

くり返しから創る人工タンパク質

芝 清隆

公益財団法人がん研究会

〒135-8550 東京都江東区有明 3-8-31

kshiba@jfc.or.jp

(Received 29 June, 2015, Accepted 6 November, 2015)

1. イントロダクション

1983年に、当時MITの生物部にいたPaul Schimmel博士は、天然に存在するアミノアシルtRNA合成酵素の素反応に対応する機能ユニットが、遺伝子の一次配列上にサブ構造(モジュール)を単位として並んでいることを発表した[1]。アラニル tRNA 合成酵素の欠失組換え体を用いたシンプルな実験であったが、天然の遺伝子(タンパク質)を、より小さな単位の集合体へと分解可能なことを実証した印象深い論文である。筆者はその後、Schimmel博士の研究室に博士研究員として参加し、イソロイシル tRNA 合成酵素のN末欠失変異体シリーズと、C末欠失変異体シリーズを、同時に当該遺伝子ノックアウト細胞株に導入し、いわゆる「シストロン内相補[2]」を示す組合せを調べる遺伝学実験をおこなった[3][4]。これらの研究から、遺伝情報翻訳の必須タンパク質であるイソロイシル tRNA 合成酵素は、細胞増殖能を失うことなく3つのミニ遺伝子へと分解できること、またイソロイシル tRNA 合成酵素上には、多数の分断可能点が存在していることが分かった。

その後、上記分断可能点の知見をベースに、イソロイシル tRNA 合成酵素遺伝子上に複数の「マイクロ遺伝子」を定義し、この複数の「マイクロ遺伝子」の組合せ的な再構成から、*in vivo*でイソロイシル tRNA 合成酵素遺伝子として機能する人工遺伝子の創製を試みた[5]。当時、遺伝子の進化仮説として注目を集めていたWalter Gilbert博士の「遺伝子のエクソン説」[6]、すなわち「エクソン」をブロック単位として、その様々な組合せから新しい遺伝子が誕生する、とした仮説のウェットな実験系を作ろうとしていたわけである。しかしながら、残念ながらこのシステムからは、1つとして活性をもった人工タンパク質は生まれてこなかった。そもそも、「遺伝子のエクソン説」は、ある程度の遺伝子セットが誕生した後に、それらをシャッフルして、多様性・複雑性を増加する過程を説明するには魅力的な仮説ではあるが、*de novo*に最初の遺伝子が誕生した機構を説明する仮説としては、やや無理がある。「エクソン・シャッフリング型進化装置の開発」で研究予算を獲得しており、途方にくれていた時に思い出したのが、大野乾の「繰り返し仮説」である。

2. 大野乾の『繰り返し仮説』

1980年代、当時、次第に蓄積しはじめていたタンパク質や遺伝子の一次構造の中に、短い配列の繰り返しを見いだすことができることを大野乾は気づく。やがて、この観察をもとに1983年に「最初の遺伝子は短いオリゴヌクレオチドの繰り返しから生まれた」とする遺伝子誕生説を発表する[7]。この大野仮説の重要なポイントは「半保存的複製をおこなうDNAにとって、ランダムな配列の状態はむしろ実現しにくい状態」とした

ところである。複製時の相補鎖アニーリングのずれから、必然的に短い配列の繰り返しが増える、と考えたところが大野乾のユニークな着眼点である。90年代に爆発的に広がる、いわゆる進化工学的実験では、「ランダムな配列集団」から出発することを重要視している[8][9]。変異は「無方向」に起こるとするネオ・ダーヴィニズム的な考えに基づく実験系であるからだ。これに対して、大野仮説では、ある種進化に「方向性」を持ち込んでいるとも捉えることができる。

「短いオリゴヌクレオチドの繰り返し」から遺伝子が誕生したとした場合に有利な点は、(1)繰り返しからその翻訳産物(タンパク質)に周期性が生まれ、 α -ヘリックスや β -シートなどといった二次構造を獲得しやすい点、(2)3つの読み枠の何れにも翻訳終止コドンを含まないオリゴヌクレオチドが繰り返しのユニットとなった場合、得られる重合体(遺伝子)は、「フレームシフト変位」に対して耐性になる点、といったところである。

大野仮説が発表された1980年代は、遺伝子の塩基配列やタンパク質のアミノ酸配列も、限られた数しか解明されていなかった。その後のゲノム、プロテオーム、構造生物学の進展は驚くべきスピードで進展し、そこから蓄積したデータを眺めてみると、確かに遺伝子・ゲノム・タンパク質の中には、いろいろなレベルの繰り返し構造が遍満していることがわかる[10]。全部とはいわないものの、一定の割合の遺伝子は、「繰り返しを原理として」誕生したと考えてもよさそうである。

