

## アンチコドンの化学構造と遺伝暗号解読系

愛媛大学工学部応用化学科

高井 和幸

### 1. はじめに

生体機能を担う多種多様なタンパク質は、mRNAに書き込まれた遺伝情報の「翻訳」によって合成される。翻訳の作業は、リボソームという細胞内巨大分子装置によって、mRNA上の3つの連続したヌクレオチド(コドン)を順番に読み取って、アミノ酸をその順番でつなげていくことによって行われる。アミノ酸をリボソームに運び、コドンを読み取るためのアダプターの役割をする分子がtRNAである。tRNAは、その一端に、酵素の働きによってアミノ酸を結合し、一方で、3つの連続したヌクレオチド(アンチコドン)によって、コドンを認識する。コドンとアンチコドンとの相互認識は、基本的には、核酸に見られる相補的な対合に依っている。しかし、DNAの複製やDNAからmRNAへの遺伝情報の転写のように、1対1の対応関係になっているわけではない。大腸菌では、コドン61種類に対して、46種類程度のアンチコドンが存在する。コドンとアミノ酸との対応を表にまとめたものが遺伝暗号表(表1)である。

本稿では、遺伝暗号解読系の特徴について概説した上で、それとアンチコドンの化学構造との関係について筆者が行った研究を紹介する。

### 2. tRNAの種類とアンチコドン一字目に見られるヌクレオシド

大腸菌の場合、アンチコドン一字目に表2に示すヌクレオシドを持ったtRNAが存在す

表1. 普遍遺伝暗号表

UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys		
UUC		UCC		UAC		UGC			
UUA	Leu	UCA	Pro	UAA	Stop	UGA	Stop		
UUG		UCG		UAG		UGG		Trp	
CUU	Leu	CCU	Pro	CAU	His	CGU	Arg		
CUC		CCC		CAC		CGC			
CUA		CCA		CAA	CGA				
CUG		CCG		CAG	CGG				
AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser		
AUC		ACC		AAC		AGC			
AUA		ACA		AAA	AGA				
AUG	Met	ACG		AAG	Lys	AGG	Arg		
GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly		
GUC		GCC		GAC		GGC			
GUA		GCA		GAA	GGA				
GUG		GCG			GAG	GAG		Glu	GGG

表2. 大腸菌 tRNAに見られる  
アンチコドン一字目のヌクレオシド

UUU	UCU	UAU	UGU
UUC G	UCC G	UAC Q	UGC G
UUA cmm <sup>2</sup> Um	UCA cmo <sup>5</sup> U	UAA Stop	UGA Stop
UUG Cm	UCG C	UAG Stop	UGG C
CUU	CCU	CAU	CGU I
CUC G	CCC G	CAC Q	CGC
CUA U	CCA U	CAA mmm <sup>2</sup> s <sup>2</sup> U	CGA
CUG C	CCG C	CAG C	CGG C
AUU	ACU	AAU	AGU
AUC G	ACC G	AAC Q	AGC G
AUA L	ACA cmo <sup>5</sup> U	AAA mmm <sup>2</sup> s <sup>2</sup> U	AGA mmm <sup>2</sup> U
AUG ac <sup>2</sup> C	ACG C	AAG	AGG C
GUU	GCU	GAU	GGU
GUC G	GCC G	GAC Q	GGC G
GUA cmo <sup>5</sup> U	GCA cmo <sup>5</sup> U	GAA mmm <sup>2</sup> s <sup>2</sup> U	GGA U
GUG	GCG	GAG	GGG C

