

大腸菌全蛋白質の網羅的な凝集の解析

Comprehensive aggregation analysis of whole *E. coli* proteins for understanding of protein aggregation

丹羽達也・上田卓也・田口英樹

凝集の形成は蛋白質のもつ宿命である。しかし、どのような蛋白質が凝集しやすいかなど、凝集についての理解は十分ではない。そこで、筆者らは、シャペロンを含まない再構築型の無細胞蛋白質合成系 PURE システムを用いて、約 4000 個の大腸菌の蛋白質すべてを個別に合成して凝集特性を解析した。定量できた約 3000 個のデータから、蛋白質は凝集しやすいグループと溶解しやすいグループとに明確に分かれることが示された。また、蛋白質の分子量・等電点・アミノ酸組成・立体構造などが凝集特性と関係することも明らかになった。

Key words ●蛋白質の凝集 ●網羅的解析 ●無細胞蛋白質合成系 ●シャペロン

はじめに

“蛋白質の凝集”と聞いて、みなさんはまず何を思い浮かべるだろうか？ 蛋白質を精製したことのある人であれば、その際に蛋白質が凝集してしまい困った経験をもつ人も少なくないだろう。また、細胞内に形成される封入体や、さまざまな疾患にかかわると考えられているアミロイド線維のような凝集を思い浮かべる人もいるかもしれない。

このように、凝集は蛋白質を扱う研究者にとってなじみ深いものであるが、どのような蛋白質が実際には凝集しやすいのか、どのような性質が蛋白質の凝集特性を支配しているのか、など、凝集に対する理解はほとんどなされていないのが現状である。そこで、本稿で述べる筆者らの研究では、無細胞蛋白質合成系を用いて、大腸菌のもつ約 4000 個の蛋白質すべてについて翻訳の際の凝集の度合いを網羅的に解析することで、蛋白質の凝集にかかわる性質を統計的に調べた¹⁾。これにより、凝集形成機構の一端やシャペロンの関与などに関する新しい知見が得られた。

I 凝集の網羅的な解析にむけて

1. 再構築型の無細胞蛋白質合成系

蛋白質を網羅的に解析する際にまず問題となるのが、蛋白質の調製方法である。ゲノム上の数千個にもおよぶ蛋白質をひとつずつ大腸菌で発現させて精製を行なうことは現実的ではない。無細胞蛋白質合成系はこうしたハイスループットな合成を可能にしてくれる。無細胞蛋白質合成系は、目的となる蛋白質の遺伝子をコードした DNA や mRNA を反応液にくわえることによって、細胞を用いずに試験管のなかで蛋白質をつくる系である。この合成系を使えば、多種類の蛋白質の迅速な調製が可能であるのみならず、翻訳過程での蛋白質の凝集特性を評価することができるという利点もある。

この研究では、無細胞蛋白質合成系として完全再構築型である PURE システム^{2,3)} (図 1) を用いた。通常、無細胞蛋白質合成系といった場合は細胞の抽出液を用いることが多いが、PURE システムは、開始因子や伸長因子、リボソームなど、翻訳に必須な因子のみから構成されている。また、PURE システムは凝集形成を抑制するシャペロンを含まないため、シャペロンが存在しない条件で翻訳された蛋白質の凝集の度合い、つまり、その蛋白質自体が本来もつ“凝集しやすさ”を評価することが可能である。この評価系は、シャペロンの混在がさけられない細胞抽出液を用いた無細胞蛋白質合成系では実現できないものである。

