



SPring-8・理化学研究所播磨研究所ハイスルーブットファクトリーにおける蛋白質大規模 X 線結晶構造解析にむけた取り組み

国島直樹 理化学研究所播磨研究所ハイスルーブットファクトリー 〒679-5148 兵庫県佐用郡三日月町光都 1-1-1
菅原光明 理化学研究所播磨研究所ハイスルーブットファクトリー 〒679-5148 兵庫県佐用郡三日月町光都 1-1-1

要旨

ヒトをはじめとする150以上の生物の遺伝子配列が解読されたが、それらの遺伝子機能の多くは未知である。構造ゲノム科学の目的はタンパク質の基本構造を網羅的かつ系統的に解析することによりそのゲノム機能を明らかにすることである。近年、遺伝子産物であるタンパク質の構造と機能を網羅的に解き明かそうとする大規模プロジェクトが日米欧を中心とした世界各国で始まった。本稿では、SPring-8 キャンパスにある理化学研究所播磨研究所のハイスルーブットファクトリーにおける高効率タンパク質 X 線結晶構造解析に向けた試みについて紹介する。

1. はじめに

蛋白質 X 線結晶構造解析は、高輝度で波長可変の放射光の利用と、検出器や計算機、解析ソフトなどの周辺技術の進歩により、回折能の高い良質な結晶が一つあれば、1週間程度で構造決定ができるまでになった。しかし、ヒトゲノムの配列完読の発表をうけて、約1万種類といわれる蛋白質の基本構造の解明を数年で達成しようとする構造ゲノム科学プロジェクトが始まり、そのための作業の効率化、自動化を含めた新たな技術開発が世界各所の放射光施設や大学、研究機関、企業が参入して国際的な競争と協調の下に進められている¹⁾。世界最大規模の第3世代大型放射光施設 SPring-8 では、1日に多数の結晶の回折チェックからデータ測定までを連続して自動で行うために理研構造ゲノムビームライン (BL26B1/B2) が建設され、運用に向けた準備が進められているところである²⁾。理研播磨研究所ハイスルーブットファクトリーは、ビームラインを有効に活用するために、これまで手間と時間のかかっていた蛋白質試料の調製から結晶化および構造決定のプロセスをハイスルーブット (高効率) に行うための施設であり、これまでおもに耐熱性の細菌由来の安定な蛋白質をパイロットとして、多数の試料の処理を可能とする蛋白質の生産から結晶化および構造解析計算の効率のよいシステムの構築および各工程からの膨大な実験データの統合、整理を行うハイスルーブットファクトリーデータベース (HTPF-DB) などのハードウェアおよびソフトウェアの自動化に向けた開発を内外の研究機関および日本の企業数社と共同でおこなってきた。SPring-8/理研構造ゲノムビームラインと連動した一通りの流れができ、順調に稼働しているところである³⁾。本稿では、理研ハイスルーブットファクトリーの大規模蛋白質結晶構造解析に向けた取り組みについて紹介したい。

2. 構造ゲノム科学

近年150を超える生物のゲノム情報が次々と解明される中で、2003年4月およそ30億塩基対からなるヒトの全ゲノム配列の完全解読が報告され遺伝子の数は32000前後であると推定された。しかし、ゲノム機能の大部分は依然として未知である。ポストゲノム研究の大きなテーマの一つとして遺伝子機能の発現主体であるタンパク質の全基本構造を網羅的かつ体系的に解析し、ゲノム機能を解き明かそうとする構造ゲノム科学プロジェクトが提唱された。2000年6月のヒトゲノムの概要配列の発表以来、推定された1万種類の全基本構造の解明を目標に掲げて、世界各国で大規模プロジェクトが発足した。2001年4月に開かれた第2回国際構造ゲノム会議において、構造ゲノム科学を国際的に協調して進める上で必要な合意がなされ、ISGO (国際構造ゲノム科学機構) がスタートした (<http://www.isgo.org/>)。

理化学研究所においては、世界に先駆けて1995年に理化学研究所が構造ゲノム/プロテオミクス研究として NMR 及び X 線結晶解析によるタンパク質の立体構造と機能を解析するプロジェクトを提案し、1997年から「タンパク質基本構造解明プロジェクト」(現横浜 GSC 蛋白質構造・機能研究グループ: 横山プロジェクトディレクター, <http://www.gsc.riken.go.jp/j/group/themeproteinJ.html>) が発足した。1999年には高度好熱菌をターゲットとしたストラクチュローム連携研究グループ (倉光グループリーダー, <http://www.thermus.org>) がスタートした。2001年にこれらを融合する形で RSGI (理研構造ゲノム/プロテオミクス推進本部; <http://www.rsgi.riken.go.jp/>) が発足し、横浜研究所ゲノムサイエンスセンター (横浜 GSC) と播磨研究所が協力してプロジェクトを推進する体制となった⁴⁾。2002年に国の委託事業

である新世紀重点研究創生プラン (RR2002) の重要テーマとして、「我が国初のゲノム創業の実現等を目指し、我が国の研究機関の能力を結集して、平成14年度から5年間でタンパク質の全基本構造の1/3 (約3000種) 以上のタンパク質の構造およびその機能を解析し、特許化まで視野に入れた研究開発を推進することを目的とする。」タンパク3000プロジェクトが理化学研究所と7大学研究機関によって開始された (http://www.mext.go.jp/b_menu/houdou/14/02/020213ha.htm)。

3. ハイスループットファクトリー (HTPF)