る。それぞれの tRNA の量はまちまちである。A の修飾体としては、イノシン (I) が存在する。I は、大腸菌の場合にはアルギニン tRNA のみにみられるが、真核生物では、セリンやイソロイシンなどの tRNA にも存在する。G の修飾体としては、遺伝暗号表の 3 列目の各コドンボックス (一, 二字目が共通で三字目が異なる 4 種のコドン) の上半分に対応する tRNA に存在するキューオシン (Q) がある。U の修飾体は、大きく分けて、2 種類ある。Val, Ser, Ala などの、コドンボックスが 1 種類のアミノ酸を指定するようなコドンに対応する tRNA には、ウリジン 5-オキシ酢酸 ( $\text{cmo}^5\text{U}$ ; 5-カルボキシメトキシウリジン) が見られることがある。枯草菌では、同種の tRNA のアンチコドン一字目は 5-メトキシウリジン ( $\text{mo}^5\text{U}$ ) になっており、一般に、バクテリアでは、5 位に酸素原子が直接ついた修飾ウリジン ( $\text{xo}^5\text{U}$ ) になっている。また、コドンボックスが上下に 2 対 2 に仕切られている場合には、5-メチルアミノメチル-2-チオウリジン ( $\text{mnm}^5\text{s}^2\text{U}$ ), 5-カルボキシメチルアミノメチル-2'-*O*-メチルウリジン ( $\text{cmnm}^5\text{Um}$ ), 5-メチルアミノメチルウリジン ( $\text{mnm}^5\text{U}$ ) となっている。これらはいずれも、5 位に直接メチレンがついた構造 ( $\text{xm}^5\text{U}$ ) をもっている。C の誘導体としては、2'-*O*-メチルシチジン (Cm), 4-*N*-アセチルシチジン ( $\text{ac}^4\text{C}$ ), およびリシジン (L) がある。L は、コドン 3 字目の A と対応するので、表 2 においては、U の位置に示してある。

一般に、1, 2 字目が同じで 3 字目が U のコドンと C のコドンとは、同じ tRNA に対応する。その tRNA のアンチコドン一字目は、G の誘導体か、I である。つまり、通常、アンチコドン一字目の A は I に修飾され、A のままになっている tRNA は生物界にほとんど存在しない (現在までに 3 例程度見つかった)。

### 3. 遺伝暗号表の特徴とコドン-アンチコドン対合規則

コドンとアミノ酸との対応関係が、遺伝暗号表のようになっている理由は、よくわかっていない。現在の地球上の生物の遺伝暗号は、基本的には表 1 の普遍遺伝暗号である (少々の多様性は見られる)。生物の初期進化の過程で、最初は種々の遺伝暗号が存在したかもしれないが、いったん普遍遺伝暗号をもつ生物だけが生き残り、その子孫が現在の地球に広がった、と考えられている (偶然凍結説) (Crick, 1968)。しかし、普遍遺伝暗号には、いくつの特徴があり、これが決して単なる偶然で生き残ったものではないことを示唆している。

1 つのコドンボックスは、1 種類から 3 種類のアミノ酸 (停止も含めて数える) に対応している。コドンボックスが 2 種類以上のアミノ酸で区切られている場合には、区切り方に特徴があり、上下に 2 対 2, 3 対 1, あるいは、2 対 1 対 1 に区切られている。しかし、U と C との間では、決して区切られていない (同じアミノ酸が上下に分断されていたり、入れ子になっていたりすることもない)。これらの特徴から、1966 年に、よろめき (wobble) 仮説が提唱された (Crick, 1966)。コドン一, 二字目はそれぞれアンチコドン三字目, 二字目とワトソクリック型の塩基対を作り、コドン三字目とアンチコドン一字目とはワトソ

ンクリック型のほかに **wobble** 型の塩基対を作ることができる。この、コドン三字目とアンチコドン一字目との対応規則が **wobble rule** である (図 1)。その後の研究で、 $\text{xo}^5\text{U}$  型ヌクレオシドがコドン三字目の U と塩基対合できることが明らかになったほかは、**wobble rule** が基本的には正しいと考えられている。

遺伝暗号表において、区切られているコドンボックス (スプリットボックス) と区切られていないコドンボックス (ファミリーボックス) とを色分けしてみると、さらに興味深いことがわかる。一字目と二字目が G と C のみからなるコドンボックスは、区切られていない。また、A と U のみから

