Å)*	42.82-2.80 (2.90-2.80)	42.78-2.60 (2.70-2.60)
$R_{\text{meas}}$ *	0.479 (5.311)	0.078 (1.317)
$R_{\text{pim}}$ *	0.101 (1.133)	0.027 (0.451)
$\langle I/\sigma(I) \rangle$ *	9.28 (0.91)	16.92 (1.59)
$CC_{1/2}$ *	0.993 (0.456)	0.999 (0.645)
Completeness(%)*	99.9 (99.9)	91.78 (94.01)
Redundancy*	22.0 (21.7)	7.9 (8.0)
<b>Refinement</b>		
Resolution (Å)	42.82-2.80	42.78-2.60
No. unique reflections*	75476 (7487)	38650 (3891)
$R_{\text{work}} / R_{\text{free}}$	0.1799 / 0.2276	0.2308 / 0.2585
No. atoms		
Protein	13840	4470
Ligand	640	—
water	3	3
Averaged B-factors(Å <sup>2</sup> )		
Protein	63.29	115.74
Ligand	85.35	—
Water	57.04	66.92
R.m.s. deviations from ideal		
Bond lengths (Å <sup>2</sup> )	0.18	0.01
Bond angles (°)	1.49	1.27
Ramachandran plot		
Favored (%)	95.91	90.50
Allowed (%)	3.34	6.63
Outlier (%)	0.75	2.87

### 3-3 結果と考察

HSA のアポ結晶構造を 2.6 Å 分解能、共結晶で得た HSA-Dalbavancin 複合体結晶から、複合体結晶構造を 2.8 Å で決定した。複合体結晶構造から、2 分子の Dalbavancin が、HSA1 分子に結合していた。(Figure 3-2) Dalbavancin は HSA の IA サブドメインおよび III B サブドメインにそれぞれ結合していた。興味深いことに、これらの部位は、HSA の化合物結合部位としてよく知られる Sudlow' s site [33] とは異なる結合部位であった。



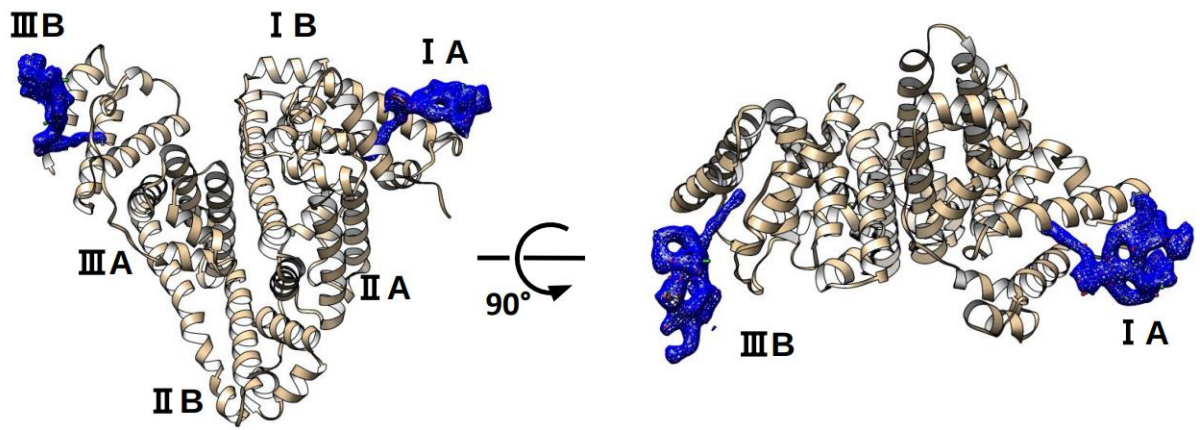


Figure 3-2 HSA-Dalbavancin 複合体構造  
各サブドメイン名を図中に示した。

一方で、ソーキングによって調製した結晶からは、複合体構造を得ることができなかった。これは、環状ペプチドの結合部位として可能性の高い場所が、結晶パッキングに寄与していたためであると考えられる。今回ソーキングに使用した環状ペプチドの多く (Caspofungin, Anidulafungin, Daptomycin, Dalbavancin) は、炭化水素鎖を有している。また、いずれも環状部分を有するため、構造の類似性が比較的高いと考えられる。これらから、今回実験に用いた、多くの環状ペプチドの結合部位は、Dalbavancin と同様の IA サブドメイン、III B サブドメインであると予想される。しかし、スクリーニングに用いたアポ結晶は、IA サブドメイン、III B サブドメインがパッキングに寄与していたため、環状ペプチドが結合不可能であったと考えられる。また、今回のスクリーニング実験で得られた結晶構造の多くは分解能が 4-5 Å であり、結晶構造から環状ペプチド由来の電子密度を確認することが難しいと予想された。

HSA-Dalbavancin 複合体構造では、IA サブドメイン、III B サブドメインの両サイトの結合様式は酷似していた。Dalbavancin の炭化水素鎖がそれぞれの結合部位中のポケットに挿入されていた (Figure 3-3)。IA サブドメインは Val17、Phe19、Val23、Phe27、Phe49、Leu66、Leu69、Phe70、および Leu251 の疎水性アミノ酸によって疎水性結合ポケットが形成されており (Figure 3-4(上段))、これによって Dalbavancin の炭化水素鎖がこのポケットに深く結合することを可能としていた。III B サブドメインについても、Phe502、Phe507、Phe509、Leu532、Phe551、Leu575 および Val576 の疎水性アミノ酸残基によって疎水性結合ポケットが形成されており (Figure 3-3(下段))、IA サブドメインと同様の様式で Dalbavancin の炭化水素鎖が深く結合することを可能としていた。Dalbavancin の結合によって引き起こされる HSA の構造変化と Dalbavancin の結合様式をさらに詳しく調べるために、アポ構造と複合体構造の重ね合わせを行うことで、両者の構造の差を解析した。Dalbavancin の環状部分は、HSA の表面を覆うことで相互作用面積を広げ ( $\approx 360 \text{ \AA}^2$ )、結合に寄与していた。特に IA サブドメインにおいては、HSA の

Ala55-Ser65 から形成されるループ領域が大きく構造変化することで、Dalbavancin の環状部分との相互作用面積を増加させていた (Figure 3-5)。アポ構造の Ala55-Ser65 は電子密度が明瞭に確認できず、このループはディスオーダーしていると考えられた。一方で、Dalbavancin 複合体のループ領域は、その電子密度が明瞭に観測された。また、当該ループ内の Glu60 が Dalbavancin の環状構造と 2 箇所の水素結合を形成し、相互作用を増加させていた。さらに、前述の疎水性アミノ酸から形成されるポケット周辺の構造についても、アポ構造と比較すると、HSA-Dalbavancin 複合体では、Dalbavancin 結合部位周辺のヘリックス (IA サブドメインのヘリックス 4 および III B サブドメインのヘリックス 3、4) がそれぞれ 2 - 5 Å 程度構造変化していた (Figure)。この構造変化によって、Dalbavancin の炭化水素鎖が HSA の化合物結合部位へのアクセスを可能にしていた。IA, III B どちらの化合物結合ポケットにおいても、HSA-Dalbavancin 複合体構造でのみ、フェニルアラニン (Phe70: IA サブドメイン、Phe551: III B サブドメイン) が大きくスイングしていた。このスイングの結果、Dalbavancin の炭化水素鎖の末端部分とこれらのフェニルアラニンが立体障害を起こすことなく Dalbavancin の HSA への結合を可能としていた。これらの構造変化による Dalbavancin の結合様式は induced-fit mechanism [34] であると考えられる。アポ構造と複合体構造の 2 つの構造を比較することで、初めて詳細な Dalbavancin の HSA への詳細な結合の議論が可能となった。

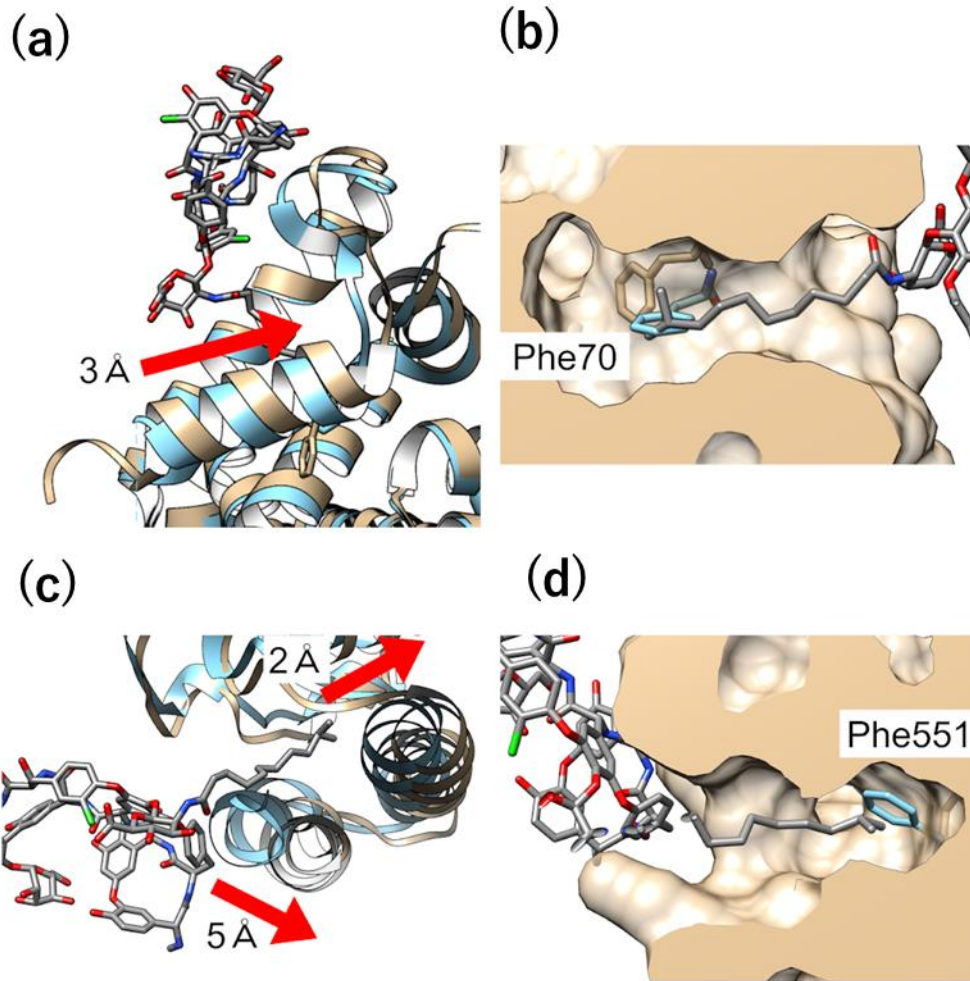
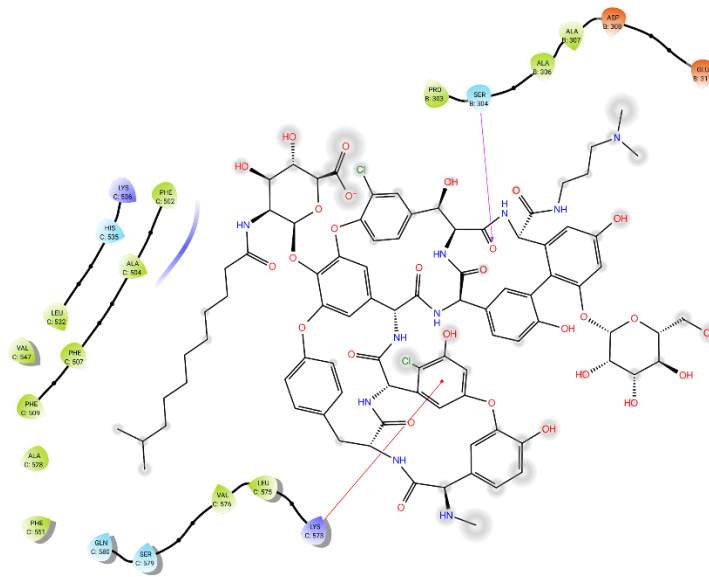
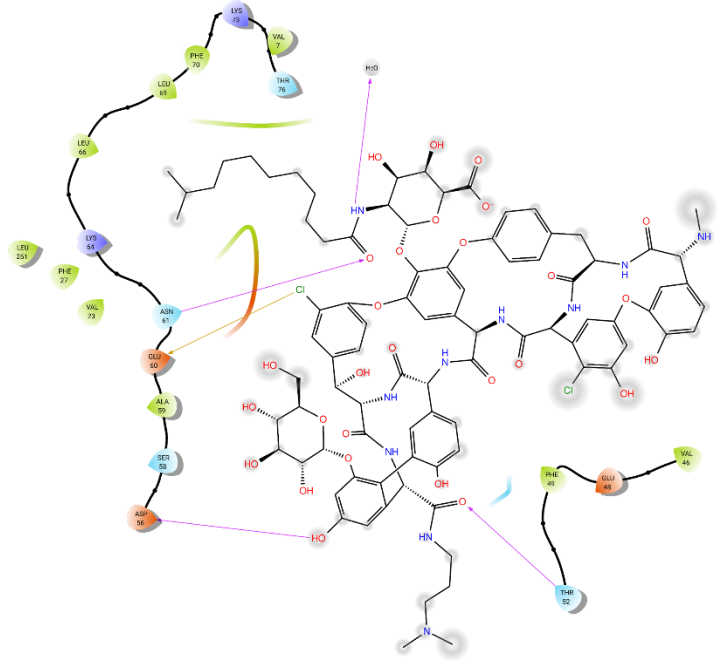