Tatsuya Niwa, Takuya Ueda, Hideki Taguchi

東京大学大学院新領域創成科学研究科 メディカルゲノム専攻分子医科学分野

E-mail : taguchi@k.u-tokyo.ac.jp

URL : <http://molbio.t.u-tokyo.ac.jp/taguchi/index.html>

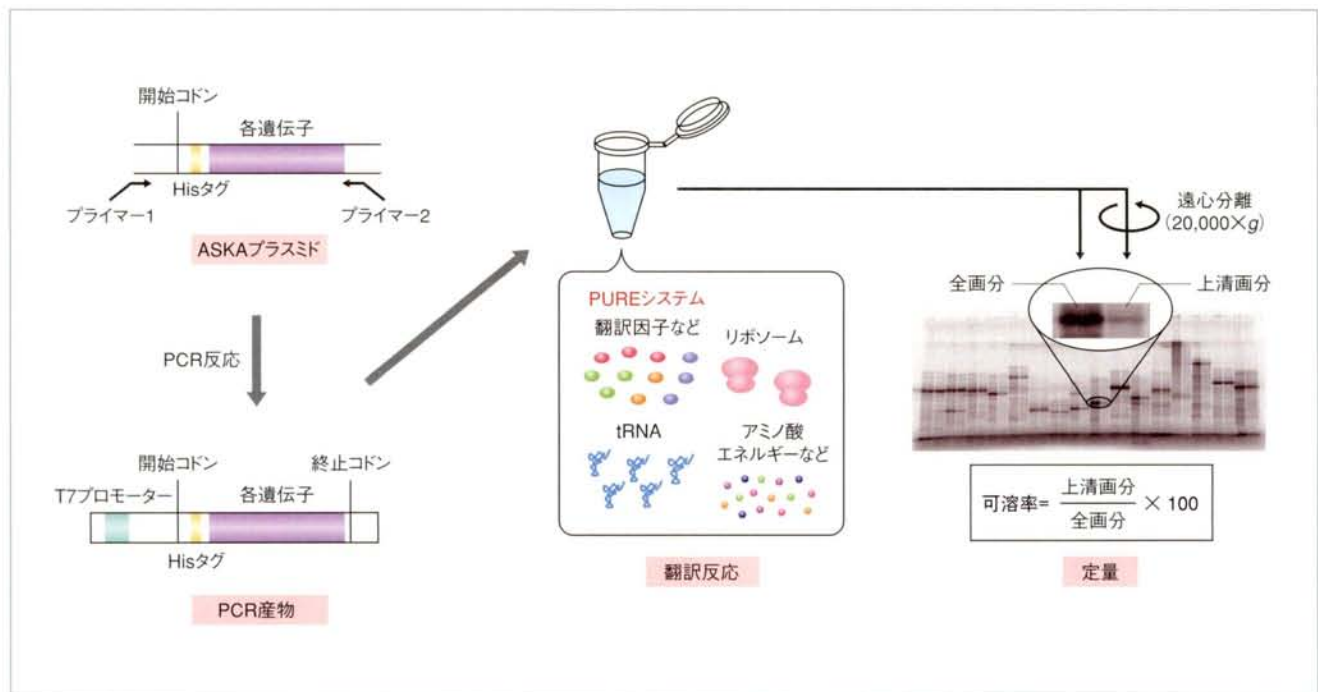


図1 実験系の模式図

ASKA ライブラリーのプラスミドをPURE システムで使えるようにPCR でそれぞれ増幅してPURE システムにくわえ、翻訳反応を37℃で1時間行なった。翻訳反応後、溶液の一部は全画分として回収し、残りの溶液は遠心分離して(20,000 ×g, 30分間)、その上清を上清画分として回収した。得られた全画分と上清画分をSDS-ポリアクリルアミドゲル電気泳動によって分離・定量し、なかにも含まれる翻訳産物の量比から可溶率を算出した。SDS-ポリアクリルアミドゲル電気泳動の際に翻訳された蛋白質のみを検出できるよう、放射性同位体標識アミノ酸 (^{35}S 標識メチオニン) を翻訳系にくわえている。

2. 大腸菌全遺伝子ライブラリー

網羅的解析の対象としては、大腸菌の予測される全蛋白質遺伝子(ORF)を網羅したライブラリーであるASKA ライブラリー⁴⁾を使い、大腸菌の全蛋白質、約4000個を合成することにした。このライブラリーはすべての遺伝子が同じプラスミドベクターに挿入されているため(図1)、PCR反応の際に同じプライマーセットを用いることで安定かつ簡便に遺伝子断片を調製することができる。

