

## ポストゲノム研究と NMR

阿久津秀雄

大阪大学蛋白質研究所

### はじめに

昨年6月にヒトゲノム塩基配列のラフトドラフトが国際ヒトゲノムプロジェクトチームとセレーラ社から発表され、大きな注目を集めた。ヒトゲノムプロジェクトは図1に示すように1985年にアメリカで始まり、1990年には日米欧の国際プロジェクトになった。頭初は2003年の完成を目指していたが、セレーラ社の参入により塩基配列の決定は著しく早まった。ヒトゲノム塩基配列の決定は生命科学における歴史的成果であり、社会的インパクトも非常に大きい。現在、ヒトにとどまらず、様々な生物種のゲノム塩基配列が明らかになりつつある。これらのゲノム遺伝子の機能解明が次の目標になる。このような現状を踏まえ、政府は「生命科学の21世紀に向けてのバイオ施策」として次のようなものを挙げている。

- 1) ヒトゲノム解析を突破口とした5大疾患の克服,
  - 2) 再生医療,
  - 3) イネゲノム,
  - 4) 安全性の確保と国民の理解の推進,
  - 5) ゲノム研究開発の加速化のための基盤整備,
  - 6) バイオ技術の応用による環境問題の克服,
- 等である。このように、ゲノム関連の施策の比重が大きくなっている。

もちろん、ゲノムの決定は基礎的な生命科学

- 1985年 ヒトゲノム計画始動 (米国)
- 1990年 国際プロジェクトとして開始
- 1991~1993年 第一段階の実施
- 1994~1998年 第二段階の実施
- 1999~2003年 第三段階の実施
- 2000年6月26日 ヒトゲノム塩基配列のドラフト発表
- 1999年 理化学研究所ゲノム科学総合研究センター発足
- 2000年 米国NIHの構造ゲノムイニシアティブ開始

図1. ヒトゲノムおよび構造ゲノムに関する動き

の研究法にも大きな影響を与えている。これらを一括してポストゲノム研究と呼ぶことができる。中でも重要なのは明らかになったゲノム塩基配列からその全遺伝子を同定し、それらの遺伝子機能を網羅的に解明することであろう。それには現在幾つかの方法が考えられている。そのうちの 하나가発現タンパク質の立体構造を決定することをとおして機能を推定しようとするものである。そのためには発現タンパク質の構造を網羅的に決定する必要がある。これを構造ゲノム科学と呼び、ポストゲノム研究の新しい

研究分野として脚光を浴びつつある。

### 構造ゲノム科学と NMR

構造ゲノム科学が提唱されるようになった背景には X 線結晶構造解析, NMR によるタンパク質の 3 次構造決定の加速的進展がある。図 2 に Protein Data Bank (PDB) に登録された生体分子の 3 次構造データの数の経年変化を示す。この図から分かるように, ここ 10 年で決定された構造の数がそのほとんどを占めており, 幾何級数的に増大してきた。このような実績を背景にして構造生物学的研究が発展し, タンパク質の立体構造と生物学的機能の関係が急速に明らかになりつつある。また, 決定されたタンパク質の構造データの解析から, タンパク質の基本的折り畳み構造 (フォールド) は限られた数しか存在しないと考えられるようになった。したがって, 有限個の基本構造を決めれば他のものはこれを基礎にして決定することができる。

このようにして, ゲノムに存在する遺伝子の発現タンパク質を網羅的に決定する構造ゲノム科学がポストゲノム研究の一つとして提唱されるようになった。2000 年にはヒトゲノムプロジェクトにならって構造ゲノムの国際的進め方を議論する会議が幾つか開かれ, ここでは今後 5 年間に約 10000 個の代表的タンパク質の構造を決定することが話し合われた。昨年 (2000 年) の 11 月には第 1 回の学術的構造ゲノム国際会議が日本で開催された。

日本は構造ゲノムにおいては世界に先行した取り組みが進められている。これは横山茂之博士の提唱により設置された理化学研究所ゲノム科学総合研究センター NMR パークによるものである。この NMR パークはタンパク質の基本フォールドのカatalog作りを目指して, 世界に先駆けて設置された。ここでは現在約 20 台の高磁場 NMR 装置が稼動しており, さらに増設される予定である。したがって, 日本における構造ゲノムは Spring 8 を用いた X 線結晶構造解析, NMR パークによる溶液構造解析が主軸

