


タンパク3000プロジェクト - 野心的なポストゲノム研究のはじまり -

理化学研究所 播磨研究所 ハイスルーフトファクトリー
構造生物物理研究室 宮野 雅司

生物の中でゲノムの情報が翻訳され、生き物が生きていく中で働く生物ナノロボットであるタンパク質の研究の中で、タンパク質の立体構造を網羅的に明らかにしようという文部科学省によるタンパク3000 (National Project on Protein Structural and Functional Analyses)、そして米国NIHによるPSI (Protein Structure Initiative) をはじめとした構造ゲノム科学がすでに世界で始まっている。

この流れは、これまで個別の対象を個々の研究として行われてきたライフサイエンス研究を網羅的に行うという新しい大きな流れであり、ゲノム (Gene+ome: 遺伝子の世界)、トランスクリプトーム (Transcript+ome: 転写産物世界)、プロテオーム (Protein+ome: タンパク質世界) という新たな生物全体としての網羅的セントラルドグマ

DNA: ゲノム: ゲノミクス  (複製)
(転写)
RNA: トランスクリプトーム: トランスクリプトミクス
(翻訳)
タンパク質: プロテオーム: プロテオミクス
(化学的実体: その集合的世界: その技術)

の研究の広がりとなっている。

核酸であるゲノムとトランスクリプトームがプロテオームとは研究において大いに違っている。まず、化学的基礎は、DNA、RNAなどの核酸は糖がリン酸エステルによってつながったひもの上に結合したACTG (RNAではTの代わりにU) の4種類の塩基の並びによってその情報が直線的にコードされているのに対して、タンパク質をコードするゲノム全体の数パーセント以下の遺伝子DNA配列の3つ並び塩基、三つ組み (トリプレット: triplet) が一つのアミノ酸に対応して巨大な翻訳ロボットシステム、リ

ボソームで脱水縮合されてアミノ酸が1本のポリペプチドになる。それだけでは、タンパク質とならず、きわめて性質の異なる側鎖をもつ20種のアミノ酸の配列によって、ヘリックス シートといった構造ユニットをもつある決められた立体構造に折り畳まれる。そこではじめて“働く形”を持った生物ナノロボット、タンパク質として機能する。

つまり、タンパク質情報を担った遺伝子が生物細胞の水溶液環境で、時にはいろいろな手助けを借りながら、その働くための立体構造という情報飛躍の上に成り立っている。この1次元ポリペプチド配列から3次元タンパク質立体構造への情報ギャップは、コンピューター技術の発達した今でもX線結晶構造解析、NMRという実験的方法によって立体構造決定したモデルとなる既知の類似タンパク質の構造なしには越えることができない。翻訳され、折り畳まれたタンパク質は生体の中でさらにペプチドが特定の場所で切断されたり、時には新たにつなぎ直されたり、そして糖、脂質、リン酸、硫酸、メチルなど多くの翻訳後修飾されて調節を受け、この生物ナノロボットは、そのときに応じて正しく制御され生命機能を担う。時には、リボソームのようにタンパク質の代わりにRNAがその主要構成要素になることがある。

タンパク質の立体構造が、狂ってしまうだけでヒトは病気にさえなることが明らかになってきている。特に、タンパク質プリオンだけで起こるクロイツフェルツヤコブ病はBSE (狂牛病) のように種を越えて感染して脅威になっている。アミロイドの蓄積によるといわれるアルツハイマー病などタンパク質の線維化変性による病気はもちろん、これ以外の多くの病気ではタンパク質の働き過ぎ、機能低下などにより病気が進行することが知られており、乳ガンのハーセプチン、慢性骨髄白血病のグリベックな

どの新しいガン治療薬は、働きすぎたために異常に増殖するようになったガン細胞のタンパク質を明らかにした上でそのタンパク質機能を抑えることで働く。この新しいアプローチを広くゲノム創薬と呼ぶことがある。しかし、これまでのクスリと同様魔法の弾丸でもないことは同じような働きによる肺がん治療薬イレッサの新聞報道でも明らかである。