Figure 3-3 化合物結合部位のアポ構造と複合体構造の比較 リボン図で示した。  
 IAサブドメインの重ね合わせ (a, b) およびIII Bサブドメインの重ね合わせ (c, d) をそれぞれ示した。アポ構造をシアン、複合体構造をタンで示す。



- |                    |             |         |                    |                |                  |
|--------------------|-------------|---------|--------------------|----------------|------------------|
| Charged (negative) | Polar       | Glycine | H-bond (backbone)  | Pi-cation      | Salt bridge      |
| Charged (positive) | Hydrophobic | Water   | H-bond (sidechain) | Pi-Pi stacking | Solvent exposure |

Figure 3-4 IA(上段)およびIII B (下段) サブドメインにおけるHSA と Dalbavancin の相互作用様式

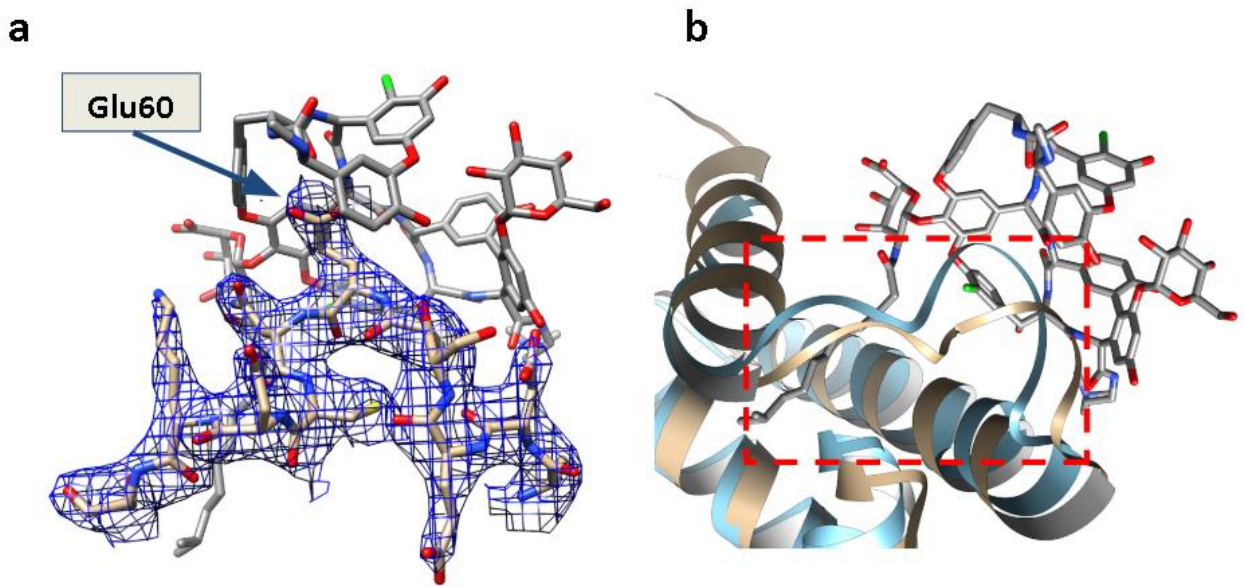


Figure 3-5 Dalbavancin 結合部位 (IA サブドメイン) におけるヘリックスの induced fit ( $1.5\sigma$   $2mFo-DFc$  マップ) (a) HAS-Dalbavancin 複合体構造 化合物結合部位近傍のループが Dalbavancin と相互作用していた。(b) アポ構造と Dalbavancin 複合体構造の重ね合わせ 複合体構造は induce fit mechanism によってループ領域が Dalbavancin の環状部分と相互作用するように大きく構造変化していた。

### 3-4 結論

本章では、DeepCentering と imaproc の応用例として、HSA アポ結晶及び HSA と種々の環状ペプチドとの複合体結晶のスクリーニングを行った。DeepCentering と imaproc を用いたスクリーニングの結果、アポ結晶においては、2.6 Å 分解能での構造決定に成功した。また、複合体結晶については、スクリーニング実験では、ソーキング法で調製した結晶では、化合物複合体構造は得られなかったものの、回折強度測定自動化により、実験の省力化が達成された。また、DeepCentering によって無人での終夜測定が可能であったため、1 度のビームタイム (12 時間) でスクリーニングを終えることができた。その後の共結晶への研究の方針転換を早い段階で決定することに貢献した。HSA アポ結晶構造と HSA-Dalbavancin 複合体結晶構造の詳細な構造比較により、Dalbavancin の結合様式を明らかにした。また、今回得られた HSA-Dalbavancin の結晶構造は、HSA と環状ペプチドの初めての結晶構造であり、この構造を基にした計算科学的研究への貢献も期待される。DeepCentering および imaproc の開発により、大型結晶を対象とする回折強度測定の高速度および省力化を達成した。

### 3-5 参考文献

- [1] Craik, D. J., Fairlie, D. P., Liras, S., Price, D. The future of peptide-based drugs. *Chem. Biol. Drug* (2013). **81**, 136-147.
- [2] Driggers, E. M., Hale, S. P., Lee, J., Terrett, N. K. The exploration of macrocycles for drug discovery—An underexploited structural class. *Nat. Rev. Drug Discov.* (2008). **7**, 608-624.
- [3] Scott, A. M., Wolchok, J. D., Old, L. J. *Nat. Rev.* (2012). **12**, 278-287.
- [4] Scott, J. K., Smith, G. P. *Science* (1990). **249**, 386-390.
- [5] Heinis, C., Winter, G. *Curr. Opin. Chem. Biol.* (2015). **26**, 89-98.
- [6] Lennard, K. R., Tavassoli, A. *Chem. Eur. J.* (2014). **20**, 10608-10614.
- [7] Travassoli, A. *Curr. Opin. Chem. Biol.* (2017). **38**, 30-35.
- [8] Passioura, T., Katoh, T., Goto, Y., Suga, H. *Annu. Rev. Biochem.* (2014). **83**, 727-752.
- [9] Suga, H. *J. Pept. Sci.* (2018). **24**, e3055.
- [10] Jongkees, S. A., Caner, S., Tysoe, C., Brayer, G. D., Withers, S. G., Suga, H. *Cell Chem. Biol.* (2017). **24**, 381-390.
- [11] Matsunaga, Y., Bashiruddin, N. K., Kitago, Y., Takagi, J., Suga, H. *Cell Chem. Biol.* (2016). **23**, 1341-1350.
- [12] Kawamura, A., Münzel, M., Kojima, T., Yapp, C., Bhushan, B., Goto, Y., Tumber, A., Katoh, T., King, O. N., Passioura, T., Walport, L. J., Hatch, S. B., Madden, S., Müller, S., Brennan, P. E., Chowdhury, R., Hopkinson, R. J., Suga, H., Schofield, C. J. *Nat. Commun.* (2017). **8**, 14773.
- [13] Passioura, T., Watashi, K., Fukano, K., Shimura, S., Saso, W., Morishita, R., Ogasawara, Y., Tanaka, Y., Mizokami, M., Sureau, C., Suga, H., Wakita T. *Cell Chem. Biol.* (2018). **25**, 906-915.



- [14] Zorzi, A., Deyle, K., Heinis, S. *Curr. Opin. Chem. Biol.* (2017). **38**, 24–29.
- [15] Maack, T., Johnson, V., Kau, S. T., Figueiredo, J., Sigulem, D. *Kidney Int.* (1979). **16**, 251–270.
- [16] Kurtzhals, P., Havelund, S., Jonassen, I., Kiehr, B., Larsen, U. D., Ribel, U., Markussen, J. *Biochem. J.* (1995). **312**, 725–731.
- [17] Koehler, M. F., Zobel, K., Beresini, M. H., Caris, L. D., Combs, D., Paasch, B. D., Lazarus, R. A. *Med. Chem. Lett.* (2002). **12**, 2883–2886.
- [18] Zobel, K., Koehler, M. F., Beresini, M. H., Caris, L. D., Combs, D. *Med. Chem. Lett.* (2003). **13**, 1513–1515.
- [19] Dumelin, C. E., Trüssel, S., Buller, F., Trachsel, E., Bootz, F., Zhang, Y., Mannocci, L., Beck, S. C., Drumea–Mirancea, M., Seeliger, M. W., Baltes, C., Müggler, T., Kranz, F., Rudin, M., Melkko, S., Scheuermann, J., Neri, D. *Angew. Chem., Int. Ed.* (2008), **47**, 3196–3201.
- [20] Malabarba, A., Goldstein, B. P. *J. Antimicrob. Chemother.* (2005). **55**, ii15–ii20.
- [21] Chen, A. Y., Zervos, M. J., Vazquez, J. A. *Int. J. Clin. Pract.* (2007). **61**, 853–863.
- [22] Bennett, J. W., Lewis, J. S., Ellis, M. W. *Ther. Clin. Risk Manage.* (2008), **4**, 31–40.
- [23] Stura, E., and I. Wilson. *J. Crystal Growth* (1991) **110** 270–282
- [24] Hassell, A. M., An, G., Bledsoe, R. K., Bynum, J. M., Carter, H. L., 3rd, Deng, S. J., Gampe, R. T., Grisard, T. E., Madauss, K. P., Nolte, R. T., Rocque, W. J., Wang, L., Weaver, K. L., Williams, S. P., Wisely, G. B., Xu, R., Shewchuk, L. M. *Acta Crystallogr Sect. D* (2007). **63**, 72–79.
- [25] Müller, I. *Acta Cryst.* (2017). **D73**, 79–92.

- [26] Ueno, G., Kanda, H., Hirose, R., Ida, K., Kumasaka, T., Yamamoto, M. *J. Struct. Funct. Genomics.* (2006), **7**, 15–22.
- [27] Sugio, S., Kashima, A., Mochizuki, S., Noda, M., Kobayashi, K. *Protein Eng.* (1999). **12**, 439–446.
- [28] Hirata, K., Ueno, G., Nisawa, A., Kawano, Y., Hikima, T., Shimizu, N., Kumasaka, T., Yumoto, H., Tanaka, T., Takahashi, S., Takeshita, K., Ohashi, H., Goto, S., Kitamura, H., Yamamoto, Y. *In AIP Conf. Proc.* (2010). 1234.
- [29] Kabsch, W. *Acta Crystallogr., Sect. D Struct. Biol.* (2010). **66**, 125–132.
- [30] McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. *J. Appl. Cryst.* (2007). **40**, 658–674.
- [31] Emsley, P., Lohkamp, B., Scott, W. G., Cowtan, K. *Acta Crystallogr., Sect. D: Struct. Biol.* (2010). **66**, 486–501.
- [32] Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., Zwart, P. H. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* (2010). **66**, 213–221.
- [33] Sudlow, G., Birkett, D. J., Wade, D. N. *Mol. Pharmacol.* (1975). **11**, 824–832.
- [34] Koshland, D. E. *Proc. Natl. Acad. Sci. U. S. A.* (1958). **44**, 98–104.