ハイスループットファクトリー (HTPF) は2001年4月に大型放射光施設 SPring-8 のキャンパス内にある播磨研究所にタンパク質の大量高速構造解析を目的として発足した (Fig. 1)。2002年からは RSGI (理研構造ゲノム/プロテオミクス推進本部) の枠組みの中でタンパク3000プロジェクト「タンパク質基本構造の網羅的解析プログラム」を受託し、SPring-8の強力な放射光を利用してX線結晶解析によるタンパク質の立体構造決定をそのハイスループット化に必要な技術開発とあわせて進めている。これまでに、おもに高度好熱菌 *Thermus thermophilus* HB8 や超好熱古細菌 *Pyrococcus Horikossii* OT3 由来の蛋白質をターゲットとして、多数の蛋白質試料に対応した効率のよい遺伝子発現から蛋白質精製までのプロセスの確立、全自動結晶化観察ロボットシステム「TERA」 (Fig. 2)、迅速X線蛋白質結晶評価装置 (Fig. 3) の開発を行い、さらさらの過程で生じる大量のデータを整理するハイスループットファクトリーデータベース「HTPF-DB」などを構造ゲノムビームラインと連携する形で開発整備し、運用を始めている。これらのシステム開発は、理研のノウハウと優れた技術をもつ日本の複数の企業の協力、共同研究開発によって進めている。各ターゲットについての発現ベクターの構築、タンパク質の発現、精製、結晶化、X線回折、立体構造のモデリング、解析結果をPDBに登録といった進捗状況は月一回、HTPF-DBの各データをもとに自動で更新され、HTPFの公式ホームページで公開している



Figure 1. RIKEN Highthroughput Factory at SPring-8. At SPring-8 campus, we have been established an integrated high-throughput synchrotron radiation protein crystallographic facility as RIKEN Highthroughput Factory including new development of many elements for the purpose in the early phase.

(<http://www.htpf.harima.riken.jp>, Table 1)。今後は、ISGO (国際構造ゲノム科学機構) の基準に準拠した形で RSGI 全体のデータベースの中で進捗報告が行われる予定である。

蛋白3000プロジェクトが開始してから2年が経過し、設備や装置、データベースなどインフラの整備がほぼ整いつつあり、簡単なターゲットについては構造解析が容易になってきた。現在までの成果を維持する一方で、これまでの体制を見直し、より解析困難なターゲットに挑戦してゆく技術開発も必要となる。内部の連携はもちろん、大学、研究機関、企業などとのよりいっそうの連携が重要である。



Figure 2. Full Automatic Protein Crystallization and Observation Robot System "TERA".

There are three components; (1) liquid-handling robot for dispensing crystallization setup robot in the right of front, (2) microscopic photo system with CCD camera and XY plate loader at the center of front, and (3) storage for 2,500 crystallization micro plates and 125 crystallization reagent plates in the back. Close-up views: parts of the microscopic photo system (bottom-center panel) and the crystallization robot (bottom-right panel). Upper right shows various protein crystals.

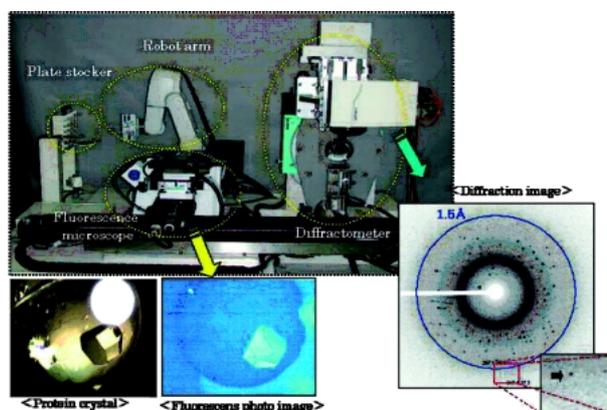


Figure 3. *in situ* X-ray diffractometer for protein crystal. The system accepts the microbatch crystallization plate and can measure X-ray diffraction of protein crystals without mounting. The system consists of two major parts: optical microscope and X-ray diffraction parts.

Table 1. The over-all tracking statistics of structural determination of the proteins from the thermophile bacteria, *Thermus thermophilus* HB8 and *Pyrococcus horikoshii* OT3 by January, 2004

	<i>T. thermophilus</i> (HB8)	<i>P. horikoshii</i> (OT3)
Expressed	215	855
Soluble	188	465
Purified	170	214
Crystallized	90	64
Diffraction quality Crystals	79	51
Native diffraction data	58	24
Phasing	51	19
Crystal structure	39	8
In PDB	38	8

4. 蛋白質の生産

X 線結晶構造解析の成否は生産されたタンパク質の性質に帰着されることもしばしばであるが、蛋白質の性質は多種多様であり精製手順も様々である。多数の試料について特に X 線結晶構造解析で要求されるグレードの高純度で大量のサンプルの自動調製は実質大変困難であり、もっとも人手と時間がかかる行程である。大規模かつ迅速にサンプルを調製する様々な努力は、大腸菌高密度培養⁵⁾や無細胞発現⁶⁾、タグ付きタンパク質試料のワンステップ精製⁷⁾など各方面でなされているところである。HTPFにおいては、おもに高度好熱菌と超高熱古細菌由来の安定な蛋白質試料を用いて、供与された大量発現ベクターのスマールスケールでの発現チェックから、大腸菌の発現系を用いて培養から精製調製、結晶化に適した試料の分析までの、効率のよいシステム構築を行ってきた。大腸菌の培養は発現チェックの結果に応じて 2L~9L の培地を用い、収量がほぼ一定になるようにしている。またいくつかの検討を経て、収量の実績の高い精製プロトコルを標準化し、作業の単純化をはかった。結晶化に用いる精製タンパク質試料について、ネイティブ電気泳動や動的溶液光散乱測定、N 末単アミノ酸分析などにより、溶液中での蛋白質分子の物理化学的均一性を確認して、結晶化の適否を判断しながら行っている (Fig. 4)。