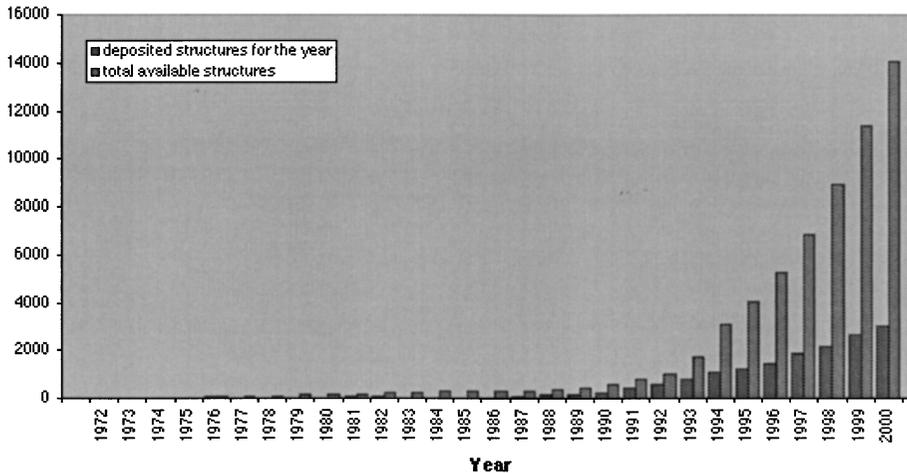


図 2. タンパク質データバンクへの登録数の経年変化  
[Protein Data Bank ホームページより (2001 年 1 月 1 日現在)]

2001 年 5 月 2 日受理

別刷請求先 〒565-0871 大阪府吹田市山田丘 3-2 大阪大学蛋白質研究所 阿久津秀雄

となって展開されている。これは世界的に見れば非常に特徴ある取り組みである。アメリカおよびヨーロッパではX線結晶構造解析が構造ゲノムの中核となっている。理化学研究所ゲノム科学総合研究センターはこれらの設備を用いて今後5年間に3000に迫る基本的タンパク質構造の決定を目指している。

### 構造ゲノム科学の方法論的問題点

構造ゲノム科学は基本的にはハイスループットによる高速度の構造決定を目指すものである。ハイスループットな構造決定に必要な要素技術はcDNAの作成、タンパク質の効率の良い発現、X線結晶解析に向くか、NMRに向くかの判断、効率の良い安定同位体標識と構造決定(NMR)、あるいは効率の良い結晶化と構造決定である。ここでボトルネックとなるのはタンパク質の発現と構造決定にむくタンパク質の選別である。後者にはNMRのドメイン化、X線での結晶化が含まれる。タンパク質の大量発現系としては大腸菌、バキュロウィルス、培養細胞、蚕の系等が研究されている。さらに、最近は無細胞の発現系が注目されている。この場

合、代謝系によるアミノ酸の変換が起こらないために安定同位体標識の効率が非常によい。無細胞系では非天然アミノ酸を取り込ませることも可能である。c型ヘムのようにヘムがポリペプチドに共有結合している場合は特殊な発現系が必要である。我々は *Shewanella* 菌を用いてこれに適した新しい発現系を開発した<sup>1)</sup>。水溶性タンパク質についてはハイスループットのためにロボットの使用を含めた大掛かりな自動化が進んでいる。

構造ゲノム科学における最大の問題はゲノムの約30%を占めるといわれている膜タンパク質であろう。これについては現在のところ、基本的な発現系自身が確立されていない。X線においては結晶化が困難であり、NMRにおいては構造解析が困難である。NMRにおいてはこのような困難を克服するために固体NMR法による構造解析法の開発が急ピッチで進んでいる。固体NMRでは本質的に分子量の制限がないので、構造解析法が確立すれば強力な武器となるであろう。固体NMRによるタンパク質の構造解析は現在のところ、配向試料を用いるものと(S. Opella, T. Crossら)と非配向試料に対してマジック角試料回転(MAS)を用いる