## 第4章

### マイクロ流路を用いた 化合物ソーキングシステムの開発

#### 4-1 マイクロ流路を利用したX線結晶構造解析への応用

マイクロ流路とは、フォトリソグラフィなどの微細加工技術を用いて作製された微小な流路を用いて、種々の生物学的実験や化学反応を行うデバイスの総称であり、マイクロ流路を用いた技術はマイクロフルイディクスと呼ばれる。1990年代後半から、マイクロ流路を用いた技術開発は多く行われており[1-21]、タンパク質X線結晶構造解析の分野においても、Slip Chip[1]と呼ばれる界面拡散型のデバイスや、1条件あたり20 nL程度の微小液滴を利用した結晶化デバイスなど、特に結晶化スクリーニングの効率化のためのデバイス開発が行われてきた。

これまでに述べたように、タンパク質-化合物複合体構造解析では、タンパク質-化合物複合体結晶の調製にソーキング法がよく用いられる。ソーキング法では、多くのアポ結晶を用いて、化合物濃度、pH、緩衝液の種類、温度などの様々なパラメータの条件検討が必要となる。また、一般的にタンパク質X線結晶構造解析は100 Kの低温条件で測定を行う[22]ため、抗凍結剤と呼ばれる化合物も併せて添加する必要がある。この抗凍結剤の種類、濃度についても多くの条件検討が必要となる。更に、タンパク質結晶は溶媒含有率が非常に高く[23]、空気中への長時間の露出や過度な物理的外力を与えると結晶が崩壊するおそれがあり、そのハンドリングには細心の注意が要求される。このような試料ハンドリングが難しい中で、実験者が手作業で膨大な数のソーキング実験を行うことは大きな負担となる。FBDD等のソーキング法を用いた複合体結晶構造解析では、この複合体結晶調製がボトルネックとなっており、ソーキング実験の簡便化、省力化が求められている。

近年、常温結晶構造解析[24]も注目され始めており、その試料ハンドリング法の一つとしてマイクロ流路の利用が検討されている。低温結晶構造解析が多くのタンパク質の機能理解に大きく貢献してきた一方で[25-35]、同一種類のタンパク質を低温、常温で構造解析を行い、構造比較した結果から、常温構造でのみ、①アミノ酸残基の alternate conformation が低温構造よりも多く存在すること ②化合物結合

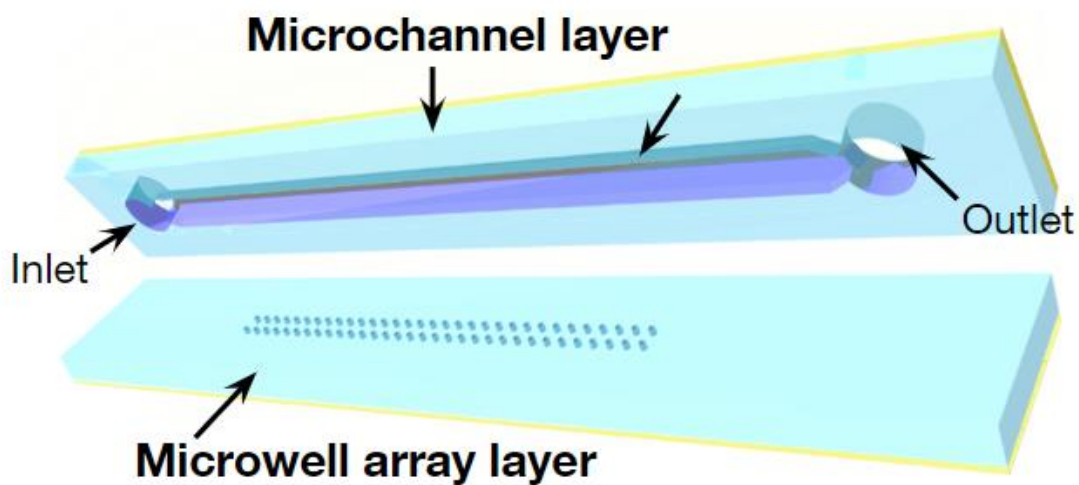
部位での conformation が低温構造と比較して異なっている場合があること ③アロステリック部位が維持されるなど[33]、低温結晶との差が報告されている。これらの常温構造と低温構造の差は、結晶凍結によるアーティファクトが一因であると考えられており、XFEL 等の第 4 世代光源を利用した常温無損傷結晶構造解析が現在活発に行われている。一方で、XFEL での実験は利用可能な実験施設が非常に限られているため、日常的な利用が難しいという問題がある。常温結晶構造解析の重要性が認識されてきている中で、利用機会の多さを勘案すれば、リング型放射光施設での常温結晶構造解析の汎用化が期待される。しかし、非凍結状態のタンパク質結晶を安定に保ちつつ、かつ、放射線損傷を低減した X 線回折強度測定が、リング型放射光施設で常温構造解析を行う上でのボトルネックとなっている。一般的な常温結晶の X 線回折強度測定は、MicroRT[36]などの、結晶を保持しつつ、測定中の乾燥を防ぐための、特殊な治具を用いる必要がある。また、このような治具を用いた実験は、凍結結晶の回折強度測定と比較して、より熟練度が要求される。さらに、常温構造解析においては、放射線損傷が大きな問題となる。一般的な常温構造解析では、重篤な放射線損傷のため 1 つの結晶から完全なデータセットを収集することは難しく、複数の結晶の回折データをマージすることで構造解析を行う場合が多い。このように、複数の結晶を、常温測定という熟練度が要求される実験を繰り返し行い、多大な労力と時間を要することが常温結晶構造解析を行う上でボトルネックとなっている。

このような背景を踏まえ、本研究ではソーキング実験の簡便化、省力化、ならびに、常温構造解析の汎用化を目標とし、Figure 4-1 に示すような、マイクロ流路を用いた結晶捕捉アレイを開発した。開発した X 線回折実験用結晶保持アレイは、流路内にピットと呼ばれる穴を作製し、そのピットに結晶を捕捉する機構を採用した。一度ピットに捕捉された結晶は、流路に溶液を追加送液しても位置は固定されたままである。捕捉された結晶に対して、化合物溶液を結晶捕捉アレイに導入することでソーキング実験が完了する。本研究で開発した結晶捕捉アレイを使用するこ

とにより、上述のように手作業で行っていたソーキング実験を簡単なピペッティング作業のみで行うことが可能となった。更に、結晶捕捉アレイをそのまま回折計に取り付け *in situ* で測定することで、常温構造解析を、従来よりも簡易的に実施することを可能にした。

本研究では、開発した結晶捕捉アレイを用いて、① リゾチームとソーマチンを用いた結晶捕捉の検証実験 ② リゾチームとトシル酸を用いた化合物ソーキング実験と常温X線回折強度測定 ③ トリプシンと種々の化合物を用いた複合体結晶スクリーニングを行った。

(a)



(b)



Figure 4-1 本研究で用いた結晶捕捉アレイの例 (a) 結晶捕捉アレイの概念図および (b) 接合後の断面図

本研究で用いた結晶捕捉アレイは、マイクロチャネルレイヤーとマイクロウェルアレイレイヤーの2つのレイヤーから構成される。あらかじめこれら2種類のレイヤーを作製し、プラズマ処理で接合して結晶捕捉およびソーキング実験に使用した。

## 4-2. 材料と方法

### 4-2-1. 結晶捕捉のための流路の設計および固定治具の作製

前節で述べた通り、本研究では結晶捕捉アレイに結晶を捕捉することで、ソーキング実験を行った。本研究で用いた結晶捕捉アレイは、マイクロチャネル層とマイクロウェルアレイ層の2層で構成される。本研究で用いた結晶捕捉アレイの作製工程をFigure 4-2に示す。はじめにマイクロチャネル層とマイクロウェルアレイ層を液状シリコンゴム (PDMS) で製作する際の鋳型をエポキシ樹脂ベースのネガフォトレジスト SU-8 3050 (以下フォトレジスト) を材料とするフォトリソグラフィーで製作した。マイクロウェルアレイ層の鋳型を作成するためのフォトマスクは矩形ピットの開口サイズに相当する縦横 400 x 100  $\mu\text{m}$  の紫外線透過窓を 100  $\mu\text{m}$  間隔で 40 個持つ。これらのフォトマスクは共同研究先の北海道大学渡慶次研究室にて作成した。鋳型の作成では、フォトレジストをシリコンウェハー上に必要な厚みとなるように注ぎ (マイクロチャネル層 : 280  $\mu\text{m}$ , マイクロウェルアレイ層 : 250  $\mu\text{m}$ ) フォトマスクを装着した後、マスクアライナー (M-1 S、三笠商事) を用いて紫外線照射を行い、フォトレジストを硬化した。照射後のフォトレジストをシリコンウェハーごとホットプレート上で焼成することでフォトレジストを架橋し、次いで SU-8 現像液に浸漬し未架橋の SU-8 を溶解除去してマイクロチャネル層とマイクロウェルアレイ層の鋳型を得た。

マイクロチャネル層とマイクロウェルアレイ層の製作では、1H, 1H, 2H, 2H-perfluorooctyl 蒸気で処理した上述の鋳型に流し込んだ PDMS を 80 度に加熱固化しマイクロチャネル層とマイクロウェルアレイ層を得た。得られたそれぞれの層をカッターで切り出し、それぞれプラズマ処理を行い表面の親水化を行った。親水処理後のマイクロウェル層とマイクロウェルアレイ層を接着し、これを結晶捕捉及び化合物ソーキング実験に用いた。結晶捕捉を行う直前に、結晶捕捉アレイをデシケータに静置した状態でダイヤフラムポンプを用いてデシケータ内を減圧することで結晶捕捉ア



レイを3分間脱気し、これを実験に用いた。

本研究では、結晶捕捉アレイをそのまま回折計に取り付け、*in situ*測定を行う。一般的な結晶の取り付け治具では結晶捕捉アレイの取り付けが困難であることから、Figure 4-3に示す取り付け治具を作製した。

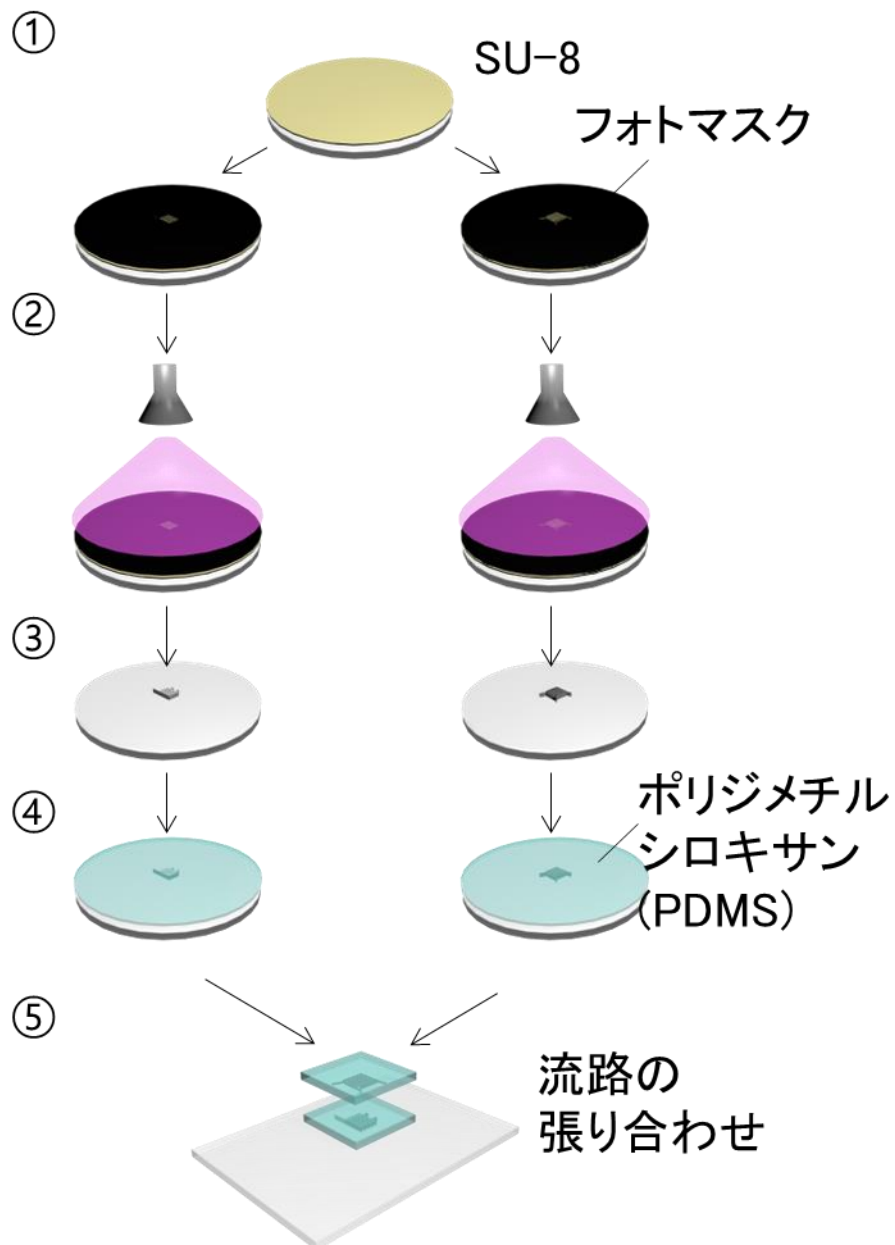


Figure 4-2 結晶捕捉アレイの作製工程

以下の手順で結晶捕捉アレイを作製した。

- ① フォトレジスト (SU-8) を Si ウェハ上にスピコートし厚みを調整する。
- ② UV で露光し硬化させる。
- ③ SU-8 ディベロッパー中で振盪して現像し鋳型の作製が完了となる。
- ④ PDMS を③の鋳型に流し込み、80°Cで硬化させる。
- ⑤ 鋳型から剥型し、二種類の流路をアライメント後にプラズマ接着する。

