

論文の内容の要旨

氏名 石田真巳

論文題目

Leader open reading frame を用いる好熱菌遺伝子の翻訳効率の改善

ゲノム DNA の GC 含量が偏っている生物には、好熱菌、放線菌、古細菌、根粒菌など様々な有用微生物が含まれている。また、結核菌やハンセン病菌などの病原菌も含まれている。これら多種の微生物に関して、タンパク質科学や生物科学の基礎研究、産業への利用効率を上げるための育種研究、病気治療のための基礎医学研究などのために、遺伝子産物であるタンパク質を高効率で生産する意義は大きい。しかし、これらの生物由来の GC-リッチあるいは AT-リッチ遺伝子は、大腸菌での発現効率がしばしば低い。現在、様々な発現系が存在するが、大腸菌の系には、増殖が速い、培養が容易、安価であるなどの利点が多い。従来、こうした GC バイアスのある遺伝子の発現効率が低い原因として、コドン使用頻度が宿主のそれと異なること、そして、発現を阻害するような二次構造が mRNA 上にできること（特に GC リッチ遺伝子の場合に多くみられる）が指摘されていた。しかし、反例が少なくないため、低発現効率の一般的な原因はこれらだけでは説明できず、そのために改善方法も不完全であった。そこで本研究では、GC-リッチな高度好熱菌 *Thermus thermophilus* 由来の数種遺伝子を主な研究材料として、大腸菌での発現効率が低い原因を明らかにし、発現効率の改善方法を作成することを目的とした。

Thermus thermophilus 遺伝子の大腸菌内発現効率に与える leader ORF の影響

T. thermophilus 由来の *leuB* 遺伝子、*trpBA* 遺伝子を実験材料として、大腸菌内発現効率に影響する遺伝子構造を分析した。*T. thermophilus* 由来の遺伝子には共通して AGG、CGG を始めとする大腸菌にとってのレアコドンが頻繁に使用されている。*leuB* 遺伝子は、好熱菌由来の 5' 非コード領域を一緒に組み込むか否かで発現効率に変化したため、発現効率に影響する構造を調べるモデルとして選んだ。*leuB* 遺伝子の 5' 非コード領域を分析した結果、leader open reading frame (leader ORF) と名付けた別の ORF が存在すると *leuB* 遺伝子の発現効率が顕著に増加することが見出された。leader ORF の開始コドンを変異で破壊すると、下流 *leuB* 遺伝子の発現は激減した。また、*leuB* 遺伝子の発現効率は、leader ORF の長さが短い方が高く、leader ORF と好熱菌 ORF の間の距離が短い方が高かった。leader ORF として *cat* ORF を好熱菌 *leuB* ORF の上流、下流に各々挿入し、上流に挿入した場合だけ好熱菌遺伝子の発現効率が上がることを明らかにした。*leuB* 遺伝子の翻訳開始部位には、翻訳開始を阻害すると思われる mRNA の二次構造が検出されたが、この二次構造を破壊するよりも leader ORF を導入する方が発現効率は高かった。また、*T. thermophilus trpBA* 遺伝子の発現効率も leader ORF で向上することを確認した。さらに、*cat* 遺伝子をレポーターとする分析や、リアルタイム定量 PCR による mRNA の直接定量を行った結果、好熱菌 *leuB* 遺伝子の転写量はプロモーター強度に依存するが、転写量の増減は翻訳産物の量には直接反映されないことが判明した。これらの実験結果から、好熱菌遺伝子は大腸菌での翻訳開始効率が低く、leader ORF には下流好熱菌 ORF の翻訳開始効率を上げる作用があると判断した。

発現向上効果が大きな leader ORF の構造

leader ORF と好熱菌 ORF との間の距離や重複が発現効率に及ぼす影響を明らかにするため、*T. thermophilus polA* 遺伝子に以下のような間隔で 36 bp の leader ORF を導入した。leader ORF と *polA* ORF との間隔を 10 bp または 2 bp 開けたもの、leader ORF と *polA* ORF とを 1 塩基重複 (TGATG モチーフ) または 4 塩基重複 (ATGA モチーフ) させたものである。好熱菌遺伝子の発現効率を比較した結果、1 塩基重複または 4 塩基重複させると発現効率の上昇が著しく、このことから、leader ORF は翻訳カップリング機構で下流好熱菌 ORF の翻訳開始効率を向上させると推定した。*T. thermophilus leuB* 遺伝子の場合には、*leuB* 開始コドンから数百 bp 上流に離れた leader ORF にも発現向上効果が認められたが、*polA* 遺伝子の場合と同様の ATGA モチーフで leader ORF