3. 遠心分離による凝集評価実験

凝集特性を評価する実験方法には、遠心分離による手法を用いた⁵⁾(図1)。まず、PCR反応でそれぞれの遺伝子断片を増幅し、放射性同位体標識したメチオニンを添加したPUREシステムによって翻訳反応を行なう。反応後、反応液の一部を分取し(全画分)、残りを遠心分離して上清を回収する(上清画分)。このようにして得られた全画分と上清画分をそれぞれSDS-ポリアクリルアミドゲル電気泳動によって分離し、蛋白質をオートラジオグラフィーによって定量

する。そして、全画分と上清画分に含まれる蛋白質の量比を可溶率と定義し、これを凝集しやすさの指標として用いた。

II 凝集しやすい集団と可溶性の高い集団の存在

ライブラリーの全遺伝子(4132個)について凝集評価実験を行なったところ、約7割(3173個)について、PUREシステムによる蛋白質の合成量および可溶率を定量することができた。残りの3割については、蛋白質の合成がみられない、SDS-ポリアクリルアミドゲル電気泳動の際に分離がうまくできない、などの理由で定量することができなかった。

定量できたものについて、可溶率の分布のヒストグラムを作成したところ、可溶率の低い部分と高い部分にそれぞれピークをもつ二峰性になることがわかった(図2a)。この二峰性は、蛋白質の合成量とは相関しないこと(図2b)、また、内膜蛋白質などを除いた細胞質の蛋白質(2277個)だけでも変化しないことから、蛋白質には凝集しやすい集団

と可溶性の高い(凝集しにくい)集団とが存在することが明らかとなった。

この二峰性についてはさまざまな解釈が可能だが、ここでいまいちど強調したいのは、この実験はシャペロンを含まない翻訳系で蛋白質を合成しているという点である。凝集しやすい蛋白質集団は正しいフォールディングにシャペロンを必要とするということを暗示している。別の見方をすれば、蛋白質というものは単独でフォールディングできるよう進化可能であったにもかかわらず、“あえて”凝集しやすいまま保存されてシャペロンへの依存性を残してきたという考察も可能である(具体的なシャペロンの基質との関係については、後述する)。

III 凝集しやすさと物理化学的な性質との関係

つぎに、蛋白質の物理化学的な性質と凝集特性との相関を調べるため、蛋白質のアミノ酸配列から得られる物理化学的な情報と可溶率との比較を行なった。以降の解析は、細胞質の蛋白質のみを対象として、可溶率0~30%の集団を凝集性集団、可溶率70%以上の集団を可溶性集団と定義している。

まず、蛋白質の分子量と等電点についてヒストグラムと

散布図による比較を行なった(図3a,b)。これらの結果より、分子量が小さいもの、また、等電点が低い(pI 5~7)ものは、可溶性が高い傾向にあることが示された。

アミノ酸組成については、負電荷をもつアミノ酸残基の割合と、芳香環をもつアミノ酸残基の割合について、凝集性集団と可溶性集団とで分布に有意な差がみられた(図3c)。一方、正電荷をもつアミノ酸残基の割合や疎水性アミノ酸残基の割合については差を見いだせなかった(図3c)。

疎水性に関しては、ヒドロパシープロット⁶⁾というアミノ酸配列の局所的な疎水性を調べるアルゴリズムを用いて可溶率との相関を調べることも行なったが、凝集しやすい集団と可溶性の高い集団とのあいだに有意な差はみられなかった。

IV 凝集しやすさと立体構造との関係

天然の蛋白質の立体構造は多様であり、それらを比較するためさまざまな分類法が存在する。この研究では、そのなかのひとつSCOPデータベース⁷⁾による分類を用いて、可溶率と立体構造との関係を調べた。SCOPデータベースは蛋白質の2次構造(α ヘリックスや β シートなど)のトポロジーによって立体構造を分類したデータベースであり、class, fold, superfamily, familyという階層構造をとっ