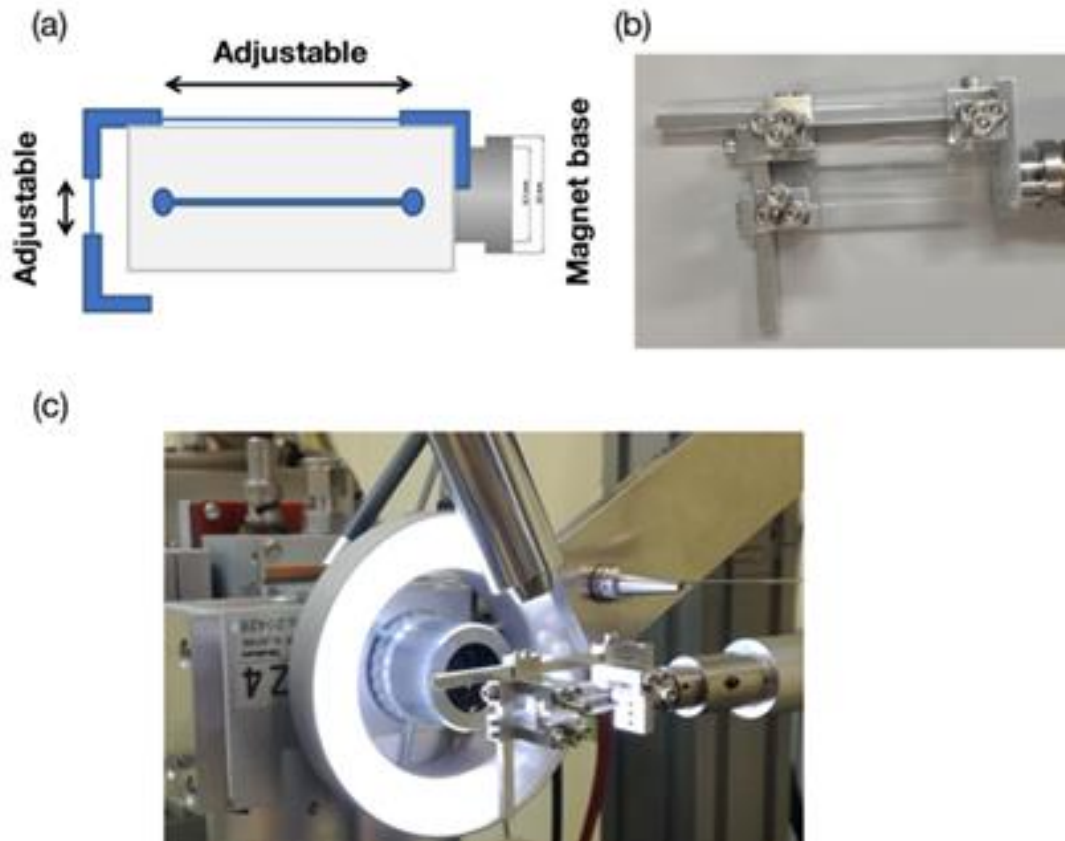


Figure 4-3 結晶捕捉アレイを回折計へ取り付けるための治具

(a) 模式図 (b) 結晶捕捉アレイを取り付けた状態 (c) 回折計に取り付けた状態

#### 4-2-2 ニワトリ卵白由来リゾチームとタウマトコックスダニエリ由来ソーマチンを用いた検証実験

作製した結晶捕捉アレイを用いて、結晶捕捉、化合物ソーキング及びX線回折強度測定が問題なく行えるか検証実験を行った。テスト試料として、ニワトリ卵白由来リゾチーム及びタウマトコックスダニエリ由来ソーマチンを用いた。リゾチームはナカライテスク、ソーマチンは和光純薬工業からそれぞれ購入し、購入品をそのまま結晶化実験に使用した。リゾチームを 50 mg/mL になるように 50 mM 酢酸緩衝液 (pH 4.6) に溶解した。結晶化にはバッチ法を用いた。100  $\mu$ L リゾチーム溶液と 100  $\mu$ L の 1.0 M 塩化ナトリウムをマイクロチューブ内で混合し、20°C で静置した。結晶化から数日後に 100 - 300  $\mu$ m の結晶が析出した。化合物溶液には、リゾチームの結晶化溶液の組成にトシル酸の終濃度が 500 mM になるよう調製した。

ソーマチン結晶については、以下の手順で試料調製を行った。ソーマチンを 0.1 M N-(2-acetamido) iminodiacetic acid (ADA) に溶解し、溶液を 15000 x g で遠心した。上清を回収し、ソーマチン濃度が 40 mg/mL となるように ADA 緩衝液で希釈した。結晶化はバッチ法を用いて行った。沈殿剤溶液には、0.1 M ADA 緩衝液 (pH 6.5)、1.6 M 酒石酸カリウムナトリウム (和光純薬工業)、50 mM HEPES (pH 7.0) を用いた。100  $\mu$ L のタンパク質溶液と 100  $\mu$ L 沈殿剤溶液をマイクロチューブで混合し、20°C で静置した。結晶化から 24 時間後に約 200  $\mu$ m の結晶が析出した。

得られたリゾチームとソーマチンの結晶を用いて結晶捕捉を行った。実験手順の概要を Figure 4-4 に示す。実験手順は以下の手順で行った。初めに、結晶捕捉アレイ (PDMS) 中に含まれる空気を取り除くため、デシケータ内に結晶捕捉アレイを静置し、ダイヤフラムポンプで 3 分間の脱気を行う。その後、4  $\mu$ L のリザーバー溶液を結晶捕捉アレイに導入し、次いで、4  $\mu$ L の結晶懸濁液を導入し 2 分静置する。流路中に残った余剰の結晶を除去するため、8  $\mu$ L のリザーバー溶液を導入する。最後に、4  $\mu$ L の化合物溶液を導入しソーキング実験が完了する。いずれのタンパク質に

についても、操作は手作業のピペッティングにより行った。リゾチーム結晶については、トシル酸溶液を 4  $\mu\text{L}$  結晶捕捉アレイに導入し、ソーキングを行った。

リゾチーム結晶、ソーマチン結晶のいずれについても、結晶捕捉アレイ中に捕捉された結晶方位の確認のために、データ処理プログラムである XDS の指数付け結果から得られる結晶方位を示す UB 行列を使用し、指数付けされた結晶方位の  $a^*$  軸および  $c^*$  軸の方向ベクトルを極座標表示することで結晶方位を解析した。また、X線回折強度測定は以下の条件で行った。複合体結晶 1 個あたりの測定範囲を  $20^\circ$  に設定し、複数の結晶から得た回折データをマージすることで構造解析を行った。各結晶の露光点の決定は BSS から行った。露光時間は 0.5 秒、1 イメージ当たりの振動角は  $0.5^\circ$  で測定を行った。常温測定では、およそ 50 - 250 kGy 程度で結晶の回折能が半減する [40]。この時の線量を  $D_{1/2}$  と定義すると、本研究では 1 結晶あたりの線量が  $D_{1/2}$  に近い値になるよう露光時間を設定した。本測定における結晶 1 個当たりの Dose は約 130 kGy であった。得られた回折の処理及びデータマージは、自動処理データ処理プログラム KAMO [41] を用いた。Phaser [42] を用いて分子置換法で位相決定を行った。サーチモデルには PDB ID: 1DPX (リゾチーム) および 1RQW (ソーマチン) を用いた。その後、Coot [43] を用いて手動でモデル改良を行い、phenix.refine [44] により精密化を行った。

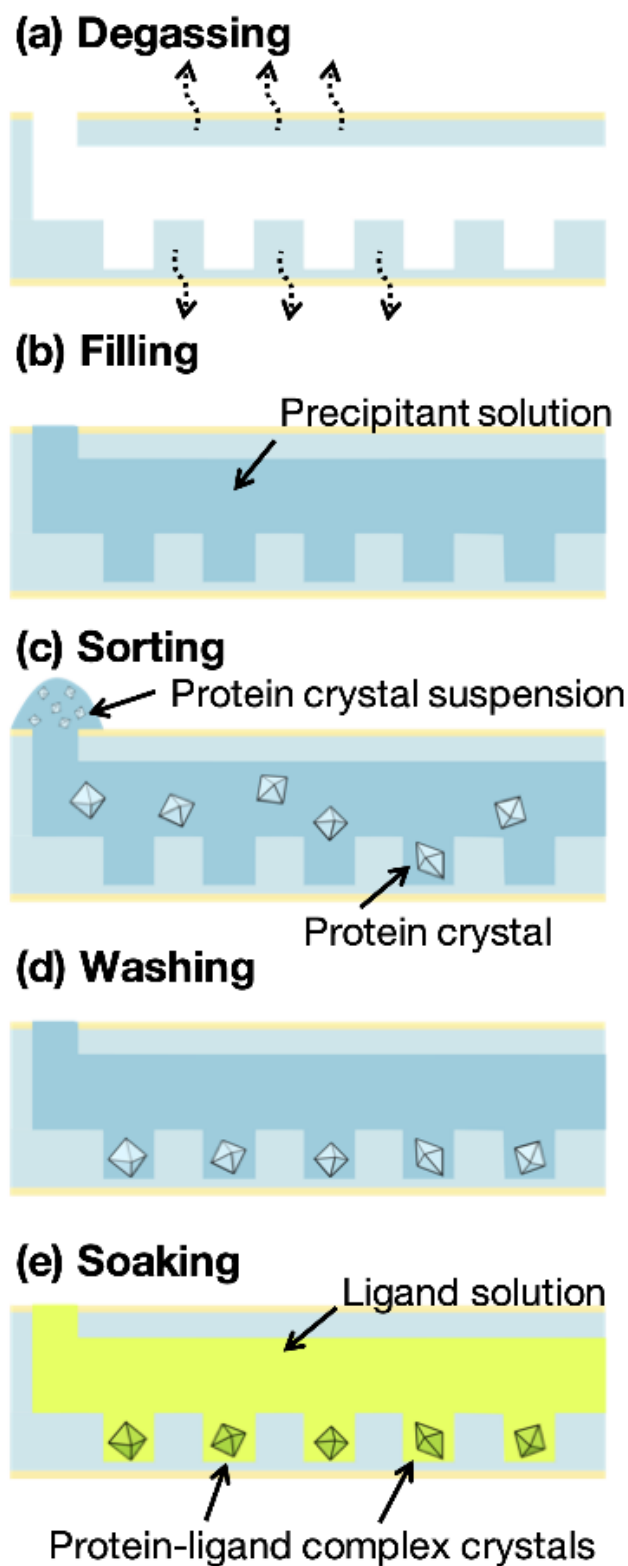


Figure 4-4 結晶捕捉アレイによる結晶補足および化合物ソーキング

(a)3 分間の脱気 (b) リザーバー溶液の送液 (c) 結晶懸濁液の導入 (d) リザーバー溶液の送液による流路内の洗浄 (e) 化合物溶液の導入によるソーキング