を重複させた方が発現効率は大きく向上した。

次に、遺伝子産物が $\alpha_2\beta_2$ 複合体であるトリプトファン合成酵素をコードする *trpA* (α サブユニット) 及び *trpB* (β サブユニット) 遺伝子の、leader ORF 存在下での発現効率を調べた。その結果、*trpB* 遺伝子からは多量の産物が生産されたのに対して、*trpA* 遺伝子からは僅かしか生産されなかった。そこで、leader ORF と *trpA* ORF の接続部分周辺の塩基配列を変異させ、発現効率との関係を調べた。その結果、leader ORF の後半部分の AT 含量が高いほど下流 *trpA* 遺伝子の発現効率が上がることが分かった。また、*trpA* 遺伝子のリボソーム結合配列として、leader ORF 内部に AGG コドンがあると発現効率が向上した。leader ORF の改良によって好熱菌 *trpA* 遺伝子の発現は改善されたが、*leuB*、*trpB* など他の遺伝子に比べると、*trpA* 遺伝子の発現効率は相対的に低く、その原因は解明できなかった。以上の結果を総合し、下流好熱菌 ORF の翻訳効率を上げる効果が大きい leader ORF として、(1) 前半部分を翻訳開始効率の高い大腸菌遺伝子と同じ配列にし、(2) 内部に AGG コドンを含むリボソーム結合配列を置き、(3) 後半部分の AT 含量を高くし、(4) 終点を好熱菌 ORF と 1 塩基または 4 塩基重複させる、という詳細な設計を提案した。

重複型 leader ORF を用いる *T. thermophilus* タンパク質の大量生産

上記の重複型 leader ORF による発現向上効果を実証するため、この構造を *T. thermophilus* 由来の数種耐熱タンパク質の生産に適用した。まず、*leuB* 遺伝子がコードする 3-イソプロピルリンゴ酸脱水素酵素 (以下 IPMDH) を実験材料とした。その結果、強力なプロモーターの下流に、前半を大腸菌 *lacZ* (α ペプチド) の翻訳開始領域と同じ配列にした全長 36 bp の重複型 leader ORF を置き、その下流に *leuB* コード領域を置くと、最も生産効率が上がった。次に、このような上流構造を簡便に導入する方法を以下の様に考案した。*lacZ* コード領域内に外来遺伝子のクローニング部位がある pUC 系プラスミドをベクターとする。これによって、*lac* プロモーターから leader ORF の前半 (*lacZ*) の領域は、ベクターの配列を使うことができる。pUC ベクターに目的遺伝子のコード領域を挿入し、PCR 介在オリゴヌクレオチド指定変異によって重複型 leader ORF の後半を導入すれば、目的の構造が完成する。この方法を *T. thermophilus* 由来の *pfk1* 遺伝子がコードするホスホフルクトキナーゼ 1 (以下 PFK1)、*atpA* 遺伝子がコードする V-ATPase A サブユニット、及び、*atpB* 遺伝子がコードする V-ATPase B サブユニットの生産に適用し、各タンパク質が高効率で生産されることを確認した。ここで、PFK1 及び V-ATPase B サブユニットは IPMDH とほぼ同レベルで

生産されたのに対し、V-ATPase A サブユニットの生産量は顕著に多かった。しかし、塩基配列の分析からは、A サブユニットの生産効率が特に高い原因は分からなかった。重複型 leader ORF の導入方法をさらに簡便にして適用範囲を広げるため、高発現化に必要な構造を組み込んだ発現ベクタープラスミドを試作した。

超好熱古細菌 *Pyrococcus furiosus* の AT-リッチ遺伝子の大量発現

AT-リッチな超好熱古細菌 *Pyrococcus furiosus* から *trpA* 及び *trpB* 遺伝子をクローニングし、大量発現のために leader ORF を導入した。*trpA*、*trpB* 両遺伝子は、*P. furiosus* ゲノム DNA から PCR 介在法で大腸菌にクローン化し、各々の塩基配列 (*trpA* 747 bp、*trpB* 1,167 bp) を決定した。コドン使用頻度は AT 側に偏っており、両遺伝子ともに AUA、AGR、CUA など大腸菌にとってのレアコドンが頻繁に使用されていた。次に、*lacZ* コード領域を欠く pUC ベクタープラスミドに各超好熱菌コード領域を挿入したところ、 α サブユニットに比して β サブユニットの生産量が著しく少なかった。そこで、第4章と同じ方法で *P. furiosus trpB* 遺伝子に重複型 leader ORF を導入したところ、 β サブユニットの生産量が 10 倍以上に向上した。この結果から、GC リッチな遺伝子のみでなく広範な遺伝子において大腸菌内の翻訳開始効率が低いものがあり、そのような遺伝子の発現効率は重複型 leader ORF の導入によって改善され得ると推定した。

まとめ

本研究において、重複型 leader ORF という構造が翻訳カップリングを通じて GC-リッチな高度好熱菌 *T. thermophilus* 由来遺伝子の大腸菌内発現効率を著しく改善することを明らかにした。また、AT-リッチな超好熱古細菌 *P. furiosus* 由来 *trpB* 遺伝子の発現も、同じ構造の導入によって改善されること確認した。これらのことから、これまで発現効率の低い原因が分からなかった GC バイアスのある遺伝子の大腸菌内発現には、翻訳開始効率が重要なネックになっていると考えられた。GC バイアスによって生ずるレアコドンも発現効率に悪影響を及ぼすかもしれないが、本研究では、重複型 leader ORF による翻訳開始効率の改善のみで、発現効率を著しく向上させることができた。これらの結果に基づいて、重複型 leader ORF の簡便な導入方法を考案し、実際に数種の好熱菌タンパク質に適用して大量生産させた。市販の発現系の多くは転写効率の向上を主な作用点としている。それに対して本研究で考案した翻訳効率、特に翻訳開始効率、を改善する方法は、従来の発現系の欠点を補い、有用生物の遺伝子が大腸菌で大量発現させるために役立つと期待される。