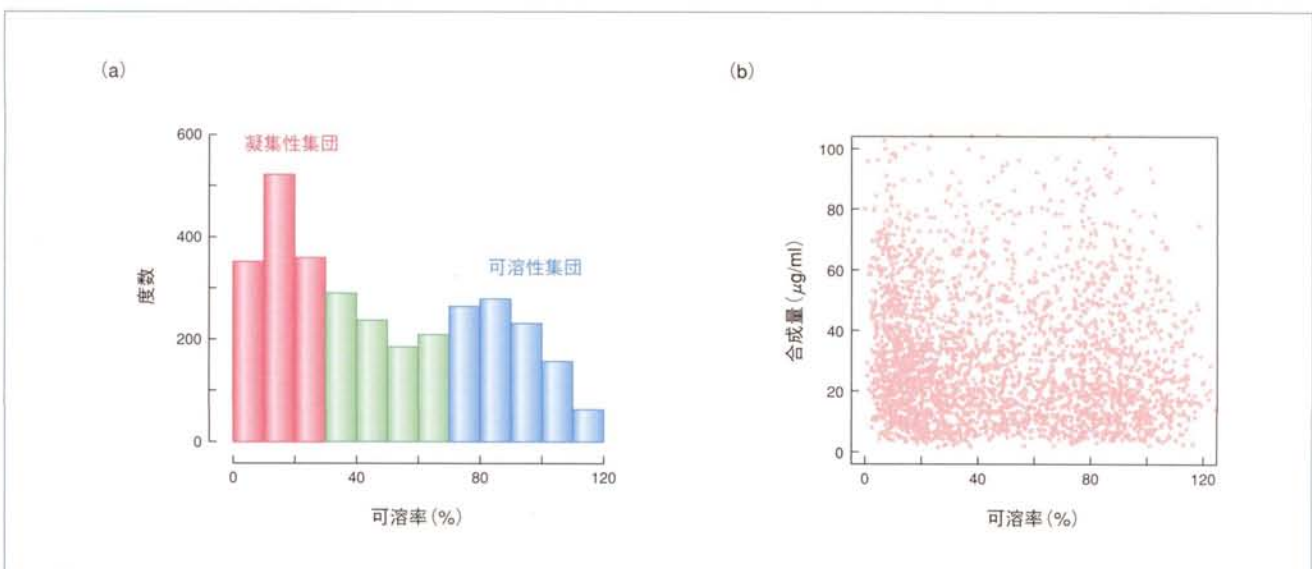


図2 可溶率の分布と合成量との関係

(a) 定量できた3173個の蛋白質の可溶率のヒストグラム。可溶率0~30%を凝集性集団(赤色)、可溶率70~100%を可溶性集団(水色)と定義し、以後の解析に用いた。

(b) 可溶率と合成量の散布図。

ている。今回は、そのなかのfoldという階層で蛋白質を分類し、それぞれのfoldのなかでの凝集しやすい集団と可溶性の高い集団の割合を調べた(図4, 表1)。これにより、たとえば、c94 (Periplasmic binding protein-like II) では凝集性集団が8割以上をしめているのに対し、c47 (Thioredoxin fold) では可溶性集団が7割以上をしめているなど、凝集しやすいfoldや可溶性の高いfoldの存在することがわかった。

立体構造は蛋白質が正しいかたちになった(フォールディングしたあとの“結果”)であり、かたちをつくるまえ(フォールディングするまえ)の凝集しやすさに関係するというのはいっけん不思議なようにも思えるが、凝集が形成されるのはフォールディングが最終段階に近くなって、ある特定のfoldを特徴づけるかたちになってから起こることを示唆しているのかもしれない。