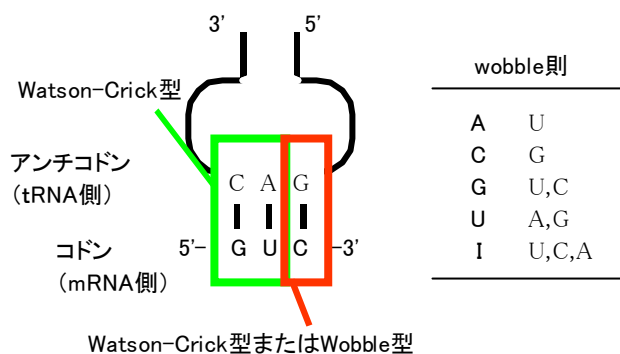


図1. コドン-アンチコドン対合とwobble則

なるコドンボックスは、区切られている。残りの 8 つのコドンボックス (一字目と二字目のうちどちらかが A または U でもう一方が G または C になっているコドンボックス) は、遺伝暗号表の右半分では区切られておらず、左半分では区切られている。これは、ファミリーボックスにおいて、コドン一、二字目とアンチコドン二、三字目との相互作用が強いいため、一字目のみの違いでアンチコドンを区別することが困難であったために、進化の初期段階において、一字目を区別しなくてもよいような遺伝暗号が有利であったからであろうと考えられている (Lagerkvist, 1981)。

#### 4. コドン使用頻度と細胞内 tRNA 存在比

1 種のアミノ酸は、複数のコドンで指定できる場合が多い。同じアミノ酸を指定する別のコドンのことを同義語コドンという。例えば、ロイシンに対応するコドンは 6 種類あり、それらを、5 種類の tRNA で解読している。同じアミノ酸に対応する別の tRNA のことをアイソアクセプターという。異なる同義語コドンは、遺伝子中に出現する頻度が異なる。特にバクテリアや菌類において、細胞内で多く作られるタンパク質の遺伝子 (高発現遺伝子) に出現する同義語コドンの使用頻度は、特定のコドンに偏っている。この偏りは、細胞内アイソアクセプター tRNA の存在量の偏りと対応している (Ikemura and Ozeki, 1977)。つまり、出現頻度の高いコドンに対応する tRNA は、細胞内存在量も多い。高発現遺伝子における同義語コドン使用頻度は、細胞内におけるそれぞれの同義語コドンを解読する活性と対応していると考えられる。

#### 5. 問題点

以上のような教科書的な事実を受け入れるのはよいが、それではなぜ、遺伝暗号解読系はそのような構成になっているのであろうか？ 遺伝暗号表は、なぜ、表 1 の普遍遺伝暗号表になったのであろうか？ tRNA のアンチコドン一字目には、なぜ、このような多様な修飾ヌクレオシドが存在するのだろうか？ 遺伝暗号表が「凍結」した過程には、偶然要因もあったはずであるし、凍結後、遺伝暗号解読のメカニズムも凍結したかどうかはわからない。従って、現在の遺伝暗号解読系の解析によってこのような問いに完全な回答を与えることはできないかもしれない。しかし、現在の遺伝暗号解読系を、物理化学的、生化学的な見地から総合的に詳細に解析することによって、遺伝暗号解読系がどのような方向に進化してきたのか、それはなぜなのか、といったことが解明できると考えられる。

しかし、筆者が研究を始めた時点では、tRNA アンチコドンに見られるヌクレオチドを単離して、そのヌクレオチドの単量体やオリゴヌクレオチドのレベルでの物理化学的測定を行い、それに基づいて遺伝暗号解読系の機能を推定した研究は多数存在したが、少なくとも実際に機能しているタンパク質合成系を用いた実験で、それらの推定が生化学的にも検証できるかきちんと定量的に調べた実験は、全く存在しなかった。また、生化学的な研究は、ほとんどが定性的なものであり、定量的な議論のできるような研究が存在しなかった。

そこで、本研究では、まず、アンチコドン一字目のヌクレオチド修飾の有無によって、tRNA とコドンとの対応関係がどのように違うのかについて、生化学的かつ定量的に解析し、さらに、全く修飾されていない tRNA のふるまいについて解析した。

