

論文 / 著書情報
Article / Book Information

題目(和文)	トリプファンを含まない遺伝暗号表およびタンパク質の進化に関する研究
Title(English)	
著者(和文)	河原晃大
Author(English)	Akio Kawahara
出典(和文)	学位:博士(理学), 学位授与機関:東京工業大学, 報告番号:甲第9682号, 授与年月日:2014年12月31日, 学位の種別:課程博士, 審査員:木賀 大介,山村 雅幸,中村 清彦,木川 隆則,瀧ノ上 正浩
Citation(English)	Degree:., Conferring organization: Tokyo Institute of Technology, Report number:甲第9682号, Conferred date:2014/12/31, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

学位論文

トリプトファンを含まない遺伝暗号表
およびタンパク質の進化に関する研究

指導教員 木賀 大介 准教授

氏名 河原 (小林) 晃大

目次

目次.....	1
略号、記号および備考	3
第一章 序章.....	5
1.1. 生命の起源と進化にみられる複雑化.....	5
1.2. タンパク質の進化における複雑化と配列最適化.....	5
1.3. 遺伝暗号表の進化における複雑化.....	6
1.4. 過去の生命現象を推測する方法	7
1.5. トリプトファンを含まない 19 種類のアミノ酸セットで遺伝暗号表とタンパク質を構築する意義.....	8
1.6. 図表.....	10
第二章 トリプトファンを含まない 19 種類のアミノ酸セットからなるタンパク質の人工進化.....	14
2.1. 導入.....	14
2.1.1. トリプトファンを含まないタンパク質およびその人工進化に関する先行研究	14
2.2. 材料と方法	14
2.2.1. DNA コンストラクション.....	14
2.2.2. フローサイトメーターによる解析とセルソーターによる細胞分取	16
2.2.3. タンパク質発現・精製.....	16
2.2.4. 精製されたタンパク質変異体の蛍光強度測定.....	17
2.3. 結果.....	17
2.3.1. トリプトファンを置換したことによる蛍光強度の低下と、置換部位周辺への Saturation mutagenesis	17
2.3.2. 1 ラウンド目の Directed evolution から得られた活性変異体の解析.....	18
2.3.3. 2 ラウンド目の Directed evolution から得られた活性変異体の解析.....	18
2.3.4. Directed evolution によって得られた変異が与える影響の解析.....	18
2.4. 結果のまとめと考察	19
2.5. 図表.....	22
第三章 トリプトファンを含まない 19 種類のアミノ酸セットからなる遺伝暗号表の構築.....	34
3.1. 導入.....	34
3.1.1. 天然および実験室内における拡張遺伝暗号表に関する先行研究.....	34
3.2. 材料と方法	34
3.2.1. DNA コンストラクションおよび <i>in vitro</i> 転写.....	34
3.2.2. 無細胞翻訳系	35

3.2.3.	タンパク質精製	35
3.2.4.	放射標識されたタンパク質の検出	35
3.2.5.	アミノ酸組成分析	36
3.2.6.	トリプトファン蛍光スペクトル測定	37
3.2.7.	無細胞翻訳系中における GFP の蛍光強度測定	37
3.2.8.	結晶化、および X 線回折データセットの収集	37
3.2.9.	原子モデルの構築、および構造の精密化	38
3.2.10.	LexA 切断アッセイ	38
3.3.	結果	39
3.3.1.	トリプトファンを含まない遺伝暗号表による翻訳の確認	39
3.3.2.	合成されたタンパク質がトリプトファンを含まないことの確認	39
3.3.3.	トリプトファンを含まない遺伝暗号表が普遍遺伝暗号表と同等の効率でタンパク質を合成できることの確認	40
3.3.4.	トリプトファンを含まない遺伝暗号表から翻訳されたタンパク質の、全体の構造および UGG コドンへの割り当ての結晶構造解析による確認	40
3.3.5.	19 種類のアミノ酸セットからなる遺伝暗号表の構築原理の一般性の確認	41
3.4.	結果のまとめと考察	41
3.5.	図表	44
第四章	総合討論	58
4.1.	結果のまとめ	58
4.2.	考察	58
4.2.1.	トリプトファンが遺伝暗号表に導入された意義	58
4.2.2.	遺伝暗号表の違いが水平伝播の障壁になる	59
4.3.	展望	60
4.3.1.	アミノ酸の種類が減少した遺伝暗号表を用いた Directed evolution	60
4.3.1.	トリプトファンを含まない 19 種類のアミノ酸セットからなる生物の創出	61
4.4.	図表	62
参考文献	65
発表論文	72
謝辞	73