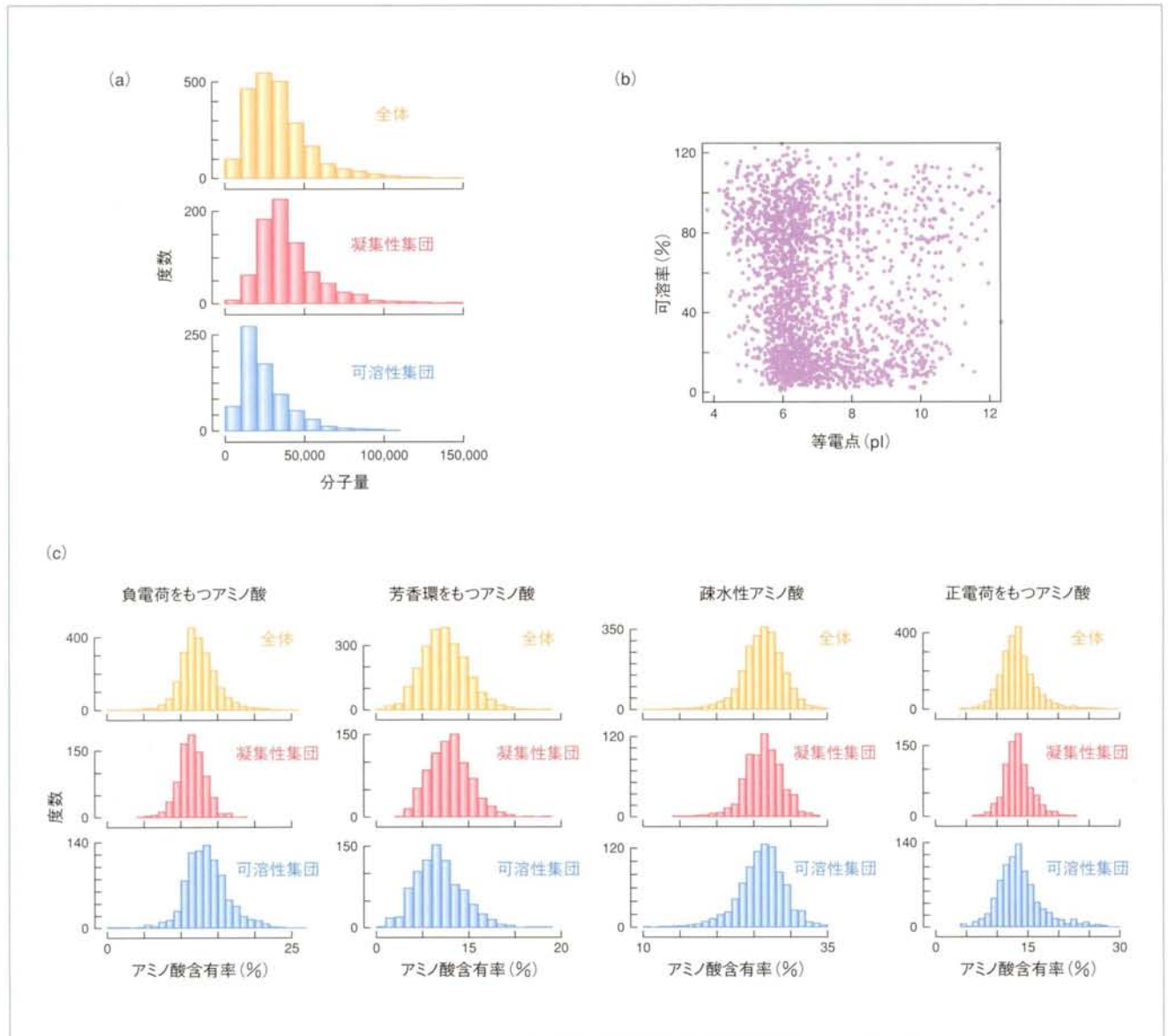


図3 可溶率と蛋白質の物理化学的な性質との関係

(a) 全体、凝集性集団、可溶性集団のそれぞれにおける分子量のヒストグラム。

(b) 可溶率と等電点の散布図。

(c) 負電荷をもつアミノ酸残基 (Asp, Glu)、芳香環をもつアミノ酸残基 (Phe, Tyr, Trp)、疎水性アミノ酸残基 (Val, Leu, Ile, Phe)、正電荷をもつアミノ酸残基 (Lys, Arg, His) の含有率を、全体、凝集性集団、可溶性集団ごとにヒストグラムで表わした。負電荷をもつアミノ酸残基と芳香環をもつアミノ酸残基では、凝集性集団と可溶性集団とのあいだで有意な差がみられた ($p < 0.001$)。

