

# タンパク3000プロジェクト

文部科学省主導によるRR2002プロジェクトの一環として、『我が国初のゲノム創薬の実現等を目指し、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プロジェクトが日本のタンパク科学を結集する形ではじめられた（図1）<sup>[1,2]</sup>。そして、構造ゲノム科学が構造プロテオミクス、機能プロテオミクスへとその研究対象を大きく広げた。



図1 タンパク3000プロジェクト(5)

ここでは、タンパク3000の網羅的解析の委託を受けた理化学研究所での、横山茂之を中心とする理研プロテオミクス研究推進（RSGI: RIKEN Structural Genomics/Proteomics Initiative）における成果のうちSPring-8を使って進んでいる部分に絞って報告する。

## SPring-8での技術開発と網羅的解析

### 1. 関連技術開発とハイスループットファクトリー

SPring-8でのビームラインまでを含めた統合的な自動回折データ収集システムの整備の一環として、まず、構造ゲノムビームラインBL26B1サンプルチェンジャーの高度化及び実験室X線装置用サンプルチェンジャーの試作を行い、自動回折強度データ収集システムの構築を進めた。自動回折強度データ収集システム用に小型の専用サンプルピンを新たに開発した。専用サンプルピンは、簡単な構成のサンプルチェンジャーにより、単純な左または右回転により確実にサンプルの

脱着を可能にし、専用サンプルトレーに52サンプルを同時に収納した状態で保存・運搬・持ち込みを行い、大量のサンプルについての管理・自動測定を効率化する<sup>[3]</sup>。放射光ビームライン用サンプルチェンジャーは基本動作テストを終了して、現在連続試験運転を行っている。実験室X線装置用サンプルチェンジャーは試作機による基本動作テストを実施し、膨大なデータのバックアップなどセキュリティの高い運用を目指してHTPF-DBとの連携テストも併せて行っている。

また、将来に向けてのR&Dとして、微小結晶の結晶性の評価と構造解析の実現可能性の検索を目標にマイクロビームと組み合わせて使用する微小結晶構造解析テストベンチの設計・製作を進めている。微小結晶構造解析のテストベンチでは、対象サンプルサイズを数十ミクロン以下に想定して回転精度サブミクロンをめざした精密ゴニオメータを備えた微小結晶用回折計のプロトタイプ製作と精度検定システムの導入を行った。芯ずれが現状の1/10であるミクロン精度の高精度回転ステージの設計・試作を行った。試作機では、芯ずれが最大値で25ミクロン以下を達成した。また、S/N比の悪い微弱な回折強度データの収集を最も高S/N比で測定可能であると予想される微小振動写真法に向けて、現在保有するCCD検出器の高速化および回折像処理ソフトウェアのR&Dを実施した。

昨年度試作で導入した完全自動結晶化観察ロボット「TERA」を商用バージョンアップしたシステムとしてさらに2セット増強し、さらに得られた結晶の迅速結晶評価を行うため、同じく昨年度試作したプレート上回折評価システムを使って、放射光タンパク質結晶構造解析を進める上でキーとなる結晶化を効率的かつ効果的に利用できる環境構築を行った。またこれらの実験情報を開発したハイスループットデータベースHTPF-DBシステムに組み込んで、発現精製、結晶構造解析の要として全体統合を図った<sup>[4,5]</sup>。さらに、結晶化の効率化、確実性をはかるためにセミマイクロ極微量結晶化分注装置の試作を行った。

結晶解析計算の自動化、簡便化のため、既存の網羅的分子置換法プログラムEPMRのグラフィカルユーザーインターフェースを作成し、各項目にはデフォルト値を表示した。また、このプログラムを実行するために既存の全ての立体構造のうち近似構造を除いた異なる立体構造のデータ（現在5230個）を用意した。入力画面に必要な事項をインプットするだけで、全ての既知構造を対象にターゲット蛋白質をサーチすることができる。

放射光施設SPring-8で、タンパク質発現プラスミド管理か

ら、解析された立体構造の機能解析データまで、すでに300万枚を超えた結晶写真イメージ、構造ゲノム自動化ビームラインではき出される1MADデータセットが数ギガバイトという莫大な回折イメージデータなど、研究実験データが日々大量に生産されている。これらの複雑多岐にわたり相互に依存した放射光蛋白質結晶構造解析を高効率で推進していくため、発現・培養・精製・結晶化・回折・解析の各段階において大量に発生するの実験データの統合研究情報管理データベースHTPF-DBの基本機能の設計、開発をした<sup>[6]</sup>。