3. 単一マイクロ遺伝子の重合体を調製するMPR法

さて、上記のように「エクソン・シャッフリング型進化装置」の開発に取り組んでいた筆者であるが、その開発過程で面白い現象に遭遇する。これは、ある条件下で、「3'末が部分的に相補鎖を形成する」一対のオリゴヌクレオチドを、鑄型なしでPCR反応をおこなうと、極めて効率よくプライマーダイマーの重合体が形成される現象である。「3'末が部分的に相補鎖を形成する」PCRプライマーのデザインは、プライマーダイマーが主産物になってしまうので、「悪いプライマーデザイン」の一例なのであるが、条件を整えると、このプライマーダイマーがなぜか、タンデムに重合してしまうのである。効率よい重合体の形成には、3'末端に、意図的にミスマッチペアを入れておくことが重要で、この特殊な反応をMPR(microgene polymerization reaction)と呼び、繰り返しを原理とした人工タンパク質創製の基盤技術としている[11]。

MPR法を用いることにより、簡単に特定の短いDNA配列のタンデム重合体を形成させることができる。繰り返しの単位となる短いDNA配列は、初期値として実験毎に設定しなければならない。

この短い DNA 配列を「マイクロ遺伝子」と呼んでいる。試しにいくつかのマイクロ遺伝子を設定し、その重合体産物を大腸菌細胞の中で翻訳してみた[12]。設定したマイクロ遺伝子は、例えば天然タンパク質の一部であったり、特定の機能に関連したペプチド・モチーフであったり、あるいは、tRNA の配列をそのままマイクロ遺伝子に使ったり、さらに大野乾が好んで取り上げていた「回文配列」[13]であったりであった。何れの場合も、ポイントは「1つのマイクロ遺伝子の3つの翻訳読み枠に終止コドンが出現しない」ように DNA 配列をデザインすることである。これは、MPR 法で得られるマイクロ遺伝子重合体の、マイクロ遺伝子の連結部位に、短い塩基の欠失や挿入がしばしば起こることと関係してくる[11]。すなわち、マイクロ遺伝子重合体は、巨視的には単一 DNA ユニットの単純なタンデム重合体だが、その翻訳産物は、翻訳の読み枠が、マイクロ遺伝子の連結部に生じる微細なノイズによりずれ、結果的に、「1つのマイクロ遺伝子がコードする3つのペプチドの組み合わせ的な重合体」となる。上で紹介した、大野仮説の「3つの読み枠の何れにも翻訳終止コドンを含まないオリゴヌクレオチドが繰り返しのユニットとなった場合、得られる重合体(遺伝子)は、フレームシフト変位に対して耐性になる」ことを利用したデザイン方法である。このようにして、半ば適当に選んだ初期値としてのマイクロ遺伝子ではあったが、意外にもその重合体翻訳産物は、「扱いやすさ」を示してくれた。タンパク質としての「扱いやすさ」を定量的に表現するのは極めて難しいのだが、「扱いにくさ」の典型例としての「細胞の中で発現しない」「タンパク質が封入体を形成してしまい、解析で困難になる」といった状況に陥る確立が、経験的に低いのである。しかも、このようにして適当に作製した人工タンパク質であるが、そこそこにタンパク質らしさを示してくれた[12][14][15]。この結果を踏まえて、さらにマイクロ遺伝子のデザイン方法を工夫し、「繰り返しを原理とした人工タンパク質創製システム、MolCraft」の構築へと進むことになる[16][17]。

