

ORIGIN OF THE GENETIC CODE

Koji Tamura¹ and Tsunemi Hasegawa²

¹Department of Biological Science and Technology, Tokyo University of Science, 6-3-1 Niijuku, Katsushika-ku, Tokyo 125-8585, Japan

²Department of Material and Biological Chemistry, Faculty of Science, Yamagata University, Yamagata 990-8560, Japan.

Dr. Tsunemi Hasegawa passed away in February, 2013. He had played important roles in the Society for the Study of the Origin and Evolution of Life, Japan. I first met him when I started my graduate study in Dr. Mikio Shimizu's laboratory at the Institute of Space and Astronautical Science in 1989. Dr. Hasegawa was an associate professor in Dr. Shimizu's laboratory. I learnt a lot from Dr. Hasegawa, from the basics to the latest knowledge of biochemistry. This manuscript was written by myself and Dr. Hasegawa in 2007. I would like to dedicate this article to the memory of Dr. Tsunemi Hasegawa as a token of my gratitude.

Koji Tamura

遺伝暗号の起源

田村浩二¹、長谷川典巳²

¹東京理科大学 基礎工学部 生物工学科

²山形大学 理学部 物質生命化学科

長谷川典巳先生は、2013年2月に亡くなられました。長谷川先生は本学会で、長年に渡り重要な役割を果たされてきました。私が東大の物理学科の大学院生として、宇宙科学研究所の清水幹夫先生のもとで研究を開始した1989年に、清水先生のところの助教授としておられたのが長谷川先生でした。長谷川先生には、生化学のイロハから最新の研究までご教授いただき、私にとって現在の礎を築いて下さった大恩人です。この論文は、2007年の時点で、私と長谷川先生との間で草稿のやり取りをしていたものです。長谷川先生への感謝と追悼の気持ちを込めて、共著論文として提出します(田村浩二)。

1. はじめに (文献 1-5)

現在の地球上に見られる生命体は、DNAの遺伝情報がRNAに転写され、それが蛋白質に翻訳されることによって機能を発揮している(セントラルドグマ)。この機構は基本的に、バクテリアから我々ヒトに至るまで共通しており、地球上の生命が、ある共通の祖先から発生し、進化したということを示している。しかし、ここで、DNAやRNAを作るためには蛋白質が必要であり、逆に蛋白質はDNAやRNAから作られる、という矛盾が発生する。いわゆる「鶏と卵」の問題である。この矛盾は1982~1983年のCechとAltmanによるリボザイムの発見により一応、解決された。テトラヒメナのグループIイントロン、および、

RNasePの構成成分であるMIRNAがRNAとして酵素活性を持つことの発見である。RNAが情報の担い手として、また酵素として機能し得たということは、生命の起源を考える上で極めて重要な発見であり、この発見以降、「RNAワールド」が生命の起源における重要な概念として登場することになった。

1953年に、WatsonとCrickによって、DNAの二重らせんモデルが提唱され、その後、生物学は質的に大きく変換してきた。それまでの分類学や博物学的生物学から、分子レベルで物事を議論する、いわゆる「分子生物学」の誕生である。分子生物学の成功はヒトの全ゲノムの解読という大きな流れに象徴される。

DNAに記されたA、T、G、Cの4つの文字がmRNAに写し取られ、3つの文字の組が1つのアミノ酸を指定することによって、蛋白質が生成される。この対応は「遺伝暗号」として知られている。「遺伝暗号」は、古細菌、真正細菌、真核生物という、現在の地球上のすべての生物系において共通に見られる、RNAの塩基配列を蛋白質中のアミノ酸の配列に変換するアルゴリズムである。しかし、この「遺伝暗号」はどのようにして決まったのだろうか?この対応関係は、いわば「暗号」でしかない。例えば、リシンというアミノ酸を指定する3つの文字はAAAであるが、ではどうしてAAAなのだろうか?なぜGGGではないのだろうか?現代の最先端の生物学においても、この対応関係をすべて「暗号」のまま置き去りにしたまま、研究が進められているのが現状である。

しかし、私達は地球上で進化してきた生命体である。その過程では、遺伝暗号がなぜそのように決まったのかという必然性を経ているはずである。生命の起源を考える上で、最も重要な問題のひとつは、このようにアミノ酸とRNAの対応がいかにかして決まったのか?という「遺伝暗号の起源」の問題である。本稿では、これまでに提唱されてきた、遺伝暗号の起源に関する諸説について概説する。

2. ガモフのダイヤモンドコード (文献 6)