理研・播磨研究所では、SPRING-8キャンパスに設置されたハイスループットファクトリー（HTPF）において、これらの技術の実践実証、統合化による効率化を目指している（図2）。真核生物により近い古細菌である超好熱菌*Pyrococcus horikoshii* OT3の網羅的解析についてもすすめており、1300のクローニングを行った。

## 2. ストラクチュローム研究

もっとも早くから進んでいる網羅的構造ゲノムプロジェクトとして、高度好熱菌 *Thermus thermophilus* HB8を用いた研究が進んでいる。

タンパク質を始めとする生体分子の立体構造を基にして、一つの細胞全体の生命現象を、原子レベル（物理化学的レベル）で理解することを目指しているのが「高度好熱菌 丸ごと一匹 プロジェクト」である。その目的に最も適したモデル生物として、(1) 遺伝子操作系が存在する生物の中で、(2) 最も高温で生息する、という2つの条件を満たす高度好熱菌

*Thermus thermophilus* HB8 を選んだ。その最終目標を達成するまでに、研究は4段階で進行する(<http://www.thermus.org>)。第1段階は、ゲノムワイドにタンパク質の立体構造を決めるステップであり、第2段階は、第1段階で得られたタンパク質の立体構造を利用する他、トランスクリプトーム解析、プロテオーム解析、メタボローム解析なども駆使する。第3段階は、ゲノムワイドな研究が始まる遥か以前から、各研究者によって行われてきた“各論的”分子機能解析を含む。第4段階は、第3段階までの情報を統合し、細胞全体を原子レベルで理解するために、シミュレーション結果と比較しつつ研究を進めることになる（システム生物学）

このうち、第1段階を、タンパク3000プロジェクトと協調して進めている。2,200種類存在する全タンパク質のうち、2003年6月末時点で、約1600種類のタンパク質量産化用プラスミドを作製し、約980種類のタンパク質量産化に成功している<sup>[6]</sup>。原則として、当該タンパク質について実績があり、論文作成や特許申請を十分に行うことのできる研究室が構造・機能解析を担当している。各チームが担当するタンパク質は、ストラクチュローム研究グループの遺伝情報チームが約400種類、代謝系チームが約400種類、機能未知チームが約500種類、その他に、大学や研究所の研究者が担当するタンパク質が約300種類、そして、HTPFが担当するタンパク質が約300種類となっている（図3）。

このようにして、これまでに結晶化まで進んだタンパク質は330種類となり、X線結晶解析が完了したタンパク質は100種類を越えた。これによって、1974年に高度好熱菌 *Thermus*