4. 繰り返しを基盤とした人工タンパク質創製装置、MolCraft

上の初期実験から、あらかじめ初期値として用

いるマイクロ遺伝子に、「機能」や「構造」を潜源化させておくことにより、重合体翻訳産物に特定の機能や構造を関連づけ可能なことがわかる。「機能」に関しては、いろいろな「機能モチーフ」が短いペプチド配列として提案されているので、これをマイクロ遺伝子のいずれかの翻訳読み枠に埋め込めばよい。MPR 法で得られる重合体は、単一マイクロ遺伝子の3つの翻訳読み「組合せ的(コンビナトリアル)」に連結した「分子多様性集団(ライブラリー)」としての性質をもち、平均すると3つの翻訳読み枠が1/3ずつ出現することになる。したがって、潜源化する「機能」も1つに限定するのではなく、積極的に3つの機能を3つの翻訳読み枠に埋め込めばよい(Fig. 1)。さらに、この機能モチーフをマイクロ遺伝子に埋め込む際に、用いるコドンを手早く選ぶことで、「構造」をマイクロ遺伝子に潜源化させることができる。すなわち、例えばロイシンやアルギニンを指定するコドンを選ぶ際には、6つのコドンの自由度がある。それぞれのアミノ酸に対する自由度は1から6といった微小な数ではあるが、これがペプチド配列となると、その組合せは莫大な数となる。すなわち、「あるペプチド配列をコードできる DNA 配列は莫大な数になる」ことを利用し、第2, 第3読み枠を積極的に使おうというものである。具体例をみていただくのが分かりやすいが、例えば「ArgLysVal-LeuGlnGlyArgMetGluAsnLeuGlnAlaGlu」の14残基ペプチドをコードできる DNA 配列は、実に530万種を数える。この530万種の DNA 配列を計算機内に発生させる。この DNA 配列は、その第1翻訳産物は、全て同一のペプチド配列をコードしていることに留意していただきたい。次に、この DNA 配列の中から、読み枠に終止コドンをもつものを除去していく。第2, 第3読み枠に終止コドンが含まれていると、重合体の翻訳がそこでストップしてしまうからである。続いて、残った DNA 配列の第2, 第3読み枠がコードするペプチド配列を計算機内で翻訳させる。この第2, 第3読み枠由来のペプチド集団に対して、計算機の中でセレクションをかけるわけだが、多くの場合「αヘリックスの形成能力」を指標にセレクションをかける。これは、経験的に「αヘリックスの形成能力」の高いペプチドをコードするマイクロ遺伝子から調製したライブラリーには、高頻度に

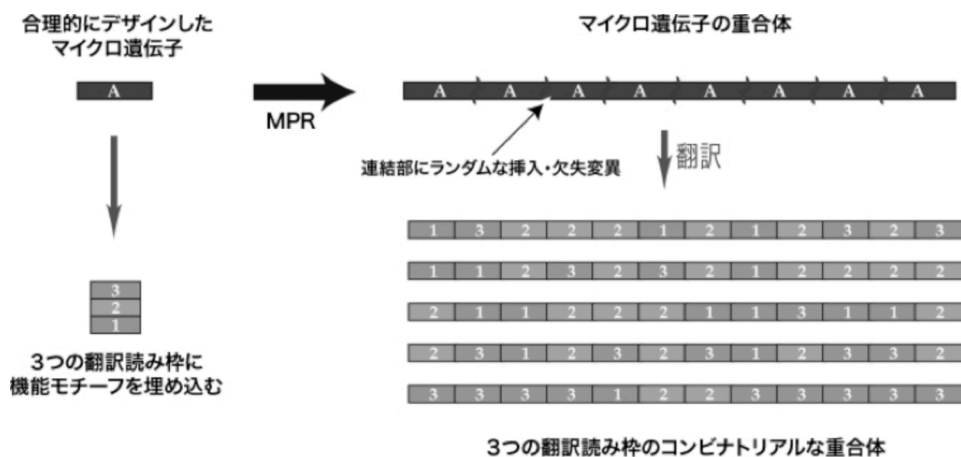


Fig. 1. マイクロ遺伝子の繰り返しから人工タンパク質ライブラリーを作製する。

「発現量が高く」「可溶性の性質」を示す人工タンパク質クローンが含まれることが分かっており、これがその後の解析を容易にするからである。この計算機内の、ある種、*in silico* 進化のようなマイクロ遺伝子のデザインは、「CyberGene」と呼ぶプログラムでおこなう。この「CyberGene」を利用した、単一マイクロ遺伝子の繰り返しを原理とした、人工タンパク質の創製法を「MolCraft」と呼んでいる[16][17]。

5. MolCraft による「モチーフ・プログラミング」

「MolCraft」では、まず「どんな機能をもった人工タンパク質を創りたいか」から出発する。そして、その機能に関連した「モチーフ」を探し出してこなければならない。注目するモチーフ(=機能)は、1つでもよいのであるが、多くの場合2つ、3つの機能を埋め込むことが多い(4つ以上のモチーフを埋め込み[18]、マイクロ遺伝子もつ裏表6つの読み枠を全て使いきったり[19]といった、MolCraft の発展型も開発したが、ここでは割愛する)。

研究も、この段階まで来ると、「いったい、どんな有用な人工タンパク質を創ってくれるのか?」といった応用への期待が高まってくる。役に立つ人工タンパク質創製システムとしてのMolCraft の真価が問われるわけである。応用を考えるなら、「天然にはない機能の組合せ」を狙うのが1つの方向である。例えば、「アポトーシス誘導に関連したモチーフ」と「細胞の自動進入に関連したモチーフ」を組み合わせる事で、「細胞に自動的に進入して、アポトーシスを誘導する」といった、天然にはない人工サイトカインのようなものも MolCraft で創製することができる[18][20][21]。