#### 4-2-3. ウシ膵臓由来トリプシンと種々の化合物を用いた複合体スクリーニング

ウシ膵臓由来トリプシンは東京化成工業から購入し、購入品をそのまま結晶化に使用した。トリプシン濃度が 30 mg/mL になるように 25 mM HEPES (pH 7.0) 及び 5 mM 塩化カルシウムに溶解した。結晶化はシッティングドロップによる蒸気拡散法を用いて、4°Cで行った。リザーバー溶液の組成は、30 % (v/v) polyethylene glycol 3350、0.2 M Lithium sulfate、0.1 M Tris-HCl (pH 8.5)であった。結晶化から約 1 週間後に 200  $\mu$ m 程度の結晶が析出した。化合物ソーキングには、aniline, serotonin, 2-methyltryptamine, 4-bromobenzamide, 5-chlorotryptamine, 5-methoxytryptamine, 6-methoxytryptamine, benzamide, 及び melatonin の 9 種類を用いた (Figure 4-5)。aniline、4-bromobenzamide、5-methoxytryptamine 及び 6-methoxytryptamine については、それぞれの化合物終濃度が 50 mM になるようリザーバー溶液と 10-20 % Dimethyl sulfoxide を用いて化合物溶液を調製した。serotonin, 2-methyltryptamine, 4-bromobenzamide, 5-chlorotryptamine、benzamide, 及び melatonin についてはリザーバー溶液を用いてそれぞれの化合物濃度が 50 mM になるように化合物溶液を調製した。

結晶化プレートのドロップから、ピペッティングにより 4  $\mu$ L の結晶懸濁液を結晶捕捉アレイデバイスのインレットに添加した。その後、結晶捕捉アレイデバイスのアウトレットからピペッティングを行うことで、結晶懸濁液を流路内に導入した。結晶懸濁液を導入後、約 2 分静置することで結晶を流路内のピットに捕捉した。結晶捕捉後の結晶捕捉アレイに化合物溶液を 4  $\mu$ L 導入し、ソーキングを行った。X線回折強度測定は SPring-8 BL26B2[37]で行い、装置制御には SPring-8 タンパク質ビームラインの標準制御プログラムである BSS[38]を用いた。結晶捕捉アレイデバイスを回折計にそのまま取り付け、*in situ*測定を行った (Figure 4-5)。結晶捕捉アレイを用いた *in situ*回折強度実験は常温で行うことから、結晶の放射線損傷による影響が大きい。従って、本研究では、化合物複合体結晶 1 個あたりの測定範囲を 20° に設定し、

複数の結晶から得た回折データをマージすることで構造解析を行った。各結晶の露光点の決定は BSS から行った。露光時間は 0.5 秒、1 イメージ当たりの振動角は  $0.5^\circ$  で測定を行った。結晶への X 線照射による Dose の算出は、RADDPOSE[39] を用いた。結晶 1 個当たりの Dose は約 130 kGy であった。得られた回折データの処理及びデータマージは、自動処理プログラム KAMO を用いた。Phaser を用いて分子置換法で位相決定を行った。サーチモデルには PDB ID: 1SQ0 を用いた。その後、Coot を用いて手動でモデル改良を行い、phenix.refine により精密化を行った。melatonin 以外の場合において、精密化後、化合物由来の電子密度が明瞭に観測された。Coot を用いて電子密度に各化合物をフィッティングし、再度 phenix.refine を用いて構造精密化を行った。データ処理および精密化の統計値を Table 4-1 に示す。

また、構造比較のため、同様の化合物複合体の構造解析を、凍結結晶を用いて行った。



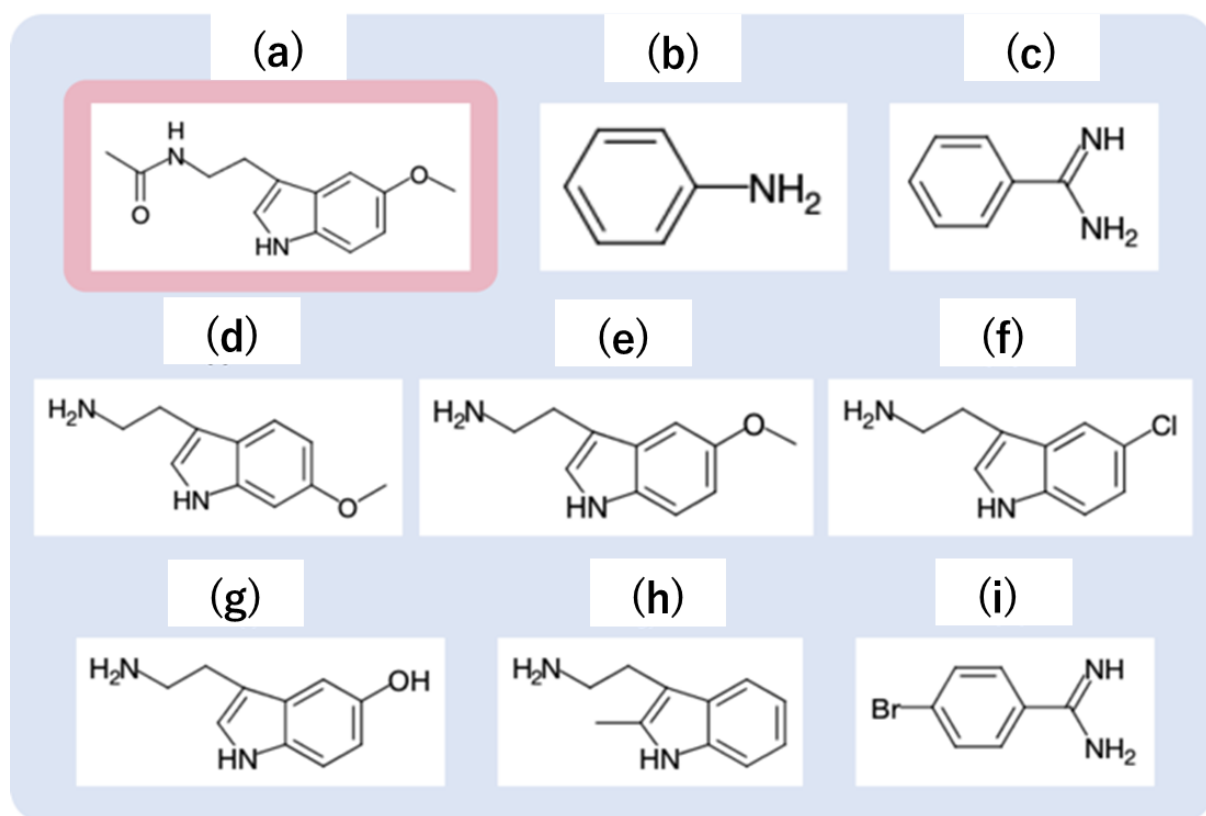


Figure 4-5 化合物スクリーニングに用いた化合物

スクリーニングに使用した化合物は(a)melatonin、(b)aniline、(c)benzamidine、(d) 5-methoxytryptamine 、 (e)6-methoxytryptamine 、 (f)5-chlorotryptamine 、 (g)serotonin、(h)2-methyltryptamine、(i)4-bromobenzamidine を用いた。すでにトリプシンの阻害剤として知られる benzamidine の構造と類似性のある化合物を選択した。図中(a)の melatonin 以外で化合物由来電子密度が観察された。

## 4-3 結果と考察

### 4-3-1 リゾチームとソーマチンを用いた検証実験

結晶捕捉アレイを用いた結晶捕捉の効率を検証するため、結晶の捕捉率を以下のよう

に定義した。

$$\text{捕捉率} = \text{捕捉された結晶} / \text{ピット数}$$

リゾチームとソーマチンの捕捉率は、それぞれ 80 % (32 結晶/40 ピット) および 90 % (36 結晶/40 ピット) であった。データマージに使用した結晶数と X 線回折強度測定

の結晶学的統計値 (completeness、redundancy、 $CC_{1/2}$ ) の関係を Figure 4-6 に示す。

測定した結晶のうち、リゾチームでは 28 結晶、ソーマチンでは 32 結晶の指数付けが自動回折イメージ処理プログラム KAMO によって自動で行われた。リゾチーム、ソーマチンのいずれの場合についても、マージした結晶数が 3 個以下であってもデータの completeness が 90% 以上であった。これは、少なくとも今回用いた結晶の対称性 (リゾチーム、ソーマチンのいずれも点群 422、多重度 8) の場合においては、completeness は構造解析を行うために十分な値であるといえる。また、分解能については、リゾチーム、ソーマチンでそれぞれ 2.0 Å、1.7 Å であった (正しく指数付けされた全結晶をマージした場合の値)。また、結晶捕捉アレイ中の結晶の配向についても、結晶方位が局在することはなかった (Figure 4-7)。このことから、今回の検証実験に用いた結晶よりも対称性の低い空間群においても、完全性の高いデータ収集が可能であると期待される。

今回行った検証実験から、本研究で開発した結晶捕捉アレイを用いた *in situ* 測定で結晶構造解析が可能であることが示唆された。

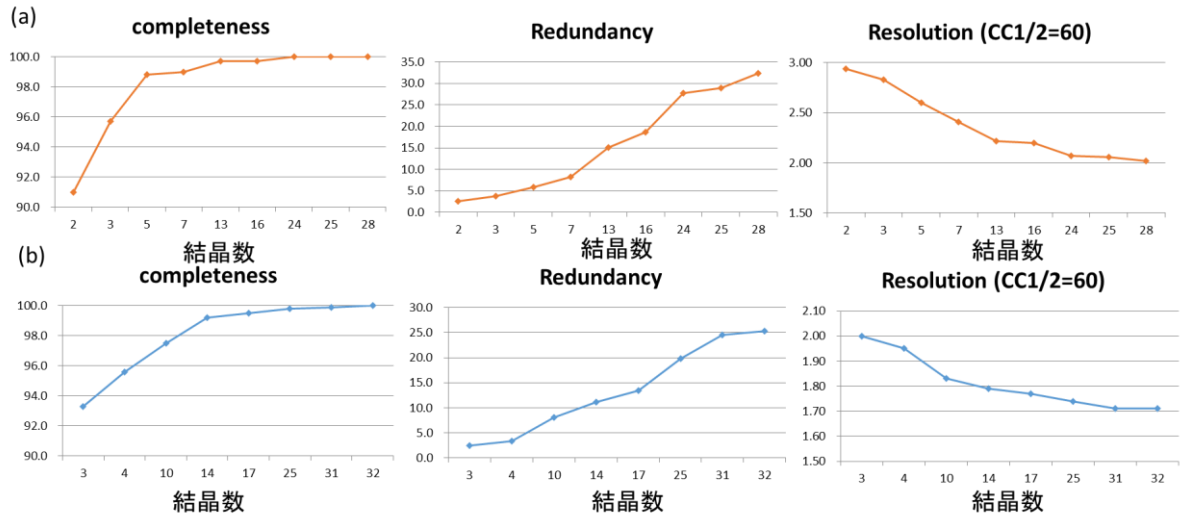
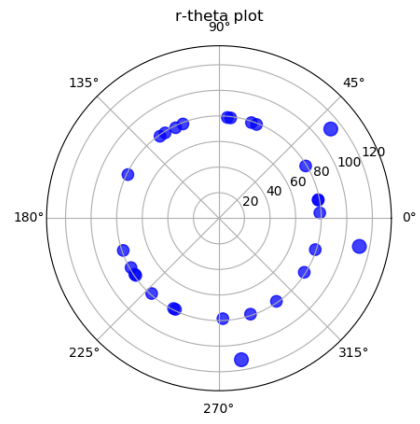
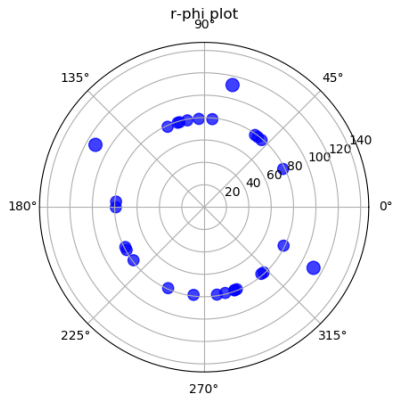
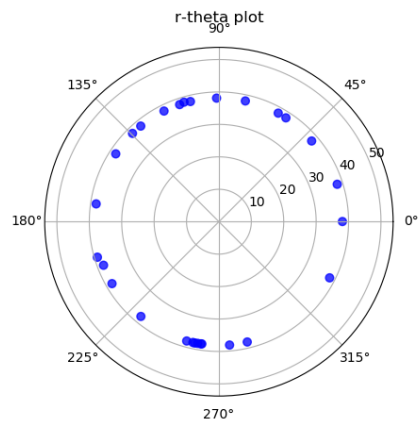
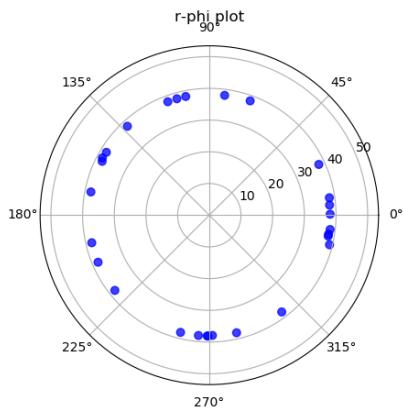