略号、記号および備考

アミノ酸の略号

A, Ala	alanine	アラニン
C, Cys	cysteine	システイン
D, Asp	aspartic acid	アスパラギン酸
E, Glu	glutamic acid	グルタミン酸
F, Phe	phenylalanine	フェニルアラニン
G, Gly	glycine	グリシン
H, His	histidine	ヒスチジン
I, Ile	isoleucine	イソロイシン
K, Lys	lysine	リジン
L, Leu	leucine	ロイシン
M, Met	methionine	メチオニン
N, Asn	asparagine	アスパラギン
P, Pro	proline	プロリン
Q, Gln	glutamine	グルタミン
R, Arg	arginine	アルギニン
S, Ser	serine	セリン
T, Thr	threonine	スレオニン
V, Val	valine	バリン
W, Trp	tryptophan	トリプトファン
Sec	selenocysteine	セレノシステイン
Pyl	pyrrolysine	ピロリジン

核酸の略号

A	adenine	アデニン
C	cytosine	シトシン
G	guanosine	グアニン
T	thymine	チミン
U	uridine	ウリジン
K	G, T/U	グアニンまたはチミン/ウリジン
S	G,C	グアニンまたはシトシン
N	A,C,G,U	アデニンまたはシチジンまたはグアノシンまたはウリジン

アミノアシル tRNA 合成酵素の略号

aaRS aminoacyl-tRNA synthetase アミノアシル tRNA 合成酵素

各 aaRS の略号は各アミノ酸の 3 文字表記を用いて表した。

例) アラニル tRNA 合成酵素の場合、AlaRS とした。

転移 RNA の略号

tRNA transfer RNA 転移 RNA

各 tRNA の略号は各アミノ酸の 3 文字表記を用いて表した。

例) アラニル tRNA の場合、tRNA^{Ala} とした。

アミノ酸の位置の略号

アミノ酸の位置は、各アミノ酸の 3 文字表記に上付き文字を用いて表した。

例) 57 番目のトリプトファンの場合、Trp⁵⁷ とした。

その他の略号

DNA deoxyribonucleic acid デオキシリボ核酸

RNA ribonucleic acid リボ核酸

第一章 序章

1.1. 生命の起源と進化にみられる複雑化

生命の起源と進化に関する研究は、単純なものから複雑なものへとシステムが発展していくことを明らかにしてきた。DNA が遺伝情報を持ち、タンパク質が触媒活性を持つ現在の翻訳機構は、もともとは RNA が遺伝情報と触媒活性を合わせ持っていた単純な機構から移行したものだという RNA world 仮説が、1980 年初期のリボザイムの発見(1,2)によって有力になった。RNA が遺伝情報と触媒活性を担っていた時代から、RNA の触媒活性をタンパク質が担う RNA-Protein world に移行し、さらには RNA の遺伝情報を DNA が担う DNA-Protein world へ移行したと考えられている。触媒機能のタンパク質への移行と、遺伝情報の DNA への移行のうち、特に前者はシステムの複雑化を招く。RNA は 4 種類の塩基配列を持ついっぽう、タンパク質は 20 種類の様々な化学的性質を持つアミノ酸を持ち、触媒できる反応が多様化したと考えられる。また、進化における複雑化の別の例として、細胞構造の複雑化を挙げることができる。真核生物は、原核生物(古細菌あるいは古細菌と真核生物の共通祖先)から 22 億年前に進化したとされている。真核生物の細胞内には、原核生物の細胞内には見られない核膜、ミトコンドリア、ゴルジ体、小胞体などの細胞小器官が存在する。これらの細胞小器官の中の膜で囲まれた空間は、細胞内の他の部分と隔離され、それぞれ異なるタンパク質とイオンの組成を持った空間を作り、1つの細胞の中で様々な機能分担を可能にしている。原核生物から進化した真核生物が、多様な種に分化し地球上に広まったことを考えると、細胞構造の複雑化は生存において有利な特性だったと思われる。以上のように、進化において、単純なものから始まった生命システムは、実現できる機能を拡張し生存により有利になるために複雑化してきたと言える。