## 6. 相対コドン認識効率

従来、天然 tRNA のコドン認識活性を測定するために、種々の測定法が開発されてきた。リボソーム結合実験は、アミノアシル tRNA のリボソームへの結合活性を測定する方法であるが、tRNA の種類によっては結合活性が低くて測定が不可能な場合があり、定量的な議論ができるような結果は期待できない。一方、無細胞タンパク質合成系を用いて、アミノアシル tRNA からのタンパク質へのアミノ酸の取り込みを測定することにより、tRNA アイソアクセプターのコドン特異性を解析することができる。類似の方法で、無細胞翻訳系の忠実度（間違いを起こす頻度）の解析も行われる。忠実度の測定では、定量的な議論が可能な結果が得られていたが、これを tRNA アイソアクセプター間の相対的な認識効率の研究に用いた従来の研究では、忠実度が必ずしも高くない（従って、本来の生体内の条件とは異なる可能性の高い）条件を用いていた。本研究では、十分に忠実度の高い翻訳系を用いて、tRNA アイソアクセプターの間での相対的な認識効率を測定することができるように、大腸菌無細胞タンパク質合成系とそれに用いる mRNA を工夫し、mRNA 上特定の 1 つのコドンが共通のアンチコドン二、三字目をもつ 2 種の tRNA のどちらをどれだけ好んで認識するかを定量的に測定する方法を開発した。具体的には、図 2 のように、2 種のアミノアシル tRNA のアミノ酸を  $^3\text{H}$  と  $^{14}\text{C}$  で別々に標識し、それらを同量だけ加えた無細胞タンパク質合成系において、どちらのアミノ酸がどれだけ効率よくタンパク質に取り込まれるか

を解析し、定量化した。これにより、用いた二種の tRNA の、特定のコードンに対する比活性の比が求められる。その値は、用いた二種の tRNA の違いが、比活性に及ぼす効果を定量的に表現する。

## 7. ヌクレオシド修飾の効果

### 7-1. 大腸菌 tRNA<sup>Ser</sup> の全修飾の効果

大腸菌 tRNA<sup>Ser</sup> は、アンチコードン一字目に cm<sup>5</sup>U をもち、コードン UCU, UCA, UCG を認識する。この tRNA には、多くの修飾ヌクレオシドがある。この中の、アンチコードン一字目のヌクレオシドだけを未修飾の U に変えた分子を作製し、これと元の分子とを上記の方法で比較すれば、アンチコードン一字目の修飾の効果が定量的に測定で

きるはずである。しかし、そのような分子の作製は、相当な手間と技術を要する。そこで、まず、修飾ヌクレオチドを持たないが、遺伝子の配列としては同じ tRNA (未修飾 tRNA という) を酵素的に合成し、天然型の tRNA と比較した (Takai *et al.*, 1996)。その結果、(1) 修飾を持たないことによって、tRNA の活性は低下するが、低下の度合いは 1 桁に満たないこと、および、(2) 未修飾 tRNA が、コードン UCC, UCG を全く認識しないこと、が、わかった。未修飾 U がアンチコードン一字目にある場合、コードン三字目として、すべてのヌクレオシドを認識するという考え方もあったが、この実験により、必ずしもそうでないことが明確に示された。また、アンチコードン一字目とコードン三字目との間に U-G 塩基対 (以下、コードン-アンチコードン対合に関する塩基対は、アンチコードン側を先に書き、ハイフンでつないでコードン側を後に書く。) ができないという事実は、wobble rule を修正する必要があることを示している。

### 7-2. アンチコードン一字目の U から mo<sup>5</sup>U への修飾の効果

次に、化学合成法により、未修飾 tRNA のアンチコードン一字目を mo<sup>5</sup>U で置き換えた tRNA を調製した。これを、同様な方法で調製した未修飾 tRNA と比較した (Takai *et al.*, 1999a)。その結果、(1) U から mo<sup>5</sup>U への修飾のみで、三字目が U および G のコードンとの対合が高まること、および、(2) この修飾により、三字目が A のコードンとの対合が弱まること、の 2 点が明らかになった。

### 7-3. アンチコードン一字目の C から Cm への修飾の効果

さらに、同様な方法により、アンチコードン一字目の C を Cm に修飾すると、三字目が G のコードンとの対合が強められることが確認された。この修飾は、他のコードンとの対合に影響を与えなかった (Satoh *et al.*, 2000)。