表1 SCOPデータベースにおけるfoldの名称と、それぞれのfoldにおいて凝集性集団と可溶性集団に属する蛋白質数

略号	fold	蛋白質数			p値
		全体	凝集性集団	可溶性集団	
c94	Periplasmic binding protein-like II	42	35	1	3.1×10^{-7}
c67	PLP-dependent transferases	35	25	2	1.0×10^{-4}
a4	DNA/RNA-binding 3-helical bundle	160	95	31	6.3×10^{-7}
c1	TIM beta/alpha-barrel	123	65	18	7.2×10^{-6}
c3	FAD/NAD (P) -binding domain	26	13	4	1.0×10^{-1}
c55	Ribonuclease H-like motif	44	20	12	4.2×10^{-1}
c37	P-loop containing nucleoside triphosphate hydrolases	198	88	39	3.7×10^{-4}
c66	S-adenosyl-L-methionine-dependent methyltransferases	39	15	7	2.5×10^{-1}
c2	NAD (P) -binding Rossmann-fold domains	86	33	28	1.0×10^0
c26	Adenine nucleotide alpha hydrolase-like	24	7	8	7.6×10^{-1}
d58	Ferredoxin-like	63	18	23	2.5×10^{-1}
a35	lambda repressor-like DNA-binding domains	24	6	8	5.8×10^{-1}
c72	Ribokinase-like	21	5	7	5.8×10^{-1}
c56	Phosphorylase/hydrolase-like	23	5	12	7.4×10^{-2}
b40	OB-fold	34	5	23	2.4×10^{-4}
c23	Flavodoxin-like	68	9	44	8.9×10^{-8}
c47	Thioredoxin fold	20	2	15	1.1×10^{-3}
合計		2081	781	669	

色の文字は $p < 0.01$

V 分子シャペロンの基質との関係

先述の凝集しやすいとされたSCOPデータベースにおけるfoldのなかに、大腸菌の細胞質ではたらく主要なシャペロンであるシャペロニンGroEL/ESの基質に多いといわれているfold (c67, a4, c1, c37など) が含まれていた⁸⁾。そこで、可溶率とGroEL/ESの基質との関係を直接的に調べるため、GroEL/ESの基質のみについて可溶率のヒストグラムを作成した(図5)。すると、GroEL/ES依存性の低い

基質(クラスI基質)では可溶率が高いものの割合が多いのに対し、GroEL/ES依存性の高い基質(クラスIII基質)では可溶率が低いものの割合が多いという傾向がみられた。この結果より、GroEL/ES依存性の高い基質は、もともと凝集しやすいという細胞内では不利ともいえる性質をもちながら、GroEL/ESの助けを得ることではじめて細胞内で機能を発揮することができ、維持されてきたのではないかと考えられる。

おわりに

本稿で述べた筆者らの研究では、大腸菌のすべての蛋白質について、翻訳時における凝集のしやすさを網羅的に解析し、蛋白質の分子量や等電点、特定のアミノ酸残基の含有率、蛋白質の立体構造などが蛋白質の凝集と関係していることを明らかにした。とくに、蛋白質には凝集しやすい集団と溶解しやすい集団のあること、また、ある特定の立体構造(SCOPデータベースにおけるfold)をもつものが凝集と強く相関しているという結果は、非常に示唆深い知見である。

また、蛋白質の凝集を数千個の単位で網羅的に解析するという研究は、これまでまったく前例がなく、この研究の結果は、凝集に対する理解につながるだけでなく、凝集やフォールディングの予測といったツールの開発、研究対象として凝集しやすい蛋白質(もしくは、可溶性の高い蛋白