もうひとつの「役に立つ人工タンパク質」を創り出す方法は、「天然のモチーフ」と「人工モチーフ」の組み合わせである。「人工モチーフ」とは、例えば、「チタンに結合するモチーフ」[22]や「カーボンナノ化合物に結合するモチーフ」[23]などであり、2000年代に大きな流れとなったバイオナノテクノロジーの分野で、いわゆる「進化工学」的手法を用いて次々と創製された「無機材料結合モチーフ」などが含まれる。

このようなMolCraftを利用して、「天然にはないモチーフ(=機能)の組み合わせ」「人工モチーフと天然モチーフの組み合わせ」をもつ人工タンパク質を創る操作を「モチーフ・プログラミング」と呼んでいる[24][25]。「モチーフ・プログラミング」で重要なポイントは、「モチーフに単純な加算性が成り立たない」ことである[24][25]。すなわち、A という機能を持ったペプチドと B という機能をもったペプチドを単純に連結しても、A+B の複合機能をもったペプチドが得られるとは限らないという点である。もちろん、単純にペプチド連結だけでなく、両方の機能をもったペプチドが得られる場合もあり[26]、そのような場合は、わざわざ「モチーフ・プログラミング」などといった回り道をしなくてよい。ところが、多くの場合は、単純な連結では、十分モチーフの機能を引き出すことができない。そのような場合に、MolCraft による「モチーフ・プログラミング」で、狙った機能を発揮する人工タンパク質を創り出すことができる場合がある[20][27][28]。埋め込ん

だモチーフが、どのようなコンテキストでタンパク質内に配置された際に、最も効率よくそのモチーフの機能が発揮されるかは、予め予想することは難しい。多様なコンテキストで配置された、ライブラリーの中から、まず「機能」で選択し、(必要ならば)その後に、機能発揮のための条件を探し出すことになる。したがって、MolCraft による人工タンパク質創製は、「合理的に機能=モチーフを埋め込む」蛋白工学的側面と、「分子多様性集団(ライブラリー)からの機能を指標とした選択」といった、進化工学的側面を併せもつシステムと捉えることができる。

6. 今後の研究

「遺伝子がどのような原理で誕生したのか」といった基礎的な興味からスタートし、「どんな有用な人工タンパク質が創れるのか」といった応用科学として展開してきた、「くり返しから創る人工タンパク質」研究は、「モチーフ・プログラミング」といった形に収束してきた。最後のステップは、実際に現場で活躍する「トランスレーション研究」となるわけだが、ここがなかなか大変な道のりとなる。「人工サイトカイン」など、研究としては魅力的な響きをもつが、実際に患者に投与するまでに成長するのかということ、抗原性が未知の人工タンパク質を、薬として使うには、かなりの高いハードルが予想される。筆者らは現在、免疫分野の研究者と、強い細胞性免疫を誘導することのできる「人工抗原」の分野で「くり返しから創る人工タンパク質」が活躍できないかと、免疫分野でのトランスレーション研究を進めている[29]。天然タンパク質の改変体としての人工タンパク質は、既に即効型インスリンを始め、多くが実用化されてはいるが、いわゆる *de novo* に創られた人工生体高分子の応用利用は、極めて数が限られている。「自由自在に遺伝子を創る」といった進化研究者の夢を、より現実的な形に深めるため、引き続きモチーフ・プログラミング研究を進めていきたい。

References

1. Jasin, M., Regan, L. and Schimmel, P. Modular arrangement of functional domains along the sequence of an aminoacyl tRNA synthetase. *Nature*, 306, 441-447, 1983.
2. Schlesinger, M. J. and Levinthal, C. Hybrid protein formation of *E. coli* alkaline phosphatase leading to in vitro complementation. *Journal of molecular biology*, 7, 1-12, 1963.
3. Shiba, K. and Schimmel, P. Functional assembly of a randomly cleaved protein. *Proc Natl Acad Sci U S A*, 89, 1880-1884, 1992.
4. Shiba, K. and Schimmel, P. Tripartite functional assembly of a large class I aminoacyl tRNA synthetase. *J Biol Chem*, 267, 22703-22706, 1992.
5. Shiba, K., Hatada, T., Takahashi, Y. and Noda, T. Guide oligonucleotide-dependent DNA linkage that facilitates controllable polymerization of microgene blocks. *J Biochem (Tokyo)*, 132, 689-696, 2002.
6. Gilbert, W. The exon theory of genes. *Cold Spring Harbor Symposia on Quantitative Biology*, 52, 901-905, 1987.
7. Ohno, S. and Epplen, J. T. The primitive code and repeats of base oligomers as the primordial protein-encoding sequence. *Proc Natl Acad Sci U S A*, 80, 3391-3395, 1983.
8. Ellington, A. D. and Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature*, 346, 818-822, 1990.
9. Tuerk, C. and Gold, L. Systematic evolution of ligands by exponential enrichment - RNA ligands to bacteriophage-T4