5. 全自動結晶化観察ロボット「TERA」

蛋白質結晶化のプロセス (Fig. 4) は、X 線結晶構造解析のプロセスの中でもっとも手間のかかる作業である。構造解析に適した分解能の高い良質な結晶を得るためには、数百から数千条件にも及ぶ結晶化条件をスクリーニングするため結晶化セットアップを行い、この結晶化ドロップの観察をする必要がある。回折測定が可能となる約 0.1 mm 以上の単結晶が得られれば、この結晶を顕微鏡下で操作して、グリセロールやオイルなどを含んだ抗凍結剤に浸したのち、強力な放射光 X 線による損傷を低減するため、液体窒素温度の気流中で瞬間凍結して X 線による回折能、

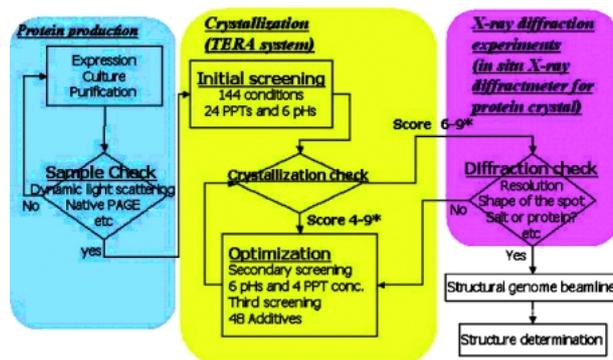


Figure 4. The crystallization protocol in HTPF.

モザイク度など結晶の結晶学的な質のチェックをおこなう。これらをすべてパスした結晶のみが結晶構造解析計算に使うための回折データセット収集の対象となる。また、こうした結晶評価の結果、X 線結晶構造解析に適しないと判定された時は、結晶化の再検討、時にはもう一度サンプルの発現・精製段階にまでさかのぼって検討し直すことになる。従来、これらの作業を 1 人の研究者が 1 つ 1 つのサンプルについて手作業で行ってきた。構造ゲノム科学では多数のサンプルに対応した標準的な結晶化プロトコル、結晶化ドロップ、生成した結晶の質の評価基準を確立し、必要に応じて再検討することが重要である。

理化学研究所では HTPF で設定した標準的なプロトコル (Fig. 4) に基づき結晶化分注および観察などの結晶化スクリーニング作業を行なう全自動蛋白質結晶化観察ロボットシステム「TERA (テラ)」(Fig. 2) を設計開発した⁸⁾。TERA はオイルマイクロバッチ法を採用した結晶化分注ロボット、結晶化ウェル自動観察装置、2,500 枚の結晶化用プレートと 125 枚の試薬プレートの保管棚、各々の装置間を連携するロボットハンドをもつプレート搬送系ロボットからなるシステムである (Fig. 4)。

結晶化分注装置については、いくつかの既存の結晶化ロボットで実際に結晶化を行い検討した結果、動作が単純で自動化に適している点、もっとも広く行われている蒸気平衡法との比較対照実験より結果が良かったオイルマイクロバッチ法で結晶化を行う英国ダグラス社製の IMPAX1-5 の機構⁹⁾を同社の了承を得て採用し、さらに分注精度と分注速度を向上させた (Fig. 2)。

結晶化スクリーニングは米国 Hampton Research 社¹⁰⁾などの複数の市販されているスクリーニングキットから、使用頻度の高い沈殿剤、緩衝剤、添加剤の条件を整理し、独自の初期スクリーニング条件を設定し効果のある条件を取り入れて評価、見直しが行えるようにデザインしている。代表的な沈殿剤および沈殿剤と添加剤の組み合わせ 24 種類と pH 4.5~9.5 までの 6 種類の緩衝剤を組み合わせた独自の 144 条件の 1 次スクリーニング、この初期結晶化スクリーニング条件のそれぞれに対応して各初期条件の

pH と沈殿剤濃度をわずかに変化させた近傍の24条件を第2ステップスクリーニングの展開用結晶化条件セットとして3456条件、結晶の質の改善に効果のあるとされる添加剤スクリーニングを48条件とした。これらの試薬は全部で約3,500種類にも及び、人手による試薬作業では時間と労力を要するばかりでなくミスや誤差の恐れも大きいので、市販の試薬分注ロボットを利用した試薬調製法を確立し、ひとりのスタッフがすべての溶液を2~3日で調製可能である。全結晶化試薬は1ヶ月ですべて交換する。

結晶化の観察は、数百から数千条件にもおよぶ結晶化セットアップの定期的な観察および評価を行う結晶化実験のなかで最も人手がかかる作業である。結晶化ドロップの顕微鏡からの像は観察者が判定を行なえる鮮明な画像イメージとして自動撮影する倒立型顕微鏡にXYZステージとCCDカメラを搭載した結晶化ドロップ観察装置を開発導入した (Fig. 2)。この装置による結晶化用プレートの結晶化ドロップ写真は複数の画像を自動再構成し、結晶がドロップのどの位置に成長しても鮮明となる。蛋白質の結晶