プロテオミクス研究はこうしたタンパク質の動態を網羅的に研究する手段を提供してくれるし、構造ゲノム科学はゲノム決定の結果予測された3割から5割もの機能のわからないタンパク質機能を推定するもっとも有力な手段となり得るので機能プロテオミクスとしても重要である。また、すでに機能がそのアミノ酸配列から予想されていても、原子レベルで立体構造がわかることで実験をデザインすることが可能になり、さらなる機能研究が大きく広がる。今やタンパク質立体構造決定はタンパク質研究のランドマーク的手法の一つである。また、遺伝子がコードするタンパク質のアミノ酸配列から機能が推定されているといっても、過半はそのタンパク質の「分子機能」が予測できるだけであり、その生物の中での具体的働き、「生物機能」については未だ多くは不明であり、タンパク質の生物機能解析にはこれまでと同様さらに膨大な地道な生物的研究が必要である。このことは、医学分野での応用研究を志向するときには特に注意を要する。全ゲノム配列が決定できて初めて網羅的研究が可能になり、これまで個別的、断片的、分析的になされてきた遺伝子・タンパク質による生物研究を、全体像として理解するための基盤の第1歩としてタンパク質の構造に関する博物学的アプローチがはじまったといえるかもしれない。

国際的なフレームからの構造ゲノム科学

構造ゲノム科学は、ポストゲノムの最も重要な分野の一つとして現在国際的競争と協調のもと始まり、進行している。第1回の横浜に続き、昨年10月ベルリンで、国際的な構造ゲノム科学組織ISGO (International Structural Genomics Organization) により開かれた第2回構造ゲノム科学国際会議 (ICSG2002) がその大きな流れを総括している^[1]。第1回である横浜の会議ICSG2000の時にSPring-8で開いたように^[2]、無細胞システムを使ったタンパク質生産、ハイスループットX線結晶構造解析のサテライト・ワークショップがICSG2002でも開かれたことが、この構造ゲノム科学の現在のもう一面を

現している。これからも、3月に英国ヒンクストンでHTPX (High-Throughput X-ray Crystallographic) データベースの、またNIH/PSIの主催で4月にタンパク質生産と結晶化のワークショップが米国で予定され、ハイスループットでのタンパク質立体構造決定に関わるあらゆる研究技術開発のための時宜を得た情報交換が急務とされている。つまり、基本的技術はすでにあるがその技術を網羅的にハイスループット技術として確立するには多くの技術的課題がある。米国で構造ゲノム科学の最大の資金援助をしているNIH/PSIの9つのプロジェクトはあくまでもパイロットプロジェクトとして位置づけられ、次の本格的センターのための技術確立をその主たる目的としていることから明らかである。

構造ゲノム科学の大きな特徴は、はじめからその応用が強く認識され、成果の知的財産権 (IPR: Intellectual Property Right) が議論されてきている。そして、すでに、米国Syrrix、SGX、英国のAstexなど構造ゲノム科学を基盤とするベンチャーが自らのための放射光ビームライン建設など大規模に活動している。遺伝子に連なる特許の問題は、国際的に注目されている。直接、構造ゲノム科学に関わるIPRとしてタンパク質の特許化の意見交換が日米欧三極でなされてきているが、その中で米国の主張するソフトウェア基準を適用し、座標データはあくまでそのためのコンテンツであるとする考えが強く、行政的取り扱いをする日本の特許庁もこの線の方向性を主張している^[3]。これは米国で主たる公的資金源であるNIH/PSIの構造ゲノム科学の援助を受けたものはタンパク質の構造決定と同時に即時公開すべきであるとのポリシーに添ったものである。日本政府のタンパク3000などの構造ゲノム科学資金がその権利化を強く要請している中で、タンパク質立体構造に基づいてさらに進んだ機能解析などなしにはIPRの確保は困難であると予想される。今後どのような形でタンパク質の立体構造のIPRが認められるかは注目すべき点である。特に、米国の特許は議論のある特許申請は先ず認めて、最終的に法廷決着するのがふつうであることを考えると予断を許さない。実際、遺伝子特許における流れは研究阻害の一面も指摘されており^[4]、その下流にある構造ゲノム科学においては、特に研究を進める基本リソースである遺伝子cDNAの流通を含めたIPRの国際協調なしには困難である。