Figure 4-6 (a) リゾチンおよび(b) ソーマチンの X線回折強度測定における結晶学的統計値 データマージにおける completeness、Redundancy、 $CC_{1/2}$ を示す。

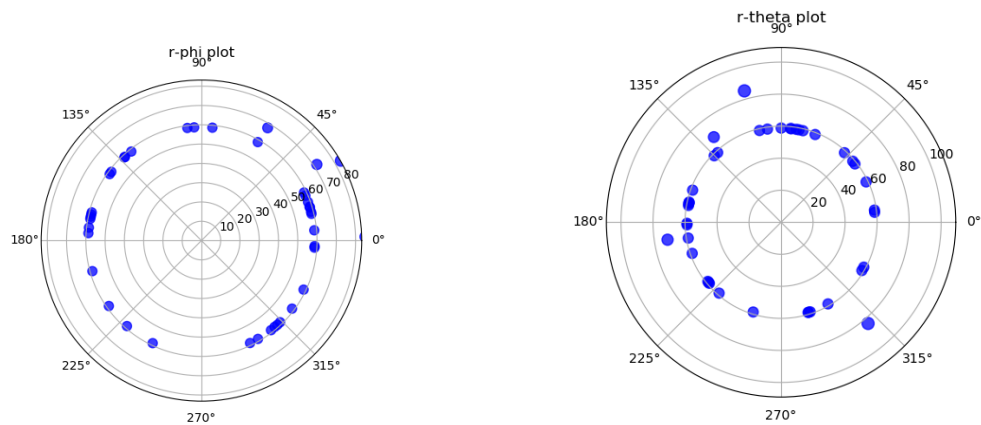
(a-1)



(a-2)



(b-1)



(b-2)

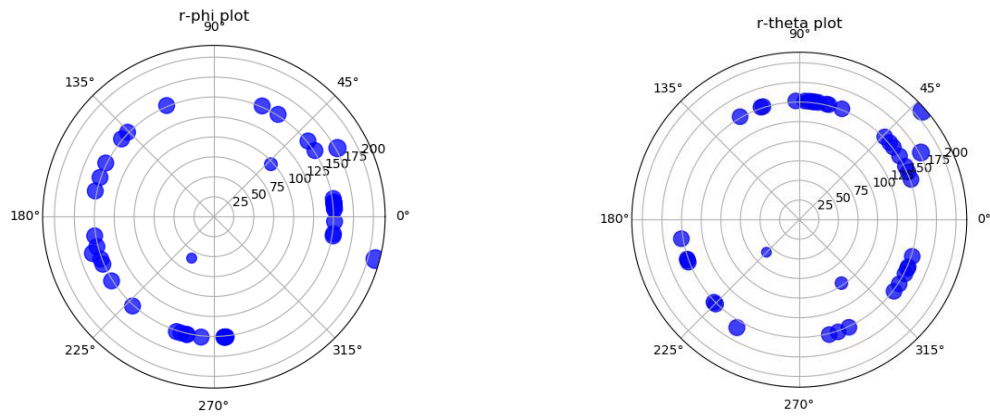


Figure 4-7 リゾチーム結晶(a)とソーマチン結晶(b)の結晶捕捉アレイ中での配向 XDS の指数付けで得られた UB 行列の a\*軸 (a-1 および b-1) および c\*軸(a-2 および b-2) の方向ベクトルの極座標プロットを示す。

青色の点が 1 つの結晶の配向を表す。リゾチーム、ソーマチンのいずれの場合においても、その配向が極端に局在することはなく、回折強度データの完全性も 90%以上と構造解析を行う上で十分な統計値を示した。

#### 4-3-2. トリプシンと種々の化合物を用いた複合体スクリーニング

それぞれのデータ測定の結晶学的統計値を Table 4-1 に示した。化合物由来の電子密度が確認されたすべての結晶構造について、分解能が 2 Å 以上の高分解能での構造解析に成功した。一方、低温構造の場合と分解能の比較を行うと、いずれの化合物複合体構造についても、低温構造の方が高分解能での構造解析が可能であった。結晶捕捉アレイを用いてトリプシンと 9 種類の化合物に対して、化合物スクリーニングを行った結果、melatonin を除く 8 種類の化合物の化合物結合部位に、化合物由来の明瞭な電子密度が観測された (Figure 4-9)。

Table 4-1 結晶捕捉アレイを用いた常温構造（上段）と低温構造（下段）の結晶学的統計値 表中の括弧内の数字は最外殻の値を示す。

	Room temperature							
	Aniline	Serotonin	2-methyltryptamine	4-bromobenzamide	5-chlorotryptamine	5-methoxytryptamine	6-methoxytryptamine	benzamide
PDB ID	7BS0	7BS2	7BRZ	7BRV	7BRW	7BRX	7BRY	7BS1
<b>Data collection</b>								
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions								
a, b, c (Å)	55.18, 59.31, 67.69	55.09, 58.52, 67.58	54.92, 58.59, 67.64	54.98, 58.77, 68.03	55.22, 58.92, 67.75	54.90, 58.60, 67.44	55.33, 57.21, 67.42	54.45, 57.88, 66.69
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Resolution (Å)	44.61-1.85 (1.92-1.85)	44.24-1.80 (1.86-1.80)	33.82-1.30 (1.35-1.30)	40.15-1.55 (1.61-1.55)	44.46-1.6 (1.66-1.6)	34.44-1.35 (1.40-1.35)	39.77-1.65 (1.71-1.65)	34.09-1.50 (1.55-1.50)
R <sub>int</sub>	0.296 (1.562)	0.207 (0.794)	0.147 (3.344)	0.185 (2.268)	0.289 (1.919)	0.154 (4.236)	0.152 (1.387)	0.128 (2.916)
<I/σ(I)>	4.52 (0.94)	7.21 (2.26)	15.47 (2.05)	10.13 (1.71)	4.91 (1.00)	21.42 (3.28)	9.42 (1.33)	15.74 (0.93)
CC <sub>1/2</sub>	0.968 (0.316)	0.979 (0.683)	0.999 (0.406)	0.998 (0.590)	0.967 (0.337)	0.999 (0.463)	0.994 (0.487)	0.995 (0.302)
Completeness(%)	94.55 (95.05)	92.61 (91.34)	99.67 (99.59)	99.10 (99.15)	96.01 (95.06)	99.02 (99.62)	95.60 (97.30)	92.00 (84.09)
Redundancy	3.6 (3.4)	3.7 (3.7)	9.3 (8.6)	10.3 (10.0)	5.4 (5.3)	12.8 (12.0)	6.6 (6.5)	15.7 (13.8)
Wilson B-factor	15.99	11.68	11.31	13.99	13.53	12.66	15.41	15.66
<b>Refinement</b>								
No. unique reflections	18476 (1804)	19327 (1877)	54271 (5306)	32276 (3162)	28648 (2791)	48000 (4761)	25238 (2523)	31682 (2864)
R <sub>int</sub> / R <sub>σ</sub>	0.1769 / 0.2107	0.1491 / 0.1936	0.1621 / 0.1837	0.1645 / 0.1759	0.1728 / 0.2143	0.1630 / 0.1815	0.1598 / 0.1961	0.1688 / 0.1898
No. atoms								
Protein	1682	1674	1701	1711	1746	1759	1675	1654
Ligand	8	19	19	25	14	15	20	10
water	202	222	205	140	208	197	174	199
Averaged B-factors(Å <sup>2</sup> )								
Protein	18.52	13.82	14.37	16.99	16.64	15.85	17.84	16.52
Ligand	24.45	25.23	20.57	32.83	18.34	25.31	28.31	16.29
Water	31.26	30.07	29.34	29.39	31.71	30.32	31.41	29.65
R.m.s. deviations from ideal								
Bond lengths (Å)	0.007	0.01	0.005	0.006	0.007	0.01	0.005	0.01
Bond angles (°)	0.79	0.77	0.78	0.82	0.8	1.3	0.79	0.84
Ramachandran plot								
Favored (%)	98.64	97.74	98.64	98.19	97.74	97.74	98.19	98.19
Allowed (%)	1.36	2.26	1.36	1.81	2.26	2.26	1.81	1.81
Outlier (%)	0	0	0	0	0	0	0	0
Ligand occupancy	0.64	0.86	0.9	0.69	0.81	0.67	0.84	0.93

	Cryogenic-conditions							
	Aniline	Serotonin	2-methyltryptamine	4-bromobenzamide	5-chlorotryptamine	5-methoxytryptamine	6-methoxytryptamine	benzamide
PDB ID	7BS7	7BS9	7BS6	7BS3	7BSA	7BS4	7BS5	7BS8
<b>Data collection</b>								
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions								
a, b, c (Å)	54.71, 58.57, 66.83	54.74, 58.53, 66.62	54.61, 58.56, 66.86	54.24, 56.90, 66.67	54.38, 58.57, 66.66	54.56, 58.67, 66.50	54.84, 56.90, 66.57	54.11, 57.07, 66.17
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Resolution (Å)	33.42-1.04 (1.08-1.04)	39.98-1.05 (1.09-1.05)	33.43-1.04 (1.08-1.04)	42.28-1.28 (1.33-1.28)	39.85-1.12 (1.16-1.12)	26.84-1.04 (1.08-1.04)	39.49-1.17 (1.21-1.17)	33.77-1.37 (1.43-1.37)
R <sub>int</sub>	0.063 (0.667)	0.053 (0.549)	0.039 (0.144)	0.127 (1.664)	0.092 (1.128)	0.032 (0.477)	0.081 (1.391)	0.040 (1.026)
<I/σ(I)>	24.90 (1.81)	28.55 (2.49)	48.71 (8.12)	12.06 (1.26)	21.24 (1.86)	43.93 (2.70)	21.31 (1.65)	42.91 (2.90)
CC <sub>1/2</sub>	0.999 (0.738)	0.999 (0.821)	0.999 (0.984)	0.999 (0.476)	1.000 (0.614)	1.000 (0.835)	1.000 (0.627)	1.000 (0.906)
Completeness(%)	88.20 (28.65)	87.46 (34.66)	87.55 (26.88)	99.91 (99.23)	92.70 (60.46)	88.00 (26.66)	99.25 (92.70)	98.99 (96.96)
Redundancy	12.1 (3.8)	12.5 (4.5)	12.1 (3.7)	7.6 (7.4)	21.24 (1.86)	11.9 (3.7)	13.8 (9.8)	14.4 (14.4)
Wilson B-factor	8.41	9.36	6.34	10.54	8.63	8.64	10.65	16.22
<b>Refinement</b>								
No. unique reflections	90800 (2916)	87417 (3404)	90512 (2748)	102008 (10148)	76437 (4930)	90569 (2714)	70079 (6479)	42915 (4142)
R <sub>int</sub> / R <sub>σ</sub>	0.1583 / 0.1690	0.1540 / 0.1703	0.1518 / 0.1619	0.1747 / 0.1907	0.1575 / 0.1752	0.1540 / 0.1718	0.1668 / 0.1807	0.1616 / 0.1781
No. atoms								
Protein	1673	1669	1658	1681	1656	1703	1681	1696
Ligand	34	35	26	59	57	77	27	23
water	383	442	394	242	331	416	360	307
Averaged B-factors(Å <sup>2</sup> )								
Protein	11.96	10.5	7.5	12.53	10.63	12.99	12.66	16.89
Ligand	9.82	14.96	10.45	23.14	14.7	10.32	15.68	22.08
Water	17.82	22.09	18.56	25.39	22.91	16.78	25.8	30.12
R.m.s. deviations from ideal								
Bond lengths (Å)	0.004	0.004	0.005	0.01	0.005	0.01	0.005	0.01
Bond angles (°)	0.84	0.83	0.88	0.88	0.85	0.86	0.82	0.85
Ramachandran plot								
Favored (%)	98.64	98.64	97.74	98.19	98.64	98.19	97.74	98.64
Allowed (%)	1.36	1.36	2.26	1.81	1.36	1.81	2.26	1.36
Outlier (%)	0	0	0	0	0	0	0	0
Ligand occupancy	0.61	0.78	0.88	0.7	0.86	0.7	0.85	0.86