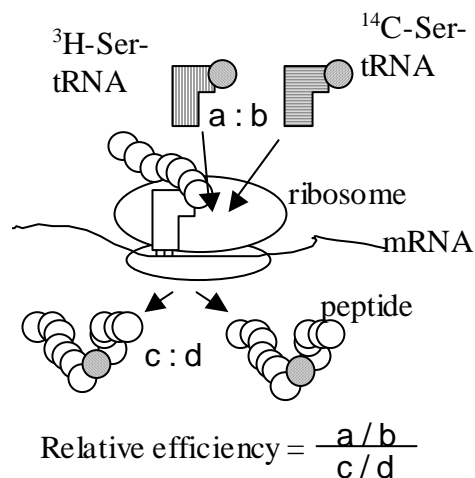


図2. 相対コードン認識効率

## 8. 未修飾 tRNA におけるコドン-アンチコドン対合規則

修飾の効果を測定するうちに、tRNA のヌクレオシド修飾が無かったら、その tRNA のコドン-アンチコドン対合規則がどうなるのか、我々が十分に理解していないことに気づいた。そこで、大腸菌セリン tRNA のアンチコドンを様々なトリプレットに変えて、コドン三字目とアンチコドン一字目との間の対合規則を確かめることにした (Takai *et al.*, 1999b)。その結果、従来の対合規則 (wobble rule) と、2 点において異なることがわかった (表 3)。第一に、アンチコドン一字目とコドン三字目との間に U-G 塩基

表 3. 未修飾 tRNA の wobble 則

Anticodon 1st	Codon 3rd
A	U*
G	U*, C
U	A
C	G

$$* A \cdot U \leq G \cdot U$$

対はできない。第二に、この位置に G-U 塩基対があるほうが、A-U 塩基対がある場合よりも、コドン-アンチコドン対合が強くなる。これらの点が天然 tRNA と異なるのは、決して未修飾 tRNA が不完全な分子だからではない。生物は、天然 tRNA で U-G 塩基対合が可能なように、アンチコドン一字目の修飾ヌクレオシドを進化させた。また、A-U 塩基対は使わずに、G-U 塩基対だけで済ませている。このように、未修飾 tRNA の規則が天然 tRNA の規則と異なるのではなく、天然 tRNA がヌクレオシド修飾などによって、未修飾 tRNA の対合規則で不十分な点をカバーしていると考えてよい。言い換えれば、アンチコドン一字目に A をもつ tRNA が自然界に稀なのは、そのような tRNA よりも、アンチコドン一字目に G をもち、それ以外の点では全く同じ tRNA のほうが、活性が高いからである。

## 9. 結語

以上のように、tRNA のアンチコドンの化学構造とその tRNA のコドン特異性との相関を詳細に議論するための素地がようやくできた。タンパク質合成系は進化の過程で非常に精密にチューンアップされてきた生体分子システムであり、特にバクテリアの系では、ほんのわずかな不具合を起こしただけでも淘汰されているようである。従って、なぜ、遺伝暗号解読系がこのようになっているのか、といったような疑問に答えるには、物理化学に基づく定量的な解析をより詳細に行うことが必要と考える。

## 謝辞

この研究は、東京大学大学院理学系研究科横山茂之教授および千葉工業大学工学部高久洋教授のもとで行われたものであり、両教授に感謝するとともに、千葉工業大学で学生として実際に実験を行ってくれた奥村修平氏、大内良介氏、佐藤旭氏に感謝する。

## 文献

- Crick, F. H. C. (1966) *J. Mol. Biol.*, 19, 548-555.
- Crick, F. H. C. (1968) *J. Mol. Biol.*, 38, 367-379.
- Ikemura, T. and Ozeki, H. (1977) *J. Mol. Biol.*, 117, 419-446.
- Lagerkvist, U. (1981) *Cell*, 23, 305-306.
- Satoh, A. *et al.* (2000) *RNA*, 6, 680-686.
- Takai, K. *et al.* (1996) *Nucleic Acids Res.*, 24, 2894-2899.
- Takai, K. *et al.* (1999a) *FEBS Lett.*, 447, 1-4.
- Takai, K. *et al.* (1999b) *Biochem. Biophys. Res. Commun.*, 257, 662-667.