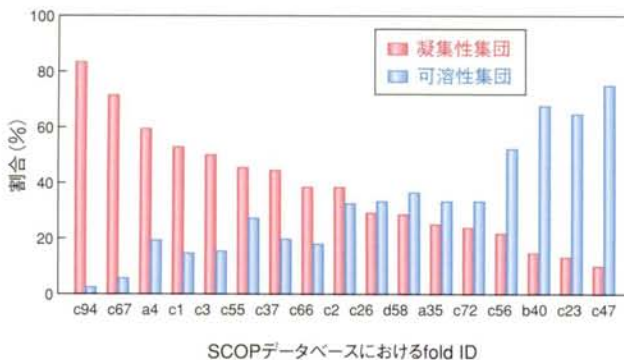


図4 SCOPデータベースにおけるそれぞれのfoldにおいて凝集性集団と可溶性集団に属する割合

それぞれのfoldの略号と名称との対応、蛋白質数は、表1を参照。

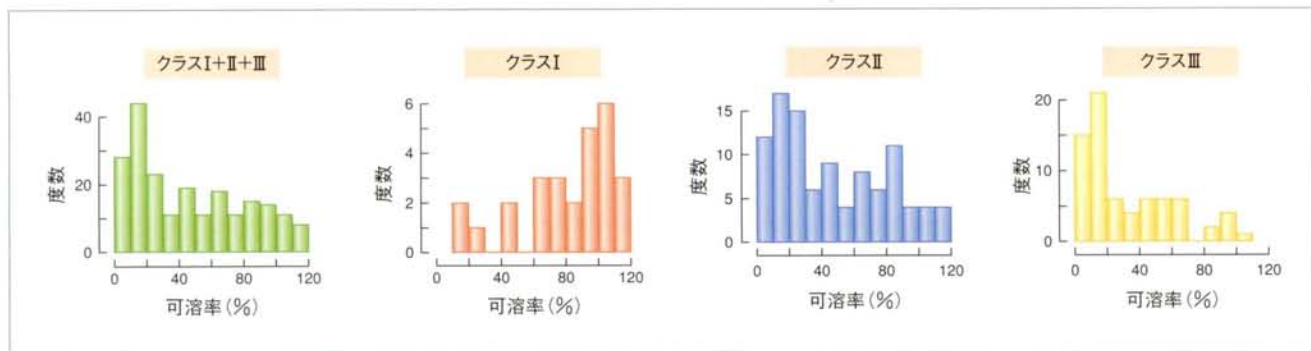


図5 GroEL/ESの基質と可溶性との関係

GroEL/ESの基質をクラスごとに分け、それぞれのクラスの可溶性のヒストグラムを示した(クラス分けは、Kernerらによるもの⁸⁾を使用)。クラスIとクラスIIはGroEL/ESへの依存性が低い基質、クラスIIIはGroEL/ESに対する依存性が高く、GroEL/ESがないと正しくフォールディングできないといわれている基質(stringent基質ともよばれている)。

質)を探したいといった際の貴重な情報リソースとなりうると考えている。現在、より多くの人にこの結果を利用してもらえるよう、可溶性や合成量などの値をデータベース化し、Webサイトで簡単に検索できるようなツールを作成しているところである。

最後に、無細胞蛋白質合成系を用いてある生物のすべての蛋白質を合成し網羅的に解析したこの研究は、ポストゲノム時代における新たなアプローチであり、“再構成プロテオーム”ということもできよう。

奈良先端科学技術大学院大学の森 浩禎博士、大阪大学の松浦友亮博士にはASKAライブラリーの取得、京都大学の高田彰二博士には実験データの解析に関してご協力をいただいた、ここに感謝したい。

文 献

- 1) Niwa, T. *et al: Proc. Natl. Acad. Sci. USA*, **106**, 4201-4206 (2009)
- 2) Shimizu, Y. *et al: Nature Biotechnol.*, **19**, 751-755 (2001)
- 3) 清水義宏・上田卓也: 蛋白質 核酸 酵素, **49**, 1520-1526 (2004)
- 4) Kitagawa, M. *et al: DNA Res.*, **12**, 291-299 (2005)
- 5) Ying, B. W., Taguchi, H., Kondo, M., Ueda, T.: *J. Biol. Chem.*, **280**, 12035-12040 (2005)
- 6) Kyte, J., Doolittle, R. F.: *J. Mol. Biol.*, **157**, 105-132 (1982)
- 7) Murzin, A. G., Brenner, S. E., Hubbard, T., Chothia, C.: *J. Mol. Biol.*, **247**, 536-540 (1995)
- 8) Kerner, M. J. *et al: Cell*, **122**, 209-220 (2005)

丹羽達也

略歴：東京大学大学院新領域創成科学研究科博士課程 在学中。