メラトニンの電子密度が観測されなかった理由として、N-アセチル基の体積が、トリプシンの化合物結合部位に対して大きいこと、荷電状態が他の 8 種の化合物と異なっており、他の化合物で観測された近傍の Asp189 と水素結合を形成することが不可能であったことなどが挙げられる。aniline を除いた、化合物由来の電子密度が確認された化合物複合体構造において、電子密度から、その官能基の違いまでを明瞭に確認することが可能であった (Figure 4-9)。aniline では、他の化合物とトリプシンの Asp189 で確認された水素結合を形成せず、結合部位のポケットにファンデルワールス力などの比較的弱い力によってトリプシンに結合していたことから、aniline 由来の電子密度が他の 7 化合物と比較して不明瞭であったと考えられる。同様の相互作用様式が aniline の低温構造においても確認された。一方で、他の 7 化合物についての電子密度は非常に明瞭に確認された。このことは、FBDD 等で用いられるような分子量の小さい化合物スクリーニングにおいても、結晶捕捉アレイを用いた構造解析が貢献できることを示している。また、凍結結晶構造と常温結晶構造の比較を行った結果、凍結結晶でのみ、化合物 (benzamidine) 由来の電子密度が、化合物結合部位以外でも観測された (Figure 4-10)。この電子密度は結晶中のタンパク質間に挟まれるような形で存在していた。Benzamidine 複合体の凍結構造の格子定数は、常温構造の格子乗数と比較して、a、b、c 軸がそれぞれ 0.62 %、1.4 %、0.8 %縮小していた。結晶凍結の際に結晶格子が縮小することで、化合物の結合が可能になったと考えられる。一方、常温構造では、このような extra site への化合物の結合は観測されなかった。トリプシンの化合物結合部位は 1 か所のみであること、今回観測された extra site への benzamidine の結合は結晶のパッキング間で観測されたことから、今回観測された extra site での benzamidine は結晶を凍結することで生じるアーティファクトであると考えられる。低温構造と常温構造との比較により、このような結晶凍結によるアーティファクトを回避可能であることが示唆された。また、化合物結合部位においても、5-chlorotryptamin 及び 5-methoxytryptamin において、凍結結晶でのみ、alternate conformation が観測された (Figure 4-11)。これについても、



結晶凍結による物理的外力によって、alternate conformationが生じたと考えられる。

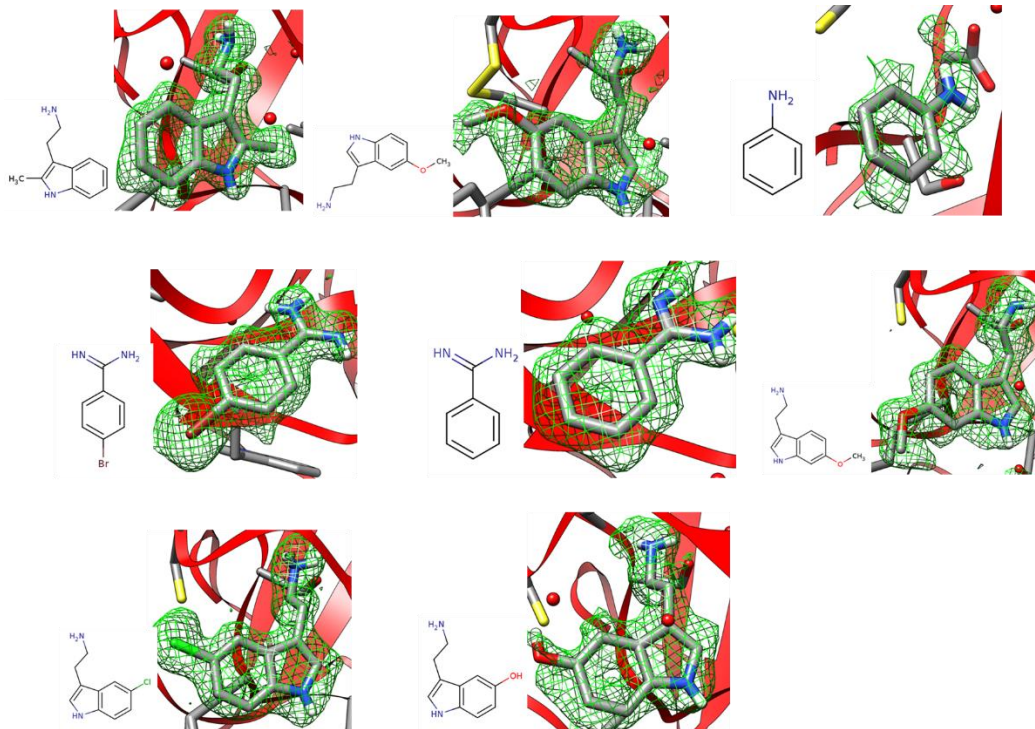


Figure 4-9 結晶捕捉アレイを用いて得られたトリプシンと種々の複合体結晶構造  
( $2.0\sigma$  mFo-DFc マップ)

化合物スクリーニングを行った 9 化合物のうち 8 種類の化合物において、化合物由来の電子密度が化合物結合部位に確認された。トリプシンと特異的な相互作用を有さない aniline 以外の化合物については、それぞれの化合物における官能基の電子密度まで明瞭に区別することが可能であった。

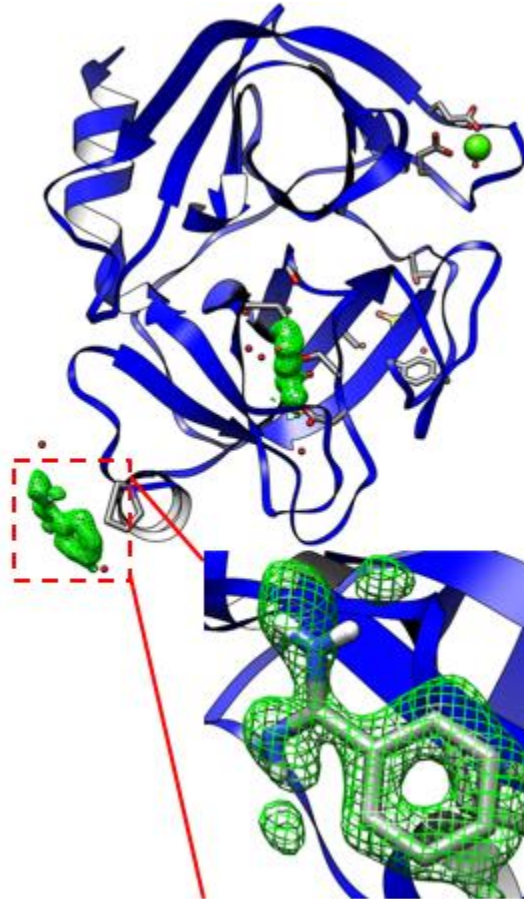


Figure 4-10 低温構造でのみ観測された benzamide の extra site への結合 ( $2.0 \sigma$  mFo-DFc マップ)

結晶中で整列しているトリプシン同士の中に挟まるように benzamide が結合していた。

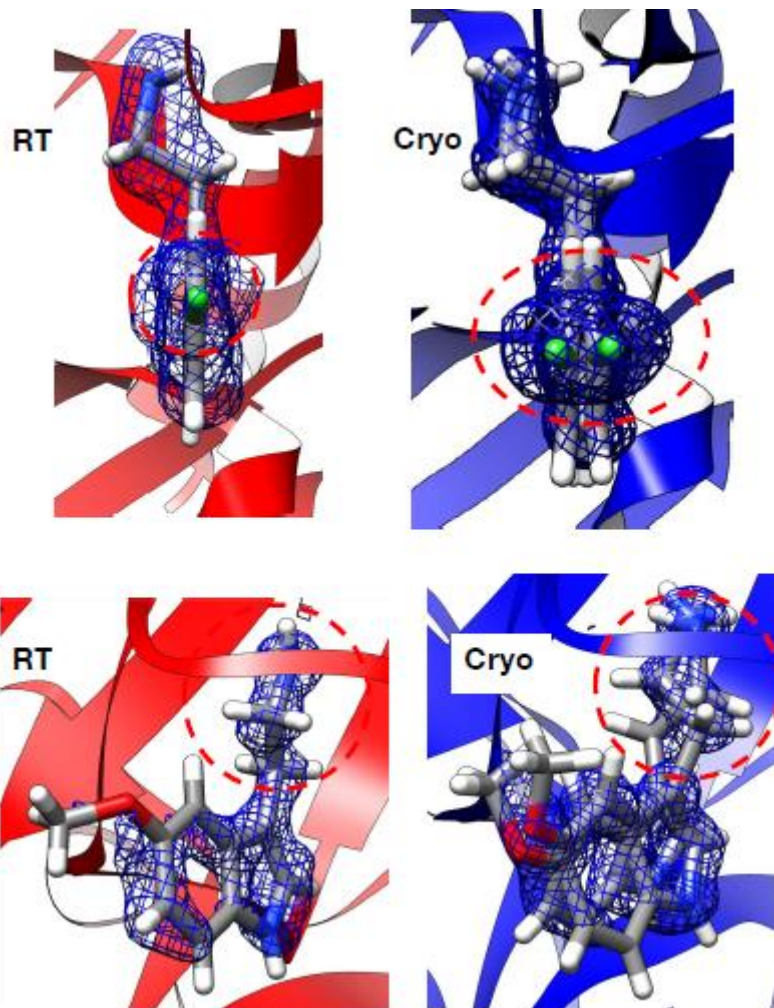


Figure 4-11 常温、低温での 5-chlorotryptamin 及び 5-methoxytryptamin 構造比較  
( $1.0 \sigma$   $2mFo-DFc$  マップ)

#### 4-4 結論

本研究では、ソーキング実験の簡便化、省力化および、常温構造解析の高速化を目指し、結晶捕捉アレイを用いたソーキング法による化合物スクリーニングの系を構築した。リゾチームとソーマチンを用いた結晶捕捉の検証実験では、いずれの結晶についても捕捉率 80%以上であった。また、結晶捕捉アレイをそのままゴニオメーターにマウントし、*in situ*で測定およびその回折強度データを用いた構造解析が可能であることを示した。トリプシンと種々の化合物を用いた化合物スクリーニングでは、化合物結合部位において、低温構造との化合物構造差を明らかにした。結晶捕捉アレイを利用することで、従来よりも常温構造解析を簡便かつ迅速に行うことが可能となった。

#### 4-5 参考文献

- [1] W. Du., L. Li., K. P. Nichols., R. F. Ismagilov., *Lab Chip*. (2009). **9**, 2286-2292.
- [2] M. Maeki, S. Yamazaki, A. S. Pawate, A. Ishida, H. Tani, K. Yamashita, M. Sugishima, K. Watanabe, M. Tokeshi, P. J. A. Kenis and M. Miyazaki, *CrystEngComm*. (2016). **18**, 7722-7727.
- [3] Li, L., Mustafi, D., Fu, Q., Tereshko, V. Chen, D. L., Tice, J. D., Ismagilov, R. F.,. *Proc Natl Acad Sci U S A*. (2006). **103**, 19243-19248.
- [4] Hansen, C. L., Skordalakes, E., Berger, J. M., Quake, S. R., *Proc Natl Acad Sci U S A*. (2002). **99**, 16531-16536.
- [5] Du, W. B., Sun, M., Gu, S. Q., Zhu, Y., Fang, Q., *Anal Chem*. (2010). **82**, 9941-9947.
- [6] Wang, J. W., Gao, J., Wang, H. F., Jin, Q. H., Rao, B., Deng, W., Cao, Y., Lei, M., Ye, S., Fang, Q., *Anal Chem*. (2019). **91**, 10132-10140.
- [7] Wang, L., Sun, K., Hu, X., Li, G., Jin, Q., Zhao, J., *Sensors and Actuators B: Chemical*. (2015). **219**, 105-111.
- [8] Zhu, Y., Zhu, L. N., Guo, R., Cui, H. J., Ye, S., Fang, Q., *Sci Rep*. (2014). **4**, 5046.
- [9] Schieferstein, J. M., Pawate, A. S., Varel, M. J., Guha, S., Astrauskaite, I., Gennis, R. B., Kenis, P. J. A., *Lab Chip*. (2018). **18**, 944-954.
- [10] Maeki, M., Yamazaki, S., Pawate, A. S., Ishida, A., Tani, H., Yamashita, K., Sugishima, M., Watanabe, K., Tokeshi, M., Kenis, P. J. A., Miyazaki, M., *CrystEngComm*. (2016). **18**, 7722-7727.
- [11] Maeki, M., Teshima, Y., Yoshizuka, S., Yamaguchi, H., Yamashita, K., Miyazaki, M., *Chem Eur J*. (2014). **20**, 1049-1056.