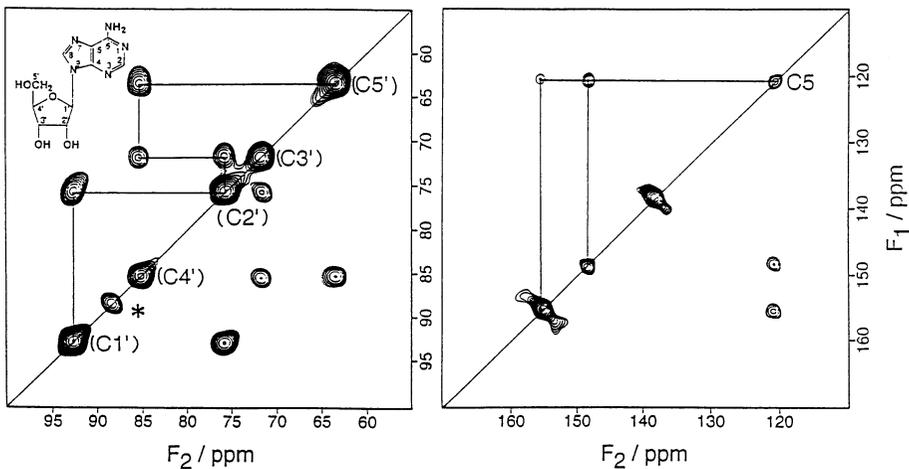


図3.  $^{13}\text{C}$ ,  $^{15}\text{N}$  標識アデノシンの二次元  $^{13}\text{C}$ -双極子 HOHAHA スペクトル<sup>2)</sup>

もの (R. G. Griffin, R. Tycko, A. McDermottら) の二つの方向で進んでいる。我々も後者の方法によるタンパク質構造解析法の開発に精力的に取り組んでいる。現在、要素技術がほぼできあがり、実際のタンパク質への適用が世界的に進んでいる。図3にこのような要素技術の一つである CP/MAS 同種核二次元相関スペクトル<sup>2)</sup>の一例を示す。

### 構造ゲノム科学と構造生物学

構造ゲノム科学はタンパク質の構造が分かれば、機能が分かるという楽観的な仮定の上に組み立てられている。しかし、タンパク質の構造が分かっただけですべての生物学的機能が分かるということはまずないであろう。ゲノムの機能を本当に理解するためには、構造ゲノム科学と構造生物学の連携は不可避である。構造生物学は Watson と Crick の DNA 逆平行二重らせんモデルとそれに基づく遺伝の分子機構仮説の提案に始まる。したがって、構造生物学は立体構造を基礎に生物機能の本質的解明に貢献し

てきた歴史がある。図2から分かるように、1990年代に入って、立体構造決定が数多くされるようになり、生命科学の中でのその比重は大きくなってきている。そして、構造を通して生物諸機能を理解するという点では NMR は最も有力な方法であることがこの間、様々な形で示されてきた。これらのことは、今後、ポストゲノム研究において NMR がいよいよ重要性を増すであろうことを示している。

### 文 献

- 1) Ozawa K, Yasukawa F, Fujiwara Y, Akutsu H : A simple, rapid, and highly efficient gene expression system for multiheme cytochrome c. *Biosci Biotechnol Biochem* 2001 ; 65 : 185-189
- 2) Fujiwara T, Sugase K, Kainosho M, Ono A, Ono A, Akutsu H :  $^{13}\text{C}$ - $^{13}\text{C}$  and  $^{13}\text{C}$ - $^{15}\text{N}$  dipolar correlation NMR of uniformly labeled organic solids for the complete assignment of their  $^{13}\text{C}$ ,  $^{15}\text{N}$  signals : an application to adenosine. *J Am Chem Soc* 1995 ; 117 : 11351-11352

## **Post-genome Sciences and NMR**

Hideo AKUTSU

*Institute for Protein Research, Osaka University  
3-2 Yamadaoka, Suita-shi, Osaka 565-0871*

Since the unveiling of the first draft of the human genome last June, the next target, to understand the biological functions of whole genes has become apparent. A new field, called structural genomics has emerged. This new fields' goal is to understand the function of expressed proteins through tertiary structures. In Japan, NMR is thought to be one of the primary methods for working on structural genomics. A large NMR facility, called NMR Park was established at RIKEN Genomic Sciences Center for this purpose. A serious international competition is developing in the field of structural genomics. Furthermore, there are many problems within structural genomics that must be overcome. To understand biological functions, structural genomics must work together with structural biology. NMR is a powerful tool in this field.