宇宙の起源の謎に一石を投じたビックバン仮説の提唱者として知られる物理学者Gamowは、遺伝暗号の起源にも興味を持ち、Crick達と「RNAタイクラブ」を作った。このクラブは20人のメンバーからなり(お揃いのネクタイを作りメンバーにはアミノ酸の名前をつけた。Gamowはアラニン、Crickはチロシンだったらしい)、遺伝暗号の解読に向けて活発に情報交換を行っていた。Gamowは、

二重らせんモデルが提唱された翌年(1954年)、自身が「ダイヤモンドコード」と呼ぶ、遺伝暗号の起源についての仮説を *Nature* 誌に提唱した。これは DNA の二重らせんを構成する塩基配列が形成する溝に、対応するアミノ酸の側鎖が直接入り込んで、認識を行うというものである。これは結局間違いではあったが、アミノ酸と核酸の対応を物理化学的な相互作用に基づく必然性に由来するとした点で画期的であり、その後の遺伝暗号の起源についての立体化学説に大きな影響を与えた。

3. 偶然凍結説と立体化学説 (文献 7-10)

DNA の二重らせんモデルの提唱者の 1 人である Crick は、その後、遺伝暗号の起源についても考察を深めた。Crick は、実験的に核酸には特定のアミノ酸との特異的な結合能が見出されていない状況を考慮し、1955 年に、酵素的に結合したアミノ酸を持ち、対応する核酸コドンと認識するアダプター分子が翻訳を仲介する、という「アダプター仮説」を提唱した。この仮説は、一般的な科学雑誌や学術専門雑誌に公表されたのではなく、「RNA タイムクラブ」のメンバーだけに配られる小冊子に出された。Crick のこの仮説は極めて秀逸であり、その後の実験で、可溶性の小さな RNA が確かにアダプター分子として機能していることが分かった。この RNA が後に、tRNA (transfer RNA、転移 RNA) と呼ばれるようになった。

遺伝暗号の起源に関して、Crick は、遺伝暗号は偶然に決まり、それが進化の過程で凍結されたという「偶然凍結説」を唱えている。しかしながら、彼は、進化の過程における遺伝暗号の成立 (RNA コードとアミノ酸との対応) には何らかの物理化学的な必然性があったという「立体化学説」を否定しているのではない。現に彼は「それゆえに立体化学説を追究するのが本質的なのだ。あいまいな相互作用ではほとんど意味がない。求めるべきは、そのような立体化学的相互作用が確かに起こるとい直接的な証拠であり、それらの特異性を説明するアイデアである。」と述べている。

「偶然凍結説」と対照的な立場をとるのが、「立体化学説」である。1967年に出された「The Genetic Code」という書物の中で、Woese はアミノ酸と核酸塩基のトリプレットの間に存在する立体化学的な対応が遺伝暗号の起源であるという説を提唱した。現に、遺伝暗号表を見ると、その並びはでたらめではないことが分かる。例えば、mRNA 上のコドン 1 文字目と 2 文字目がアミノ酸の特異性に対して重要であり、特に、2 文字目がピリミジン (C または U) の場合、疎水側鎖を持つアミノ酸に対応し、2 文字目がプリン (A または G) の場合、極性側鎖を持つアミノ酸に対応している、などの特徴がある。コドンが直接、アミノ酸を認識した可能性を指摘したものである。

一方、コドンではなく、コドンと相補的である tRNA のアンチコドンがアミノ酸を直接認識し、これが遺伝暗号の起源であると主張する人達もいる。Lacey Jr.らは、アンチコドン 2 文字目、3 文字目のヌクレオチドがアミノ酸と直接相互作用することを示唆する実験結果を示している。特に、アンチコドン 2 文字目、3 文字目と、それがコードするアミノ酸の疎水性には対応関係が見られる。RNA を構成する 4 種類の塩基は A>G>C>U の順に疎水性が小さくなるが、例えば、アンチコドン 2 文字目、3 文字目が AA (疎水性：大) のコードは、フェニルアラニン (疎水性：大) に対応している。

現在の生物系において、RNA とアミノ酸とがはじめて出会うステップは「tRNA のアミノアシル化」反応である。対応する tRNA に、正しく対応するアミノ酸だけを結合させる蛋白質がアミノアシル tRNA 合成酵素 (aaRS) であり、この反応を「tRNA のアミノアシル化」と呼んでいる。この反応において、tRNA のアンチコドンは aaRS の重要な認識部位になっている。「立体化学説」をとる立場の人においても、「コドンとアミノ酸」との直接対応の立場に立つ人と、「アンチコドンとアミノ酸」との直接対応の立場に立つ人と二分されており、確固たる結論は得られていないのが現状である。