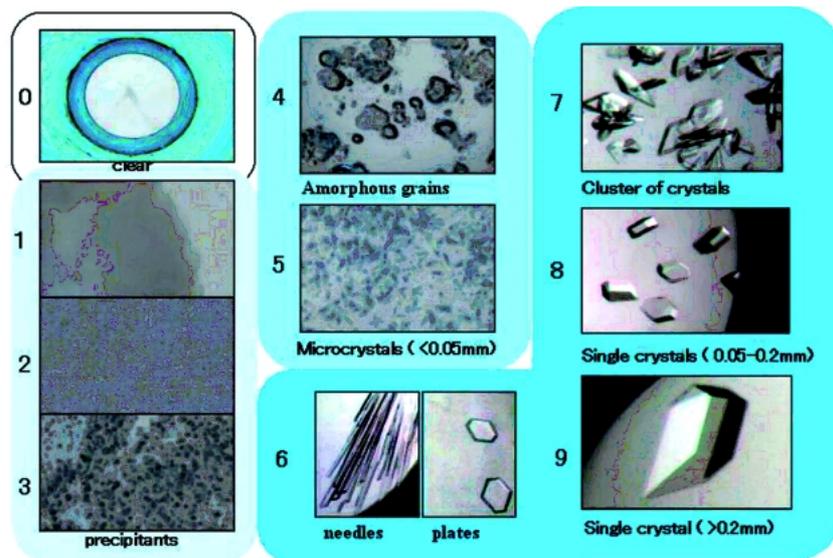


Figure 5. The score for the evaluation of crystallization setting.

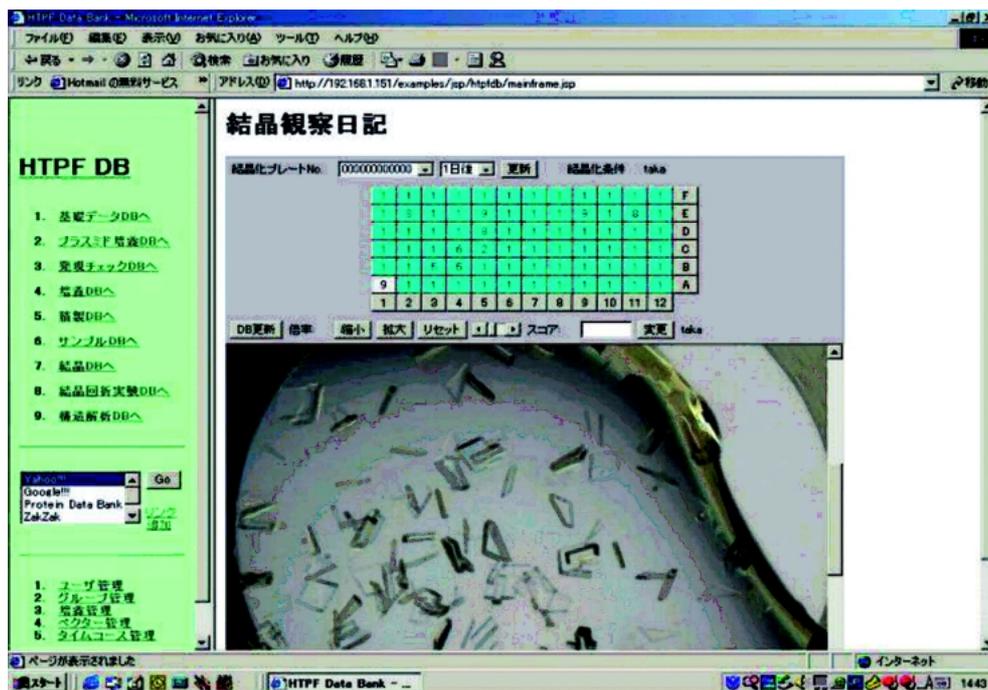


Figure 6. Observation of crystallization setting on the web browser of HTPF-DB.

化ドロップはさまざまな色・形をした沈殿、非晶質の凝集物および結晶性の沈殿物やそれらの組み合わせた複雑な様相を呈するので、実験記録は煩雑で、統一をしておかないと実験記録の間で比較不能となり、データの整理さえ困難になる。そこで、評価観察者の判別のしやすさに留意した上で、結晶化ドロップ中に典型的に現われる様相を 0：透明，1～3：沈殿（茶色，白色，透明），4：非晶質の粒，5：大きさ50 μm 以下の微結晶，6：針状，板状結晶，7：結晶のクラスター，8, 9：50 μm 以上の3次元単結晶の10段階とした独自の判定基準を設定した（Fig. 5）。現在，TERA はパソコンのコンソールから各構成ユニットを一括管理し，初期から3次スクリーニングまでの結晶化作業，結晶化ドロップの状態の経時変化を自動的に画像イメージとして定期的に撮影保存できる結晶観察作業をスケジュール化して実行できる。全自動結晶化システムからのあらゆる情報はハイスループットファクトリー全体の実験データベースに自動的に保存され，観察者はインターネットを通じて web ブラウザーから結晶化作業の要求や結晶化ドロップの観察画像の判定を随時行なうことができる（Fig. 6）。

全自動蛋白質結晶化・観察ロボットの導入によって，いままで手作業で行なっていた結晶化作業，結晶観察作業が24時間自動で行なえるようになり，大幅に作業の効率と結晶化の再現性が上がった。現在の TERA の作業速度から，僅か数人で年間千以上の試料の結晶化作業が可能となった。さらに，結晶化作業の履歴，結晶化ドロップの観察画像など結晶化作業をするうえで扱わなければならない膨