国際協調のフレームとしてNPOであるISGO (“イ

スゴ”と発音する)がある。準備委員会の幹事であった横山茂之、Tom TerwilligerそしてUdo Heinemanの3人に加えてベルリンの会議で倉光成紀、Ian Wilson、David Stuartの3人が追加された。この中で国際的なガイドラインが決められつつある(<http://www.nigms.nih.gov/news/meetings/airlie.html>)。座標については、時宜を得た(構造決定後3週間以内)公開をPDB(日本のPDBサイトは大阪大学蛋白質研究所にあるPDBj: Protein Data Bank Japan; <http://pdbj.protein.osaka-u.ac.jp/>)に寄託することになっており、特にIPR確保が必要であるような重要タンパク質については構造決定後6ヶ月を越えないうちに公開することになった。このとき同時に、構造決定研究に使ったX線ならば構造因子データなどの付随的情報も出来る限り同時に出すべきであるとしている。この間に必要なIPRの確保、またそのための機能解析を行うことになっている。そして、進行中の研究進捗はターゲットから構造決定、PDB寄託までを決められた形でインターネット上に公表することになっている。

タンパク質構造決定の科学的側面、そしてISGOのポリシーで言及すべき点が2点ある。一つは、2000年以降のいくつかの構造ゲノム科学のポリシーに関わる国際会議を経て大きく変わった点であり、それはタンパク質の立体構造の終了は研究者自身が決定するという点であり、その質はそのときの学問的レベルとタンパク質の原子座標の持つ影響の大きさを考慮して最大限その質を確保したうえで、決定されるべきであるとした。初期のバイオインフォマティクス関係者が議論の流れを作っていたときには、タンパク質のフォールドが決まればよいとしていたことと大いに異なる。

もう一点は、その研究範囲の広がりである。初期にあったこのプロジェクトに対する批判に対応した結果であり、特に現時点でハイスループットな立体構造解析がほとんど現実でない超分子複合体、膜タンパク質もターゲットから除外されなかったことである。タンパク質構造研究を進める上で必要となるタンパク質機

能研究全般も“立体構造を決定するため”、“IPRを確保できるため”と限定されてはいても、ターゲット選択、ターゲット優先度決定のためのバイオインフォマティクスはもちろん、タンパク質発現のためには、実験的プロテオミクスなしには本当に生体内で発現しているタンパク質の確定も十分でないことはライフサイエンス研究の急進展の中でますます明らかとなっている。

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、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プロジェクトが日本のタンパク科学を結集する形ではじめられた(図1)^[5]。そして、構造ゲノム科学がその視野を大きく広げたように、日本での構造ゲノムプロジェクトはゲノム配列を基礎として網羅的タンパク質立体構造解析をするばかりでなく、より生物的視点を重

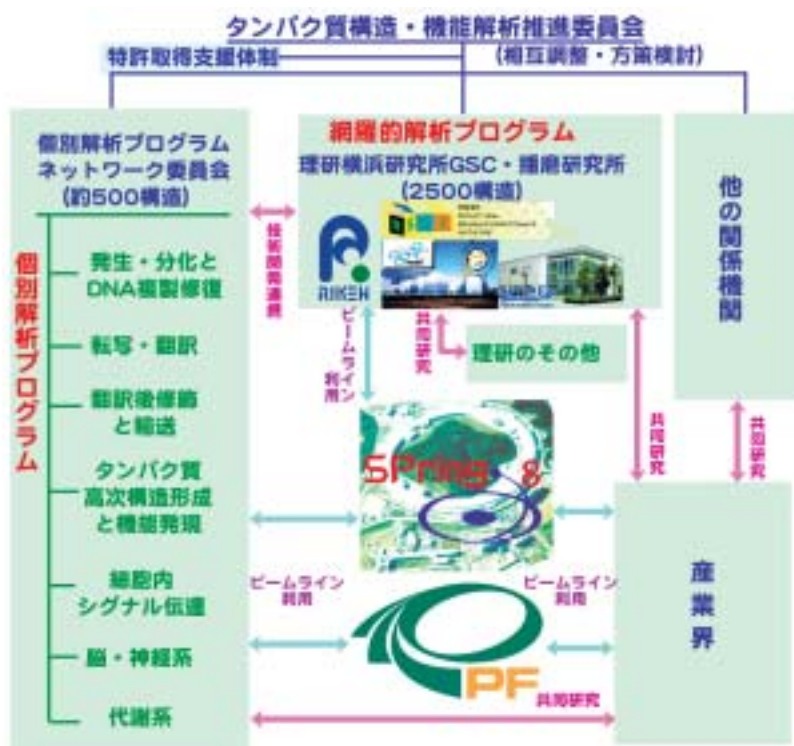


図1 タンパク3000プロジェクト^[5]