4. C4N 仮説 (Complex of four nucleotides) (文献 11-14)

tRNA は、かつて Crick が予言したように mRNA 上の情報をアミノ酸配列に変換するアダプター分子であるが、現在の tRNA には、アンチコドンの他にも特徴的な部分が存在する。すべての tRNA は 3' 末端に CCA という一本鎖の共通配列を持つが、この CCA の隣 (3' 末端から 4 番目) のヌクレオチドは「ディスクリミネーター」と呼ばれ、特徴的な部分である。ディスクリミネーターと呼ばれる所以は、この部分の塩基の分布がランダムではなく、アミノ酸の化学的な性質に関連して、ほぼ決まった塩基が使われていることが多いからである。例えば、疎水性のアミノ酸に対応する tRNA のディスクリミネーターの塩基は A であることが多いのである。現在の生物系において、多くの aaRS は、このディスクリミネーターを、アミノ酸特異性を決めるための重要な認識部位として使用している。

1982 年に、宇宙科学研究所の Shimizu は、「アンチコドンの 3 塩基+ディスクリミネーター塩基」が複合体を形成し、この複合体が作るポケットが、アミノ酸に対して鍵と鍵穴の関係を持つとする、C4N 仮説を提唱した。C4N 複合体は、ディスクリミネーター塩基上の (水素結合可能な) 3 カ所が、各アンチコドン上の (水素結合可能な) 1 カ所の部分と水素結合することによって形成される。250 種類以上もあるアミノ酸の中から、なぜ 20 種

類が選ばれたのかという問題に対して、C4N 仮説は、鍵穴に合う制限のためであるという明確な答えを呈示している。また、ヌクレオチド中の糖が、五炭糖の中でも、なゼリボースに限られているのかという問題に対しても、リボースの 2'-OH がスタックしたアンチコドン塩基の間にクサビのように入り、C4N 全体の構造を安定化しているためであると説明がつく。

C4N において、塩基はアミノ酸をどう分子認識しているのだろうか？親水的なアミノ酸に対しては、アンチコドンの塩基がスタックするので、水素結合の位置が決まり、アミノ酸を直接認識できることになる。一方、疎水的なアミノ酸に対しては塩基が形成する二面性を利用する。つまり、塩基が3つあれば、疎水性の平面部から錐状のポケットが形成され、この面を利用したポケットの疎水性相互作用で、アミノ酸側鎖と相互作用するのである。

C4N は非常に画期的なアイデアであるが、今後、この仮説を実験的に検証して行く必要がある。

5. *In vitro* セレクションを用いたアミノ酸認識 RNA の選択 (文献 15-17)

Cech, Altman らによるリボザイムの発見に引き続き、Szostak, Joyce らによって RNA の試験管内選択の手法が開発された。この手法により、数々の結合活性を持った「アダプター」と呼ばれる人工 RNA が選択取得され、「RNA ワールド仮説」を裏づける証拠の1つと考えられてきた。遺伝暗号の起源を探索する目的でも、この手法が利用され、アミノ酸を直接認識する RNA の人工アダプターが取られてきた。Yarus らによって、アルギニン、チロシン、イソロイシンなどのアミノ酸に対してアダプターが得られたが、それらのアダプターの配列の中には、そのアミノ酸のコドンに対応する配列が見られた。この結果は、立体化学説のうち、「コドンとアミノ酸」との直接対応の立場であり、コドンがアミノ酸を直接認識することによって、遺伝暗号の起源が形成されたという立場である。これらは、先に述べた、アミノ酸の疎水性とアンチコドンの関連やC4Nの結論(「アンチコドンとアミノ酸」との直接対応)とは矛盾している。しかしながら、アダプター中にコドンに対応する配列が見られたとは言え、その出現の判断は極めて恣意的にも見える。いずれにせよ、この辺の問題は更なる実験的追究が必要である。

6. 共進化理論 (文献 18-19)

現在の遺伝暗号は 20 種類のアミノ酸に対する対応関係を表したものであるが、Wong は、遺伝暗号は最初から 20 種類のものに対して決まったのではなく、最初は少数の原始的なアミノ酸が存在し、それらのアミノ酸に対して、原始的な遺伝暗号が出来た、と考えた。