大なデータが自動的に電子記録されるので，必要なデータをすぐにアクセスして利用できるように整備を進めている。今後，蓄積した結晶化ウェルの写真を使って画像処理による特徴量抽出などの手法を使って自動判定ソフトウェアの開発や蓄積される実験データをもとに蛋白質結晶化率の向上を目指したい。

6. 迅速 X 線結晶評価装置

結晶の評価は実際に X 線を照射し，回折能により判断しなければならない。放射光での測定には凍結した結晶を用いるが，このためには結晶化プレートのウェルから顕微鏡下で一つ一つ結晶を取り出さなければならない。このための抗凍結剤の種類や濃度などの条件により，結晶が損傷することもしばしばであり，結晶化プレートに生成した状態での結晶の評価は作業の効率化に大変重要である。HTPF では迅速 X 線結晶評価装置を理学電気と共同で開発した（Fig. 7）。このシステムによりマイクロバッチ用の結晶化プレート中の結晶を取り出さずに X 線回折チェックを行うことが出来る。迅速 X 線結晶評価装置は紫外線光学顕微鏡と X 線回折装置の2つの部分で構成されており，蛋白質の結晶は紫外線照射により蛍光を発するので，顕微鏡の画像により蛋白質の結晶かどうかの確認とその位置を CCD カメラからの画像を解析することにより決定する。X 線発生源として CuK α 線より波長の短い MoK α 線を用いプラスチックのプレートと結晶化母液による X 線の吸収を抑えている。X 線回折装置の X 線源と CCD 検出器は垂直方向に配置され，結晶は XYZ ステア

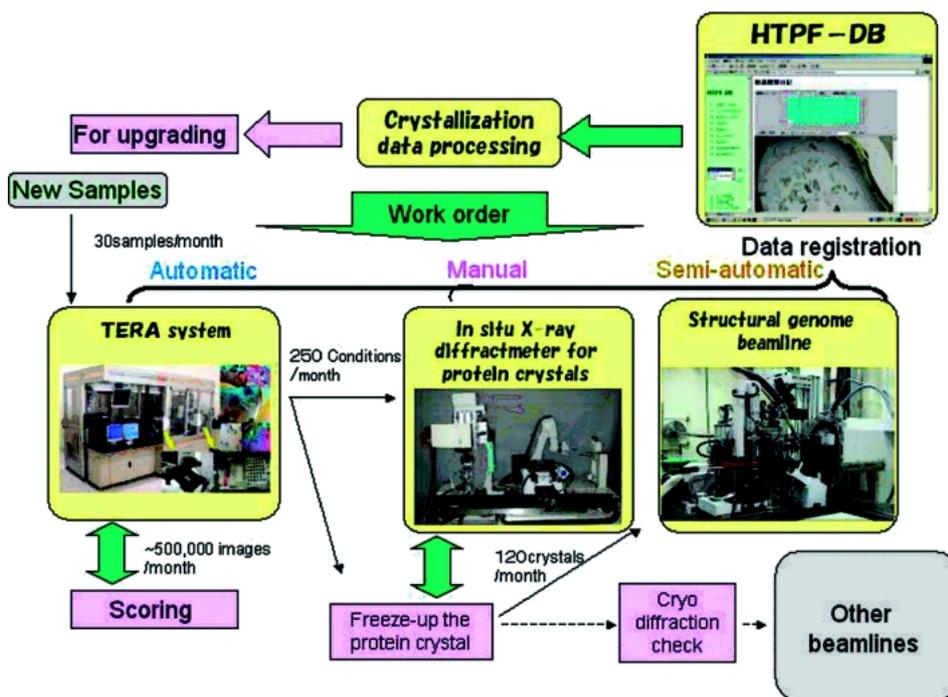


Figure 7. The view showing a frame format of work flow in HTPF.

ジ上で自動的に X 線に当たるように配置され、静止あるいは振動写真の撮影が可能である。この装置により迅速に分解能などの評価を行い、質のよい結晶を選別して凍結し、ビームラインでの測定に用いている。現在はまだテスト段階であり、一月に迅速結晶評価装置により約250個の結晶をチェックし、ビームラインで約120個の結晶のチェックおよびデータ測定を行い検討している。

7. X 線結晶構造解析

高分解能の非常によいデータがあれば Tom Terwilliger が開発した SOLVE/RESOLVE¹⁴⁾などのソフトウェアによって構造構築まではほぼ自動的に行えるようになった。しかし多くの場合、電子密度図に立ち返った全域にわたるタンパク質構造の精密化と決定された構造の評価はまだまだ時間のかかる段階であり、経験が必要なところである。精密化したタンパク質の立体構造の評価は複数の構造正当性のチェックをするソフトウェアを使って総合的に行うべきである。結晶化から回折データの再収集など実験の継続、部分構造モデル構築、イメージ処理や位相改良などの多くの結晶解析ソフトウェアをうまく組み合わせた具体的実験経験と結晶構造解析的センスが構造解析の成功には必須である。蓄積したノウハウをアルゴリズム化することで自動化が可能となる。

大規模な構造解析を効率良く行うためには、解析結果から有用な統計的情報が容易に抽出できることや、解析担当者が変わった時の実験データ引継ぎが容易なこと等が要求される。そのためには、各研究者の実験情報を一定の客観的なフォーマットで常時管理することが必要である。一方、研究の活性化のためには、各研究者の個別の手法をあ