- [12] Ferreira, J., Castro, F., Rocha, F., Kuhn, S., *Chemical Engineering Science*. (2018). **191**, 232–244.
- [13] Sui, S., Wang, Y., Dimitrakopoulos, C., Perry, S., *Crystals*. (2018). **8**, 3082–3096.
- [14] Sui, S., Wang, Y., Kolewe, K. W., Srajer, V., Henning, R., Schiffman, J. D., Dimitrakopoulos, C., Perry, S. L., *Lab Chip*. (2016). **16**, 3082–3096.
- [15] Maeki, M., Pawate, A. S., Yamashita, K., Kawamoto, M., Tokeshi, M., Kenis, P. J., Miyazaki, M., *Anal Chem*. (2015). **87**, 4194–4200.
- [16] Maeki, M., Yamaguchi, H., Tokeshi, M., Miyazaki, M., *Anal Sci*. (2016). **32**, 3–9.
- [17] Ren, Z., Ayhan, M., Bandara, S., Bowatte, K., Kumarapperuma, I., Gunawardana, S., Shin, H., Wang, C., Zeng, X., Yang, X., *Lab Chip*. (2018). **18**, 2246–2256.
- [18] Yang, H., Chen, W., Peczulis, P., Heng, J. Y. Y., *Crystal Growth & Design*. (2019). **19**, 983–991.
- [19] Gerard, C. J. J., Ferry, G., Vuillard, L. M., Boutin, J. A., Ferte, N., Grossier, R., Candoni, N., Veessler, S., *Crystal Growth & Design*. (2018). **18**, 5130–5137.
- [20] Fukuyama, M., Akiyama, A., Harada, M., Okada, T., Hibara, A., *Analytical Methods* (2015). **7**, 7128–7131.
- [21] Gerdts, Cory J. Mark Elliott, Scott Lovell, Mark B. Mixon, Alberto J. Napuli, Bart L. Staker, Peter Nollert & Lance Stewart. *Acta Cryst D*. **64**, (2008). 1116–1122.
- [22] David J. Haas. *IUCrJ*. (2020). **7**, 148–157.
- [23] B. W. Matthews. *ibid.* (1968). **33**, 491–497.
- [24] Henderson, R. *Proc. R. Soc. Lond. B*. (1990). **241**, 6–8.

- [25] James, S. Fraser, Michael, W. Clarkson, Sheena, C. Degnan, Renske, Erion, Dorothee, Kern, Tom, Alber. *Nature*. (2009). **462**, 669-673.
- [26] Burnley, T. B., Afonine, P. V., Adams, P. D. & Gros, P. *eLife*. (2012). **1**, e00311.
- [27] Roedig, P. et al. *J. Appl. Crystallogr.* (2016). **49**, 968-975.
- [28] Edlund, P. et al. *Sci. Rep.* (2016). **6**, 35279.
- [29] Naitow, H. et al. *Acta Crystallogr D Struct Biol.* (2017). **73**, 702-709.
- [30] Fenwick, R.B., van den Bedem, H., Fraser, J.S. & Wright, P.E. *Proc. Natl. Acad. Sci. USA*. (2014). **111**, 445-454.
- [31] van den Bedem, H., Fraser, J. S. *Nat. Methods*. (2015). **12**, 307-318.
- [32] Fraser, J. S., van den Bedem, H., Samelson, A. J., Lang, P. T., Holton, J. M., Echols, N., Alber, T., *Proc Natl Acad Sci U S A*. (2011). **108**, 16247-16252.
- [33] Fischer, M., Shoichet, B. K., Fraser, J. S., *Chembiochem*. (2015). **16**, 1560-1564.
- [34] Colletier, J. P., Bourgeois, D., Sanson, B., Fournier, D., Sussman, J. L., Silman, I., Weik, M., *Proc Natl Acad Sci U S A*. (2008). **105**, 11742-11747.
- [35] Skrzypczak-Jankun, E., Borbulevych, O. Y., Zavodszky, M. I., Baranski, M. R., Padmanabhan, K., Petricek, V., Jankun, J., *Acta Crystallogr D Biol Crystallogr.* (2006). **62**, 766-775.
- [36] Kalinin, Y., Kmetko, J., Bartnik, A., Stewart, A., Gillilan, R., Lobkovsky, E., Thorne, R. *J. Appl. Crystallogr.* (2005). **38**, 333-339.
- [37] Ueno, G., Kanda, H., Hirose, R., Ida, K., Kumasaka, T., Yamamoto, M., *J Struct Funct Genomics*. (2006). **7**, 15-22.
- [38] Ueno, G., Kanda, H., Kumasaka, T., Yamamoto, M., *J Synchrotron Rad*. (2005). **12**, 380-384.



- [39] Paithankar, K. S. & Garman, E. F. *Acta Cryst D.* (2010). **66**, 381-388.
- [40] Southworth-Davies, R. J., Medina, M. A., Carmichael, I. & Garman, E. F. *Structure.* (2007) **15**, 1531-1541
- [41] K. Yamashita, K. Hirata, and M. Yamamoto. *Acta Cryst* (2018). D74, 441-449.
- [42] McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. *J. Appl. Cryst.* (2007). **40**, 658-674.
- [43] Emsley, P., Lohkamp, B., Scott, W. G., Cowtan, K. *Acta Crystallogr., Sect. D: Struct. Biol.* (2010). **66**, 486-501.
- [44] Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., Zwart, P. H. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* (2010). **66**, 213-221.

## 第5章

### 総括

## 5-1 本研究の要約

本研究では、X線結晶構造解析における、回折強度測定実験の高速化、効率化、省力を目指し、①大型結晶を対象とする深層学習を用いた自動回折強度測定システム ②結晶捕捉アレイを用いた化合物ソーキングシステムを開発し、上記の問題解決に取り組んだ。

第2章では、昨今画像処理分野で大きな技術革新があった深層学習を用いて、自動結晶センタリングシステム、DeepCenteringを開発した。それまで画像処理技術を用いて開発されてきた自動センタリングと比較して、高精度のセンタリングを達成した。また、DeepCenteringと自動分子置換、精密化のためのccp4プログラムであるdimpleをシームレスに使用するための自動構造解析パイプラインimaprocを開発した。リゾチームを用いた自動構造解析パイプラインの検証実験では、手動センタリングの場合と比較して遜色ない構造精密化の統計値であった。

第3章では、DeepCenteringとimaprocの応用例として、ヒト血清アルブミンのアポ結晶および、ヒト血清アルブミンと環状ペプチドの複合体結晶スクリーニングを行った。溶媒含有率が76%と非常に高いアポ結晶に対して、合計128個の結晶からスクリーニングを行い、最終的に2.6 Åでの構造解析を行った。また、ヒト血清アルブミンと環状ペプチドの複合体結晶のスクリーニングにおいては、DeepCenteringとimaprocによる自動構造解析では複合体結晶の取得には至らなかったが、その後の共結晶実験への方針転換を迅速に行うことに貢献した。アポ構造とDalbavancin結合構造とを詳細に比較することにより、Dalbavancinの結合部位や、induced-fit機構によるDalbavancinの結合様式を明らかにした。

第4章では、ソーキング実験の簡便化、省力化および、常温構造解析の高速化

を目指し、結晶捕捉アレイを用いたソーキング法による化合物スクリーニングの系を構築した。リゾチームとソーマチンを用いた結晶捕捉の検証実験では、いずれの結晶についても捕捉率 80%以上であった。また、結晶捕捉アレイをそのままゴニオメーターにマウントし、*in situ*で測定およびその回折強度データを用いた構造解析が可能であることを示した。トリプシンと種々の化合物を用いた化合物スクリーニングでは、化合物結合部位において、低温構造との化合物構造差を明らかにした。

## 5-2 今後の展望

本研究では、自動結晶検出システム DeepCentering および結晶捕捉アレイを用いたタンパク質-化合物複合体のスクリーニング系の構築を行った。既に述べたように、これまでに、凍結結晶を用いた回折実験の自動化は達成されている。一方で、常温での回折強度測定については、放射線損傷による影響により、ラスタースキャンの使用が難しいことから、これまでに自動化はなされていない。そこで、自動化のために、画像処理技術を用いた結晶検出システムである DeepCentering を用いることで、常温構造解析におけるセンタリングの自動化が可能になると期待される。同様に、すでに多くの放射光施設で利用がなされている、結晶化プレートをそのままゴニオメーターに取り付けて常温で *in situ* 測定を行う plate scan への応用も期待される。具体的には、DeepCentering と結晶捕捉アレイによる結晶捕捉および *in situ* 測定と組み合わせることで、より効率的な常温 X 線回折実験が新たな実験の選択肢になり得ると考えられる。また、本研究で開発を行った結晶捕捉アレイは、流路中に 1 列のピットを配置し、1 種類の化合物に対するソーキングを行うために開発されたが、複数列の流路を 1 つの結晶捕捉アレイとして統合したデバイスを作製(Figure 5-1)することで、一度の X 線回折実験で複数の化合物ソーキング条件の X 線回折強度測定が可能になると考えられる。実験の更なる自動化についても改善が必要である。本研究では、結晶捕捉アレイへの結晶懸濁液および化合物溶液の導入はピペッティングによる手作業で行ったが、この工程についても、シリンジポンプ等を利用して自動化を行い、実験の省力化のためにさらなる開発が望まれる。

(a)

(b)



Figure 5-1 二次元化を行った結晶捕捉アレイ

(a)二次元化流路の模式図および(b)試作品を示す。結晶捕捉アレイの二次元化により、一度に複数条件の化合物ソーキングを行う。

## 謝辞

本研究を行うにあたり、ご指導、ご鞭撻を賜りました山本雅貴教授に心より感謝いたします。社会人博士課程という不規則な形態での研究でしたが、先生のお気遣いによって無事に研究を遂行することができました。ありがとうございました。

直前のお願いにもかかわらず、学位審査の副査を務めていただきました、兵庫県立大学西谷秀男教授、久保稔教授、高輝度光科学研究センター熊坂崇先生に深く感謝いたします。

私がプログラミングに興味を持つきっかけとなった、理化学研究所 上野剛博士に深く感謝申し上げます。「こういうプログラムが作りたい」と相談すると、そのコンセプトを絶対に否定せずに、開発を行う機会を与えていただけたからこそ、曲がりなりにもプログラミングができるようになりました。思ったことをいつでも相談できる環境を作っていただけたことも感謝しております。ブリ釣りに行けなかったことだけが心残りです。

結晶学を理解した気になっていた私にその奥深さを示していただいた、理化学研究所 山下恵太郎博士（現 MRC）に深く感謝いたします。日々是精進で山下さんに追いつけるよう研鑽をつづけていきたいと思えます。

Wet の実験ご指導および、HSA-環状ペプチド論文執筆の際にご指導を賜りました、理化学研究所 吾郷日出夫博士に深く感謝申し上げます。吾郷先生の日々の研究活動を間近で拝見し、サイエンスと真摯に向き合う大切さを教わりました。ありがとうございました。

学位申請の手続きおよび本論文のご確認においてで多大なご協力をいただきました、山本研究室松浦滉明博士に深く感謝いたします。

論文執筆時の的確なアドバイスや普段の研究時に議論いただきました山本研究室の皆様にも深く感謝いたします。

本研究を遂行する上で不可欠な結晶捕捉アレイの設計、作製のノウハウをご教授いただきました、共同研究先の北海道大学 真栄城正寿先生および竹田怜央氏に深く感謝いたします。

本研究で開発を行ったシステムの検証実験の際にサンプル提供をしていただきました、東京大学 津本浩平先生、長戸石暁先生、妹尾暁暢氏に深く感謝いたします。

普段から研究におけるアイデアのディスカッションや、互いに研鑽をしてきた、東京工業大学 生命理工学研究科所属時の先輩である和久井直樹博士、後輩の吉田浩平博士にも深く感謝いたします。当時、膜タンパク質の構造解析を目指し、もがき苦しんだのは懐かしい思い出です。構造解析を生業にしたのは私だけですが、またいずれ互いの専門分野を活かして一緒に仕事をしましょう。

社会人博士課程での研究を快諾いただいた、また、入学のための諸手続きをしていただきました、株式会社リガク応用技術センター 山野昭人博士に深く感謝いたします。また、構造解析に関する議論をしていただきました、株式会社リガク応用技術センター 単結晶解析グループ 山崎幹男博士、松本崇博士、佐藤寛康博士、菊池貴博士にも深く感謝いたします。



最後に、毎日欠かさずお弁当を作ってくれたり、他愛ない会話で元気づけてくれた妻に感謝いたします。