具体的には、グルタミン酸、アスパラギン酸、バリン、セリン、ロイシン、アラニン、グリシンなどであり、その後、その他の遺伝暗号は、アミノ酸の生合成系の進化と共に決まってきた、という考え方である (1975 年)。事実、上に挙げたアミノ酸の生合成は、現在のアミノ酸の生合成過程において数少ないステップ数から構築されている。

7. オペレーショナル RNA コード (文献 20-24)

現在の生物系において、aaRS による tRNA のアミノアシル化の解析が詳細に行われてきて、tRNA のアンチコドンの aaRS の認識部位としての重要性がクロズアップされてきた。しかしながら、tRNA の中にはアンチコドンをまったく認識に用いないものも見出されてきた。その典型的な例が tRNA^{Ala} である。

現在の tRNA は L 字型の構造を有している。しかし、この L 字型のそれぞれの腕に相当する部分は同じように進化してきたのではなく、「ミニヘリックス」と呼ばれる片方の腕に相当する部分が先にでき、その後、アンチコドンを含むもう片方の腕に相当する部分が付け加わったものと考えられる。実際、これまで明らかにされてきた各アミノ酸に対応する aaRS の構造を比較すると、ミニヘリックスを認識する領域は、アミノ酸の種類に関係なく構造が保存されているのに対し、アンチコドン部分を認識する領域の構造は、アミノ酸の種類により大きく異なっていることが分かっている。また、ミニヘリックス部分だけの RNA でも、aaRS の基質になることが多くの場合において明らかにされている。

ミニヘリックスには、アンチコドンに先立って確立された Operational RNA Code というものが存在し、それが遺伝暗号の起源に関わっていると考えられる。

tRNA^{Ala} のミニヘリックスには、G-U というウォブル塩基対が存在し、この塩基対の存在が、アラニンの特異性を決めている。tRNA^{Ala} が最も顕著な例であるが、ミニヘリックスは多くの tRNA 種において、アミノアシル化活性が認められており、ミニヘリックスに存在している構造部位 (Operational RNA Code) はアンチコドンに先立って確立された遺伝暗号である可能性がある。de Duve は、G-U の発見を、The Second Genetic Code (第2の遺伝暗号) と称している。

8. tRNA のアミノアシル化と L-アミノ酸の起源 (文献 25-28)

遺伝暗号の起源を考える前に、生物が使っているアミノ酸が、L-アミノ酸であることの謎を明らかにしておく必要がある。現在の生物における tRNA のアミノアシル化は aaRS によって活性化されたアミノ酸 (アミノアシル AMP) を対応する tRNA に結合させるという 2 段階反応である。ここで注目すべきは、反応中間体のアミノアシル AMP であり、この分

子は前生物学的環境においても生成されることが実験的に示されている。アミノアシルAMPをミミックしたアミノアシル-リン酸-オリゴヌクレオチドを用いて、蛋白質が一切存在しない状態でのアミノアシル化のモデルがTamuraとSchimmelによって作製された。

ミニヘリックス、オリゴヌクレオチド、架橋分子からなる三体が形作る空間が、反応選択性を生み出し、L-アミノ酸選択的なミニヘリックスのアミノアシル化を可能にしている。これまでのさまざまな実験事実から、地球の歴史上において、RNAワールドというものが存在したのは、まず間違いないと思われる。そして、はじめに「RNAワールド」ありき、という立場に立つと、現在の生物系がなぜL-アミノ酸を用いているのかという、アミノ酸のホモキラリティーの起源について、極めてはっきりしたストーリーとして描き出すことができる。

9. おわりに (文献 29)

DNA 二重らせんモデルの提唱者の1人であるWatsonは、その著書「Genes, Girls, and Gomow」の中で、次のように書いている。

「とにかく先には、まだまだ突き止めなくてはならない遺伝学上の大きな獲物が横たわっているのだ。そのさらなる追跡に参加しないなどは、まさに毫釐ぶりを晒すようなものだった。目指す先には「生命のロゼッタ石」すなわち遺伝コードがあり、それこそがDNA分子内に符号化されている遺伝情報を、生きた細胞の分子レベルの働き手であるタンパク質の言語に翻訳するルールを、ぼくらに示してくれるはずなのだ。」

DNAの二重らせんモデルの解明、そして、遺伝暗号の解明の後、かなりの年月が過ぎた今日でさえ、遺伝暗号の本質に関する理解は十分ではない。遺伝暗号はあくまでも、まだ「暗号」であり、それがどのように、地球上の生命に備わってきたのかということに関しては、やはり謎のままである。真の意味において、「生命のロゼッタ石」はまだ解読されてはいないのである。現代の最先端の生物学においても、この対応関係をすべて「暗号」のまま置き去りにしたまま、研究が進められているのが現状である。