る程度尊重しなくてはならない。ハイスループロットファクトリーでは、これらの要求を満たすための実験フォーマットの確立を行った。実験フォーマットは、回折予備実験 (Preliminary experiment)・回折データ (Data set)・位相付け (Phasing)・モデル構築および精密化 (Model building and refinement)・モデル確認 (Model check)・最終報告 (Final report) の6種類から構成され、各研究者に実験法の画一化を強いることなく解析情報の標準化ができるよう工夫されている。さらに、解析進捗の把握を正確に行うための進捗報告フォーマットも用意されており、各解析担当者から毎月提出されファイルされている。これらのフォーマットは実験データベースに組み込まれた形で構造解析のインフラとして利用されており、解析担当者にとって欠くべからざるものになっている (Fig. 8)。これまでの技術開発により、蛋白質結晶の回折チェックから PDB 登録までの解析情報の規格化がほぼ実現したといえる。これからは蓄積されたデータをいかに効率化に結びつけるかを考える必要があるが、一つには、標準化された解析手順をスクリプトにより自動化することが挙げられる。この解析標準スクリプトの開発により、経験の浅い担当者でも正しい手順で解析ができるようになることに加え、手順が標準化されていることから指導者による助言が容易になると期待される。もう一つには、解析プロトコルを改善するためのデータマイニングツールの開発が挙げられる。現在、蛋白質のアミノ酸配列を入力すると試すべき重原子化法が優先順位付きで出力されるシステムの開発を進めている。

解析された結晶構造を学術論文・特許・創業などに結びつけるための基盤整備もまた、タンパク3000プロジェクトの重要な目的である。構造と機能の関係を明らかにする

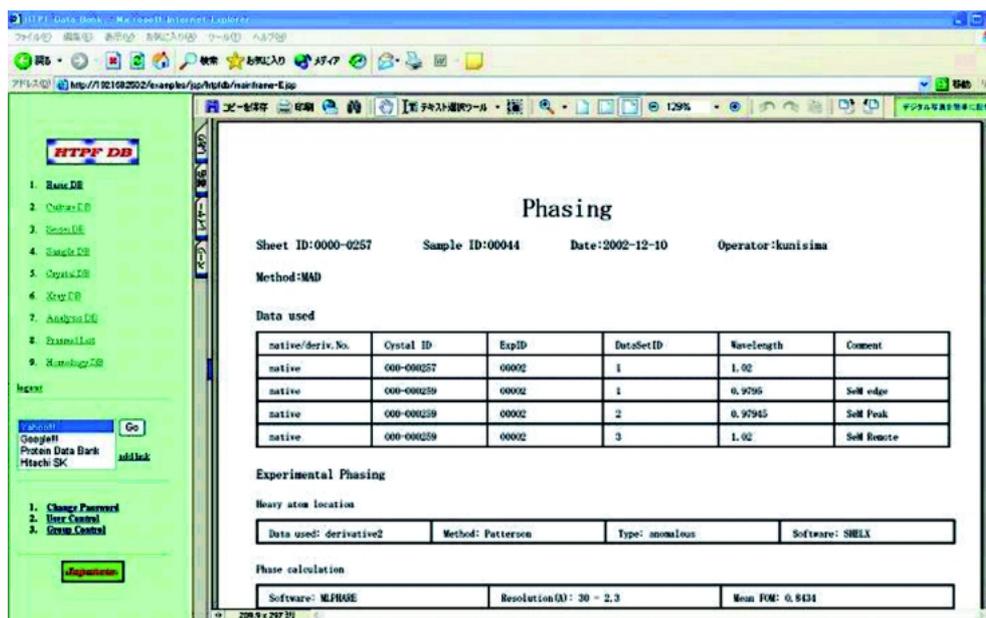


Figure 8. Experimental database for structural analysis in HTPF-DB.

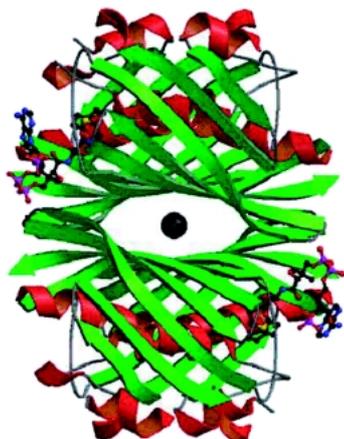


Figure 9. Ribbon diagram of the crystal structure of PaaI homotetramer from *Thermus thermophilus* HB8 (TT0310/HTPF00044) in hexanoyl-CoA liganded form.

ための機能研究の取り組みの例として、以下に *Thermus thermophilus* HB8 由来プロジェクトコード310番蛋白質の構造解析について紹介する。本蛋白質の構造解析はいろいろな意味でハイスループレットファクトリーにおけるモデルケースとなっており、これまで述べた効率化の試みの多くはこの解析例に基づいて行われている。また、我々はこれから機能解析の作業も標準化・効率化する必要があると考えている。具体的には、論文作成や特許出願の際に必要な構造評価基礎データの標準化を現在進めており、近い将来自動化を行う予定である。