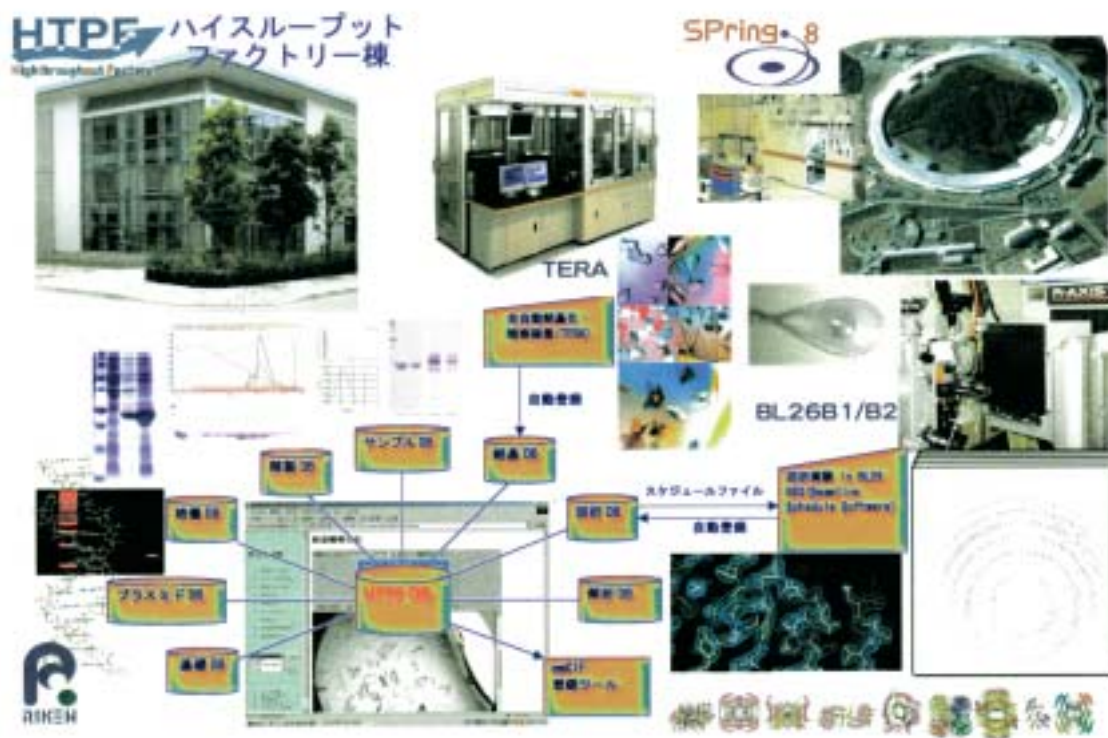


図2 ハイスループットファクトリー



*thermophilus* HB8 が発見されて以来、合計150種類以上のタンパク質の立体構造解析が、国内外の研究者の協力によって成し遂げられたことになる。



図3 ストラクチュローム研究グループ

### 3. 具体的構造解析成果の例 (図4)

微生物由来のタンパク質については、平成14年度中に複合体などを含め71個の立体構造をX線結晶構造解析により決定した。そのうちのいくつかについて具体的に説明する。

フェニル酢酸分解経路は*paal*遺伝子群と呼ばれる一連の蛋白質によって構成されるが、これらのうちPaal蛋白質の機能は不明である。TT0310、Paal蛋白質の結晶構造をMAD法により1.7 分解能で決定した (PDB code: 1J1Y)。構造既知蛋白質との構造比較、部位特異的変異体の作成およびその酵素学的解析を行った結果、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、活性残基Asp48が同定された。また、予想に反して活性が低下したN33D変異体の結晶構造を決定したところ、活性低下の理由が活性部位における水素結合様式の変化であることが明らかとなった。次に、リガンド (コエンザイムA誘導体) との複合体結晶の構造解析を行った。ヘキサノイルCoAとの複合体は酵素-基質複合体の構造を、CoA-SHとの複合体は

酵素-生成物の構造を示している。一方、リガンドなしの状態では3つの異なる結晶型が得られた。本酵素はホモ四量体構造をとっているが、I4<sub>2</sub>2<sub>2</sub>型では非対称単位に1サブユニット、P4<sub>3</sub>2<sub>2</sub>型では2サブユニット、さらにC22<sub>2</sub>型では8サブユニット (四量体2個) を含んでいた。これらの結果から、本酵素はリガンドなしの状態では様々な形を取り得るが、リガンドの結合により別の形に固定されることが示された。TT0499は、 $\alpha$ -グルコシダーゼはグリコシルヒドロラーゼファミリー1に属する糖鎖の加水分解酵素であり、分子置換法により0.99 分解能で決定した (PDB code: 1UAH)。TT0127、プリンヌクレオチドホスホリラーゼ (PNP) は核酸合成において重要な、ヌクレオチドホスホリラーゼ1ファミリーに属する酵素である。PNPの障害剤はT細胞白血病の治療薬として有効であることから、薬剤設計のターゲットとして重要である。これまでに2種類の細菌由来の酵素について結晶構造が報告されている。TT0127の結晶構造を分子置換法により1.9 分解能で決定した (PDB code: 1ODK)。さらに、構造機能相関を明らかにするため、アデノシンおよび硫酸イオンとの複合体 (PDB code: 1ODJ)、グアノシンおよび硫酸イオンとの複合体 (PDB code: 1ODI) の結晶構造を決定した。これらの複合体の構造から、PNP障害剤設計に有用な塩基認識に関する情報が得られた。TT0368、3-ヒドロキシイソブチル酸 (HIBA) はバリン代謝の重要な中間代謝産物であり、NADP依存性3-ヒドロキシイソブチル酸脱水素酵素 (HIBADH) によりメチルマロン酸セミアルデヒドに酸化される。HIBADHの結晶構造 (PDB code: 1J3V) を初めてMAD法により1.8 分解能で決定した。

高度好熱菌由来の機能未知タンパク質TT1542について、SeMet 化タンパク質の結晶構造をMAD法を用いて、2.0 分解能で決定した。このタンパク質のホモログは、原核生物から真核生物に至るまで広範囲にわたり存在しており、真核生物におけるホモログについてはGlcNAc-PI de-N-acetylase 活性を持つことが示唆されている。しかし、原核生物におけるホモログについてはそれらの機能は未だ知られていない。構