視した個別プログラム、さらには幅広くそのターゲットをおき、基礎生物学の基盤としてばかりでなく、その産業応用を展望して、可能なIPRの確保も重要な課題となっている。

プロジェクトには大きく2つのプログラムがあり、
 ・タンパク質基本構造の網羅的解析プログラム
 ・タンパク質の個別的解析プログラム
 に分かれている。

の網羅的解析は理化学研究所が、の個別的解析は7つの大学を中心としたグループが担当している。

7つの個別プログラム（カッコ内：中核機関、実施代表者）は、

1. 発生・分化とDNA複製・修復（東京大学大学院農学生命科学研究科、田之倉 優）
2. 転写・翻訳（北海道大学大学院理学研究科、田中 勲および横浜市立大学、西村善文）
3. 翻訳後修飾と細胞内輸送（高エネルギー加速器機構、若槻壮市）
4. タンパク質高次構造形成と機能発現（京都大学大学院理学研究科、三木邦夫）
5. 細胞内シグナル伝達（北海道大学大学院薬学研究科、稲垣冬彦）
6. 脳・神経系（大阪大学たんぱく質研究所プロテオミクス総合研究センター、中川敦史）
7. 代謝系（大阪大学大学院理学研究科、倉光成紀）となっている。

タンパク質3000の網羅的解析の委託を受けた理化学研究所では、その当初よりこの構造ゲノム科学を先導してきたタンパク質グループを率いる横山茂之を中心として理研プロテオミクス研究推進（RSGI: RIKEN Structural Genomics/Proteomics Initiative）が、本部長のもとにNMRによるタンパク質立体構造決定を目指す横浜研究所ゲノム科学センター（GSC）そして放射光によるハイスループットX線結晶構造解析を目指すSPring-8キャンパスにある播磨研究所により進められている。

これ以外の日本での構造ゲノム科学プロジェクトとしては、経済産業省による生物情報解析研究センター（JBIRC: Japan Bio-Informatics Research Center）がより困難な膜タンパク質をターゲットとして始まっている。また、農水省ではイネゲノム研究の一環として農業生物資源研究所で進められている。

SPring-8での構造ゲノム科学：

ハイスループットファクトリー

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）、そして初期から構造ゲノム科学を高度高熱菌由来タンパク質を網羅的に解析しようとしている倉光成紀によるストラクチュローム、そして協力研究室群、さらに側面からビームライン建設などファシリティー研究開発・建設を支えてくれる物理グループと高輝度光科学研究センター（JASRI）など多くのグループが関わってその極めて挑戦的なプロジェクトに取り組んでいる。

ハイスループットファクトリーに限って言えば、必要なハイスループット技術開発を行うとともに、その技術を実践実証するハイスループットファクトリーがある（図2）。これまでに、全自動結晶化観察ロボットシステムTERAを協力企業とともに開発し、その実際の統合的利用システムが確立され、1年ほど順調な稼働を果たしてきている^[6]。統合化には膨大な結晶化写真を蓄積管理するとともに、結晶評価できる「結晶化日記」を含む放射光結晶構造解析統合研究情報管理システムであるハイスループットファクトリー・データベース(HTPF-DB)をやはり企業との共同研究で開発して利用している^[7]。また、結晶の評価を結晶化プレート上でそのまま出来る「その場結晶評価装置」を別の企業と共同研究で開発試作して試験を始めている。SPring-8での中心的課題のひとつであるビームラインの自動化・高度化は理研構造ゲノムビームライン（BL26B1/B2）において山本雅貴を中心としたグループによって順調に開発が進んでいる^[8]。ここでも発生する膨大な量の回折イメージを先ほどのHTPF-DBに統合した利用計画を進めている。ハイスループットファクトリーの活動はISGOに準拠できるように順次努力しており、RSGIへリンクしたホームページを開設している。ターゲットは現在、ストラクチュロームから提供されたプラスミドを利用した真性細菌である高度好熱菌 *Thermus thermophilus* HB8のタンパク質である。その発現精製のしやすさ、結晶化率の高さなど、例外的に構造ゲノム科学プロジェクト向きの優れたターゲットを使って結晶化ロボットなど必要な技術開発を行ってきた。今後は、真核生物とより近い古細菌である超好熱菌 *Pyrococcus horikoshii* OT3の網羅的解析を目指して、理研・和光本所の林崎グループと遺伝子クローニングなどに