310番蛋白質は、*paa* 遺伝子群によってコードされる一連のフェニル酢酸分解関連蛋白質のうちの機能未知蛋白質 PaaI である¹²⁾。本蛋白質の結晶構造は、セレノメチオン化蛋白質結晶を利用した MAD 法により決定された (Fig. 9)。分子モデルは 2.3 Å 分解能の初期電子密度図をもとにプログラム ARP/wARP¹³⁾ を用いて自動構築され、最終的に 1.7 Å 分解能で精密化された。構造解析自体にかかった時間は 1 週間以内であり、それ以後のほぼ 1 年間は機能解析に費やされた。まず、プログラム MATRAS¹⁴⁾ を用いて構造既知蛋白質との構造比較を行ったところ、hotdog fold 蛋白質と呼ばれるコエンザイム A 関連酵素群との構造類似性が見出された。このプログラムはハイスループレットファクトリー内のデータベースを利用しており、外部へのデータ漏洩の心配なく構造検索ができる。次に、この結果に基づき、各種アシルコエンザイム A を用いた酵素学的解析を行った。その結果、本蛋白質はフェニルアセチル CoA に特異的な加水分解酵素であることが示唆された。さらに、活性部位周辺のアミノ酸残基の部位特異的の変異体 8 種類をハイスループレットファクトリー内で作成し、それらの酵素学的解析を行った結果、本酵素の活性残基 Asp48 が同定された。次に、リガンドであるコエンザイム A 誘導体との複合体の結晶化を、自動結晶化口

ボット TERA⁸⁾ を用いて試みたところ、リガンドなしのものとは全く異なる条件で 2 種類の複合体結晶が得られた。また、TERA を用いて得られた 3 つの異なる結晶型の結晶構造比較から、本酵素がリガンドなしの状態では様々な形を取り得ることが示された。このようなリガンド添加を含む膨大な結晶化条件の検討は、TERA を用いれば比較的手軽に行うことができる。これらの情報に基づき、反応機構を含む本酵素の詳細な構造機能相関について考察することができた。また、本蛋白質が塩素イオンを結合していることを利用し、新しいタイプの塩素イオンキレート剤の開発法に関する特許を出願した (特願2003-050153)。

8. ハイスループレットファクトリーデータベース (HTPF-DB)

複数の研究グループが共同で網羅的研究を推し進めるには、進行状況などの情報を正確かつリアルタイムに共有してゆく必要がある。また、研究過程で生産される膨大な実験データの蓄積はこれまでに得られなかった多面的な情報を含んでおり、それ自体として価値がある。複数のデータベースを統合した研究情報管理・利用システムとして HTPF-DB の設計開発を日立ソフトウェアエンジニアリング㈱との共同研究により行ってきた。HTPF では、各グループ間でデータ共有、交換や情報の一元管理のために、そのレポートのフォーマットを統一化している。タンパク質は各々に性質が大きく異なり、個別に最適な実験方法や条件を選択する必要があるため、画一的なデータベース化が非常に困難であるので、データベースには変更拡張に柔軟に対応できる画面インターフェースの自動生成と木構造による動的モデルとモデルからのユーザインタフェースの自動構築を採用した。HTPF-DB は、基礎 DB、プラスミド DB、培養 DB、精製 DB、サンプル DB、自動結晶化観察ロボットシステムと連携した結晶 DB、構造ゲノムビームラインからの回折 DB、解析 DB の 8 つから成り立っており、DB 利用ツールとして mmCIF 登録ツールがある。

9. おわりに

これまで述べたように、蛋白質結晶構造解析の高速化のためには、構造解析のそのものの効率化に加え、得られた立体構造を評価するための機能解析環境も効率化する必要がある。今後、横浜研究所や内外の研究機関、企業と協力して、応用上重要なヒト、マウスなどの高等生物由来のタンパク質の 3 次元構造解析によって生命現象を理解する基礎的情報を得て、創薬など応用分野への貢献を視野に入れて研究を進めていく方向である。そのために必要な装置や技術の開発や理研のもつ情報を企業が利用できるために必要な制度として、理化学研究所では融合的産業界連携制度 (<http://www.riken.go.jp/r-world/info/release/press/2004/040216/index.html>) や理研パートナー制度 (<http://www.riken.go.jp/r-world/info/release/press/2004/040216/index.html>) や理研パートナー制度 (<http://www.riken.go.jp/r-world/info/release/press/2004/040216/index.html>)

//www.rsgi.riken.go.jp/partner/index.html)等の法整備も進められているところである。現在ハイスルーブットファクトリーでは年間50構造を解析できるが、これら効率化のための技術開発を通じて、より経験の浅い人員で多くの構造解析ができるようにしていきたい。2007年にタンパク3000プロジェクトが終了した時には構造解析の専門家でない研究者が手軽に蛋白質結晶構造解析を行えるようになっているようお願いしつつ我々は作業を進めている。

謝辞

本稿に記述した高度好熱菌由来サンプルの発現ベクターは倉光成紀、横山茂之両プロジェクトディレクターをはじめとするストラクチュローム研究グループから供与され、超好熱古細菌由来サンプルの発現ベクターは林崎良英グループディレクターの横浜GSC遺伝子構造・機能研究グループにより作成された。理研構造ゲノムビームラインの建設および運用はX線干涉光学研究室の山本雅貴副主任研究員、石川哲也主任研究員によってなされたものである。試料の調製および分析、構造解析は、飯塚哲太郎播磨研所長、宮野雅司ファクトリー長のもとでハイスルーブットファクトリーのタヒールタヒロフ、湯谷克英の各チームリーダー、瀧尾擴土上級研究員をはじめとするメンバーによってなされたものである。また、装置の開発と実際の運用はHTPF、理研播磨研究所、大型放射光施設SPring-8の研究、支援スタッフにより行なわれた。TERAの開発には竹田理化学工業、エステック、アドバンソフト、迅速X線蛋白質結晶評価装置の開発および構造ゲノムビームラインの運営にはリガク、HTPF-DBの開発には日立ソフトウェアエンジニアリングの企業各社との共同研究により進めてきた。ここに述べた多くの方々をはじめSPring-8および理研の関係者の方々に深く感謝の意を表したい。