- DNA polymerase. *Science*, 249, 505-510, 1990.
10. 芝 清隆. 蛋白質にひそむくり返し構造--くり返し原理による人工蛋白質の試み. *蛋白質核酸酵素*, 46, 16-25, 2001.
 11. Shiba, K., Takahashi, T. and Noda, T. Creation of libraries with long open reading frames by polymerization of a microgene. *Proc Natl Acad Sci U S A*, 94, 3805-3810, 1997.
 12. Shiba, K., Takahashi, Y. and Noda, T. On the role of periodism in the origin of proteins. *J. Mol. Biol.*, 320, 833-840, 2002.
 13. Ohno, S. Of palindromes and peptides. *Hum Genet*, 90, 342-345, 1992.
 14. Shiba, K., Honma, T., Minamisawa, T., Nishiguchi, K. and Noda, T. Distinct macroscopic structures developed from solutions of chemical compounds and periodic proteins. *EMBO reports*, 4, 148-153, 2003.
 15. Shiba, K., Shirai, T., Honma, T. and Noda, T. Translated products of tandem microgene repeats exhibit diverse properties also seen in natural proteins. *Protein Eng*, 16, 57-63, 2003.
 16. Shiba, K. MolCraft: a hierarchical approach to the synthesis of artificial proteins. *J. Mol. Catal. B*, 28, 145-153, 2004.
 17. 芝 清隆. MolCraft: 階層的進化を模した人工蛋白質創出法. *蛋白質核酸酵素*, 48, 1506-1510, 2003.
 18. Saito, H., Minamisawa, T. and Shiba, K. Motif programming: a microgene-based method for creating synthetic proteins containing multiple functional motifs. *Nucl. Acids Res.*, 35, e38, 2007.
 19. Kashiwagi, K. and Shiba, K. Combinatorics of peptide sextets encoded by a single microgene. *J. Mol. Catal. B*, 28, 215-221, 2004.
 20. Saito, H., Honma, T., Minamisawa, T., Yamazaki, K., Noda, T. et al. Synthesis of functional proteins by mixing peptide motifs. *Chem Biol*, 11, 765-773, 2004.
 21. Saito, H., Kashida, S., Inoue, T. and Shiba, K. The role of peptide motifs in the evolution of a protein network. *Nucl. Acids Res.*, 35, 6357-6366, 2007.
 22. Sano, K. and Shiba, K. A hexapeptide motif that electrostatically binds to the surface of titanium. *J Am Chem Soc*, 125, 14234-14235, 2003.
 23. Kase, D., Kulp, J. L., III, Yudasaka, M., Evans, J. S., Iijima, S. et al. Affinity selection of peptide phage libraries against single-wall carbon nanohorns identifies a peptide aptamer with conformational variability. *Langmuir*, 20, 8939-8941, 2004.
 24. Shiba, K. Natural and artificial peptide motifs: their origins and the application of motif-programming. *Chem Soc Rev*, 39, 117-126, 2010.
 25. Shiba, K. Exploitation of peptide motif sequences and their use in nanobiotechnology. *Current opinion in biotechnology*, 21, 412-425, 2010.
 26. Yoshinari, M., Kato, T., Matsuzaka, K., Hayakawa, T. and Shiba, K. Prevention of biofilm formation on titanium surfaces modified with conjugated molecules comprised of antimicrobial and titanium-binding peptides. *Biofouling*, 26, 103-110, 2010.
 27. Tsuji, T., Oaki, Y., Yoshinari, M., Kato, T. and Shiba, K. Motif-programmed artificial proteins mediated nucleation of octacalcium phosphate on titanium substrates. *Chem Commun (Camb)*, 46, 6675-6677, 2010.
 28. Murakami, T., Kashiwagi, K. and Shiba, K. Creation of novel signalling modulators from existing cytokine using scanning motif-programming. *Chem Commun (Camb)*, 47, 9357-9359, 2011.
 29. Ito, M., Hayashi, K., Adachi, E., Minamisawa, T., Honma, S. et al. Combinatorial Contextualization of Peptidic Epitopes for Enhanced Cellular Immunity. *PLoS One*, 9, e110425, 2014.