図2 ハイスループットファクトリーとそのための技術開発^[6,7]

において協力している。機能解析は外部と協力できることは出来るだけ一緒に進められるよう努力している。これからはじまる超好熱菌そしてより困難が予想されるマウスなど哺乳動物由来のタンパク質解析に向けたプロテオミクス研究も可能な新たな体制づくりを痛感するとともに、遺伝子源である林崎グループと可能な協力は出来る限り拡げる。限られた研究リソースが許す限りにおいて、必要な研究技術開発と利用可能な遺伝子源確保にもさらに積極的に取り組む。

これからの研究と応用への展開

構造ゲノム科学は、ゲノム配列研究が、技術的各論の困難を別とすれば、一義的にATGCの並びを決め続けなければならないことと対照的である。それは、すでに述べたように、マウスのトランスクリプトームであるcDNAライブラリーの結果からも明らかであり^[9]、正確にその翻訳領域を確定することにおいても、その予想を超える頻度と数の複数のオールタナティブスプライシングによる遺伝子産物が存在することから、すべきことがまだまだ残されている。このことはもっとプリミティブである細菌からの予測遺伝子のクローニングでも半分から70%程度を大きく越えることができず、さらに発現できるタンパク質はク

ローニングされた予測遺伝子の40~70%にすぎず、現時点の知識に基づいたORF (Open Reading Frame: タンパク質を実際にコードしているDNA配列領域) 予測の信頼性など確定すべきことは決して少なくないのである。またタンパク質そのものとしても、複合体としてしか安定に存在しないようなタンパク質も少なくなく、このようなときには相互作用する相手のタンパク質との複合体で初めて立体構造決定が可能となる。現在のところ、投入した予想ORFに対してせいぜい1~2%そして歩留まりのいい好熱菌が10%強でそのタンパク原子構造解析に成功している。ハイスループット化は必須であるが、この歩留まりをどこまであげられるかが、本当の意味での構造ゲノム科学の目標であり、より広い分野への利用そして産業応用へと結びつくためのカギである。このためには、ライフサイエンスとしての総力戦それも長期化を覚悟すべきことが予感され、タンパク質の立体構造研究の基礎となる広範なトランスクリプトーム、プロテオーム研究、そして細胞世界、セローム研究、生物個体研究など生物機能解析には大変に地味な時間のかかる研究が必要であることは明白である。今私たちがハイスループットファクトリーで超好熱菌を進める中で、超好熱菌自身の生物学を避けられないことに直面しているように、それ

それぞれのターゲットに従ってふさわしい生物種の選択とその適応進化してきた生物種に見合った研究アプローチの適応放散が構造ゲノム科学にも必要でこれこそ生物論理の多様性、個性を明らかに出来る唯一の道ではないだろうか。

参考文献

- (1) ISGO ICSG2002 special issue: J. Struc. Func. Genomics(2003)in preparation.
- (2) 宮野雅司、熊坂 崇 : SPring-8利用者情報 Vol.6, No.4(2001)287.
- (3) 特許庁 : 第 部 特定技術分野の審査基準(改定案)7 タンパク質立体構造関連発明に関する事例集(案)(2002).
- (4) J. P. Walsh, A. Arora, W. M. Cohen : Science **299**(2003)1021.
- (5) 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.
- (6) 菅原光明、宮野雅司 : 現代化学 2003年2月号 No.383, (2003)41.
- (7) 宮野雅司、高 秀幸 : 現代化学 2003年3月号 No.384, (2003)43.
- (8) 山本雅貴、竹下邦和、石川哲也 : SPring-8利用者情報 Vol.6, No.3 (2001)202.
- (9) FANTOM consortium and RIKEN Genome Exploration Group Phase & Team Nature **420**(2002)563-573.

宮野 雅司 *MIYANO Masashi*

理化学研究所 播磨研究所 ハイスループットファクトリー
構造生物物理研究室
〒679-5148 兵庫県佐用郡三日月町光都1-1-1
TEL : 0791-58-2815 FAX : 0791-58-2816
e-mail : miyano@spring8.or.jp