参考文献

- 1) R. F. Service: Structural genomics. Tapping DNA for structures produces a trickle, *Science* **298**, 948 (2002).
- 2) 山本雅貴, 後藤俊治, 竹下邦和, 石川哲也: 構造ゲノムビームライン (BL26B1/B2) の計画, SPring-8 利用者情報 **6**, 202 (2001).
- 3) 宮野雅司, 菅原光明, 山本雅貴: タンパク質の大規模結晶構造解析または, ハイスルーブットということ, 日本結晶学会誌 **45**, 267 (2003).
- 4) S. Yokoyama, H. Hirota, T. Kigawa, T. Yabuki, M. Shirouzu, T. Terada, Y. Ito, Y. Matsuo, Y. Kuroda, Y. Nishimura, Y. Kyogoku, K. Miki, R. Masui and S. Kuramitsu: Structural genomics projects in Japan, *Nat Struct Biol* **7**, 943 (2000).
- 5) S. A. Lesley et al.: Structural genomics of the *Thermotoga maritima* proteome implemented in a high-throughput structure determination pipeline, *Proc. Natl. Acad. Sci. USA* **99**, 11664 (2002).
- 6) T. Kigawa and S. Yokoyama: A continuous cell-free protein synthesis system for coupled transcription-translation. *J Biochem (Tokyo)* **110**, 166 (1991).
- 7) J. A. Sigrell, P. Eklund, M. Galin, I. Hedkvist, P. Liljedahl, C. M. Johansson, T. Pless and K. Torstenson: Automated multi-dimensional purification of tagged proteins. *J Struct Funct Genomics* **4**, 109 (2003).
- 8) 菅原光明, 宮野雅司: ハイスルーブット自動結晶化・観察システムの開発, 蛋白質核酸酵素 **47**, 1026 (2002).
- 9) N. E. Chayen, P. D. Shaw Stewart, D. L. Maeder and D. M. Blow: An Automated System for Micro-batch Protein Crystallization and Screening, *J. Appl. Crystallogr.* **23**, 297 (1990).
- 10) J. Jancarik and S. H. Kim: Sparse matrix sampling: a screening method for crystallization of proteins, *J. Appl. Crystallogr.* **24**, 409 (1991).
- 11) T. C. Terwilliger: Automated main-chain model-building by template-matching and iterative fragment extension, *Acta Cryst. D* **59**, 34 (2002).
- 12) A. Ferrandez, B. Minambres, B. Garcia, E. R. Olivera, J. M. Luengo, J. L. Garcia and E. Diaz: Catabolism of phenylacetic acid in *Escherichia coli*. Characterization of a new aerobic hybrid pathway, *J. Biol. Chem.* **273**, 25974 (1998).
- 13) V. S. Lamzin, A. Perrakis and K. S. Wilson: *International Tables for Crystallography Vol. F* (Kluwer Academic Publishers, Dordrecht, Netherlands, 2001) 720-722.
- 14) T. Kawabata and K. Nishikawa: Protein structure comparison using the markov transition model of evolution, *Proteins* **41**, 108 (2000).

著者紹介



国島直樹

理化学研究所播磨研究所ハイスルーブットファクトリー構造解析第一チーム
E-mail: kunisima@spring8.or.jp
専門: X線結晶構造解析

略歴:

1994年 大阪大学大学院博士課程修了 博士(理学)
2001年まで 榊生物分子工学研究所 研究員
現在まで 理化学研究所ハイスルーブットファクトリー
チームリーダー



菅原光明

理化学研究所播磨研究所 構造生物学
理研究室・ハイスルーブットファクトリー結晶化チーム
E-mail: sugah@spring8.or.jp
専門: 構造生物学・X線結晶構造解析

略歴:

2000年 大阪大学大学院博士課程修了 博士(理学)
2000年 理化学研究所播磨研究所 基礎科学特別研究員
2002年 理化学研究所播磨研究所 研究員
2003年 理化学研究所ハイスルーブットファクトリー
チームリーダー(兼任) 現在に至る

The Efforts toward High-throughput Protein X-ray Crystallography in Highthroughput Factory, SPring-8/RIKEN Harima institute

Naoki KUNISHIMA Highthroughput Factory, RIKEN/SPring-8,
1-1-1 Koto, Mikazuki-cho, Sayo-gun, Hyogo 679-5148, JAPAN
Mitsuaki SUGAHARA Highthroughput Factory, RIKEN/SPring-8,
1-1-1 Koto, Mikazuki-cho, Sayo-gun, Hyogo 679-5148, JAPAN

Abstract

Although genome sequences of over 150 bio-organisms including human-being are being rapidly determined, most of their genome functions are unknown. The goal of structural genomics is to bring out the genome functions by analyzing cyclopedically and systematically all of the protein basic structures. Now a lot of people are making every effort to solve all structures and functions of the existing proteins as gene product by every means imaginable. Here we introduce some trials including technological development toward the high-throughput protein X-ray crystallography in our Highthroughput Factory, Spring-8/RIKEN Harima institute.