しかし、私達は地球上で進化してきた生命体である。その過程では、遺伝暗号がなぜそのように決まったのかという必然性を経ているはずである。今後の実証的な科学研究が、新たな英知の創造につながることを期待したい。

参考文献

1. Kruger K., Grabowski P. J., Zaug A. J., Sands J., Gottschling D. E. and Cech T. R. (1982) Self-splicing RNA: autoexcision and autocyclization of the

ribosomal RNA intervening sequence of Tetrahymena. *Cell* 31: 147-157

2. Guerrier-Takada C., Gardiner K., Marsh T., Pace N. and Altman S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35: 849-857

3. Gilbert W. (1986) The RNA world. *Nature* 319: 618

4. Watson J. D. and Crick F. H. C. (1953) Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* 171: 737-738

5. International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431: 931-945

6. Gamow G. (1954) Possible relation between deoxyribonucleic acid and protein structures. *Nature* 173: 318

7. Woese C. R. (1967) *The Genetic Code*. Harper & Row, New York

8. Crick F. H. C. (1968) The origin of the genetic code. *J. Mol. Biol.* 38: 367-379

9. Crick F. H. C. (1955) On degerate templates and the adaptor hypothesis. A note for the RNA Tie-Club

10. Lacey Jr. J. C. and Mullins Jr. D. W. (1983) Experimental studies related to the origin of the genetic code and the process of protein synthesis. *Origins of Life* 13: 3-42

11. Crothers D. M., Seno T. and Söll D. G. (1972) Is there a discriminator site in transfer RNA? *Proc. Natl. Acad. Sci. USA* 69: 3063-3067

12. Shimizu M. (1982) Molecular basis for the genetic code. *J. Mol. Evol.* 18: 297-303

13. 清水幹夫 (1989) RNA のアミノ酸認識仮説. *日経サイエンス* 4月号: 68-79

14. Szathmáry E. (1993) Coding coenzyme handles: A hypothesis for the origin of the genetic code. *Proc. Natl. Acad. Sci. USA* 90: 9916-9920

15. Szostak J. W. (1992) In vitro genetics. *Trends Biochem. Sci.* 17: 89-93

16. Joyce G. F. (2002) The antiquity of RNA-based evolution. *Nature* 418: 214-221

17. Yarus M. (1998) Amino acids as RNA ligands: a direct-RNA-template theory for the code's origin. *J. Mol. Evol.* 47: 109-117

18. Wong J. T. (1975) A co-evolution theory of the genetic code. *Proc. Natl. Acad. Sci. USA* 72: 1909-1912

19. Wong J. T. (1981) Coevolution of genetic code and amino acid biosynthesis. *Trends Biochem. Sci.* FEB: 33-36

20. Schimmel P. (1987) Aminoacyl tRNA synthetases: general scheme of structure-function relationships in the polypeptides and recognition of transfer RNAs. *Annu. Rev. Biochem.* 56: 125-158

21. Ibba M. and Söll D. (2000) Aminoacyl-tRNA synthesis. *Annu. Rev. Biochem.* 69: 617-650

22. Schimmel P., Giegé R., Moras D. and Yokoyama S. (1993) An operational RNA code for amino acids and possible relationship to genetic code. *Proc. Natl. Acad. Sci. USA* 90: 8763-8768

23. Musier-Forsyth K. and Schimmel P. (1999) Atomic determinants for aminoacylation of RNA minihelices and relationship to genetic code. *Acc. Chem. Res.* 32: 368-375

24. de Duve C. (1988) Transfer RNAs: the second genetic code. *Nature* 333: 117-118

25. Paecht-Horowitz M. and Katchalsky A. (1973) Synthesis of amino acyl-adenylates under prebiotic conditions. *J. Mol. Evol.* 2: 91-98

26. Tamura K. and Schimmel P. (2004) Chiral-selective aminoacylation of an RNA minihelix. *Science* 305: 1253

27. Tamura K. and Schimmel P. R. (2006) Chiral-selective aminoacylation of an RNA minihelix:

- Mechanistic features and chiral suppression. Proc. Natl. Acad. Sci. USA 103: 13750-13752
28. 田村浩二 (2007) RNA ワールドから見たアミノ酸のホモキラリティーの誕生. 生物の科学 遺伝 61: 92-92
29. Watson J. D. (2001) Genes, Girls, and Gamow. Oxford University Press, Oxford