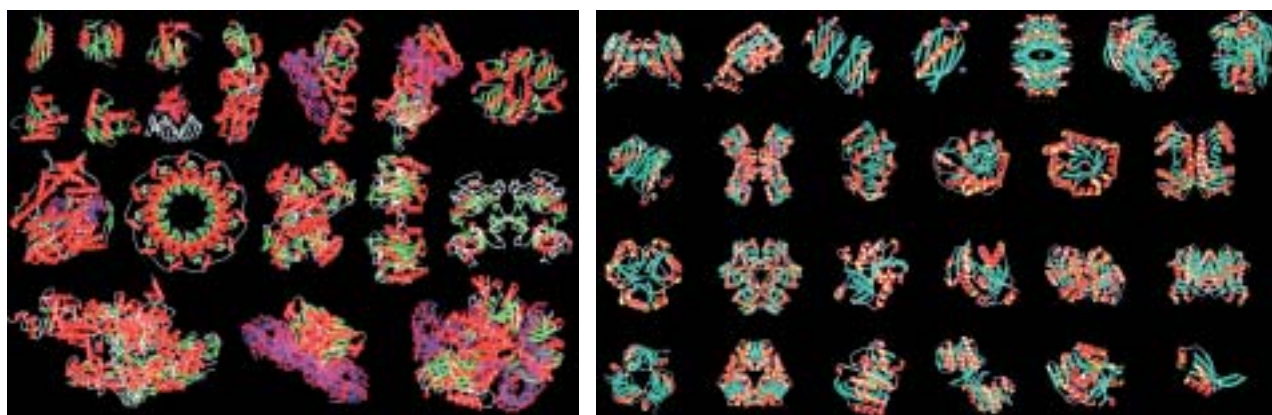


図4 構造決定されたタンパク質の例

造解析の結果、このタンパク質は N 端側に Rossmann-fold に似たドメインを持つ、糖鎖のプロセッシングに関与する酵素である可能性が示された。また、他にも機能未知タンパク質 TT1751 について、SeMet 化タンパク質の結晶構造を MAD 法を用いて、2.0 Å 分解能で決定した。このタンパク質は、Pfam データベースでは DUF302 (Domain of Unknown Function 302) に属している。このグループに属するタンパク質の立体構造を明らかにしたのは、本タンパク質が初めてである。さらに、機能未知タンパク質 TT1725 については、大腸菌無細胞タンパク質合成系で合成した SeMet 化タンパク質の結晶構造を MAD 法を用いて 1.7 Å 分解能で決定した。

また、T7 ファージ由来の RNA ポリメラーゼの伸長複合体 (転写中間体) の結晶構造を解明した。T7 RNA ポリメラーゼとヘテロデュプレックスの複合体を形成させることによって、伸長複合体の構造を捉えることに成功し、この構造が真核生物の RNA ポリメラーゼに共通することが明らかとなった。

tRNA のプロセッシングを触媒する酵素に関しては、tRNA 前駆体の 3' 末端側の余計な配列をトリミングする RNase PH (Aquifex aeolicus 由来) の結晶構造を 2.6 Å 分解能で決定し、tRNA 前駆体のトリミング機構を提唱した。また、tRNA の修飾に関わる酵素では、古細菌に特異的なアーケオシン tRNA グアニントランスグリコシラーゼ (ArcTGT) 単独の結晶構造を 2.2 Å 分解能で、さらに tRNA との複合体の結晶構造を 3.3 Å 分解能で決定し、修飾酵素が tRNA の 3 次構造を大きく変形し修飾を導入する機構を世界で初めて明らかにした。

大腸菌のタンパク質としては、DNA 複製開始を正に制御する DnaA タンパク質の DNA 結合ドメインと DnaA ボックス DNA との複合体の立体構造を、X 線結晶構造解析法により 2.1 Å 分解能で決定し、複製開始調節タンパク質による認識機構などを明らかにした。

高等生物由来のタンパク質としては、DNA 組換え反応に関わる酵素などについて X 線結晶構造解析を行い、計 4 種類の立体構造を決定した。

#### 参考文献

- [ 1 ] 文部科学省：タンパク 3000 プロジェクト発足記念シンポジウム要旨集 2002 年 8 月 19 日 理研・横浜研究所.
- [ 2 ] 宮野雅司：利用者情報 **8** (2003) 78-88.
- [ 3 ] K. Ida, M. Sugahara, M. Yamamoto and M. Miyano : SPring-8 Research Frontiers 2001/2002 in press (2003).
- [ 4 ] 菅原光明、宮野雅司：現代化学 **41** (2003) 383.
- [ 5 ] 宮野雅司、高秀幸：現代化学 **43** (2003) 384.
- [ 6 ] 大谷直人他：蛋白質核酸酵素 **47** (2002) 1009-1013.

理研 播磨研究所  
 ハイスルーフトファクトリー  
 宮野 雅司

理研 播磨研究所  
 ストラクチュローム研究グループ  
 倉光 成紀

理研 横浜研究所  
 ゲノム科学総合研究センター  
 タンパク質構造・機能研究グループ  
 横山 茂之