1.2. タンパク質の進化における複雑化と配列最適化

20 種類のアミノ酸セットを構成要素とする現在のタンパク質は、進化の初期においては 7-13 種類のみのアミノ酸セットから構成されていたと考えられている(3,4)。この主張は、現在のタンパク質に用いられる 20 種類のアミノ酸セットのうち、原始環境においては限られたアミノ酸しか存在しなかったであろうとの推測に基づいている。原始環境におけるアミノ酸の存在を推測した報告として、還元大気(5)や中性大気(6)を模した環境での放電実験、熱水噴出孔の環境を模した実験(7,8)がある。また、隕石の成分からもアミノ酸が検出されることが分かっている(9)。しかし、それぞれの系において、検出されるアミノ酸の種類に小さな違いがみられるものの、現在のタンパク質に用いられるアミノ酸セットの 20 種類すべてが生成・発見されることはない。特に、アルギニンなどの塩基性アミノ酸や、トリプトファンなどの芳香族アミノ酸はほとんど生成されない。タンパク質が 10 種類前後のアミノ酸セットで機能することは、9 種類のアミノ酸で活性を持つ AroQ chorismate mutase を構成した研究(10)や、13 種類のアミノ酸で活性を持つ Orotate phosphoribosyl transferase を構成した研究(3)によって実証されている。

タンパク質を構成するアミノ酸セットの種類は、細胞の成立後、生合成経路の発達によって徐々に増加したと考えられている。Wong は、現在タンパク質に用いられている 20 種類のアミノ酸セットのうち、原始環境を模した実験から生成されないアミノ酸の由来を、生合成経路の発達によって説明した(11)。例えば、原始環境中に豊富に存在したと考えられるセリンを反応の出発点として、トリプトファンとシステインを生合成し獲得できることが、現在のアミノ酸代謝系から理解できる。また別の研究からは、ヒスチジンの生合成に関わる酵素である HisA に対してアミノ酸置換を 1 か所だけ導入することで、トリプトファンの生合成に関わる酵素である TrpF の活性を持たせることができることが示されており(12)、新規アミノ酸の代謝は既存のアミノ酸の代謝に関わる遺伝子をもとに実現できた可能性が示唆されている。

1.3. 遺伝暗号表の進化における複雑化

遺伝暗号表が、タンパク質の世界と核酸の世界をつなぐ。DNA にコードされた遺伝情報は mRNA に転写され、遺伝暗号表 (図 1.1.)に従ってアミノ酸に翻訳(図 1.2.)される。翻訳の過程において、mRNA 上の 3 文字の塩基であるコドンに、アミノ酸が結合した tRNA が対合することで、核酸の情報がアミノ酸の情報に変換される。mRNA と tRNA の正しい結合を保証しているのは核酸塩基どうしの水素結合であり、いっぽうで tRNA とアミノ酸の正しい結合を保証しているのはアミノアシル tRNA 合成酵素 (aaRS)である。tRNA および aaRS はアミノ酸ごとに専用の分子が存在し、mRNA 上の遺伝情報が間違っても翻訳されないよう、厳格な対応付けを実現している。大腸菌を用いた研究では、コドンに対するアミノ酸のタンパク質への誤った導入は、 10^{-3} から 10^{-4} に抑えられていることが報告されており(13,14)、進化の過程で翻訳系、特に aaRS がより正確性を増すように最適化されたことを想像させる。

遺伝暗号表に含まれるアミノ酸の数は変わりうるものが、天然で発見された例および実験室内で構築された例で知られている。遺伝暗号表のアミノ酸セットの数や配置はほとんどすべての生物に共通であることから、遺伝暗号表は「普遍」遺伝暗号表と呼ばれた(15)。しかし、普遍遺伝暗号表には含まれない 21 番目のアミノ酸であるセレノシステイン(16)や 22 番目のアミノ酸であるピロリジン(17)が見つかり、遺伝暗号表に含まれるアミノ酸の種類は天然において変わりうる事が分かった。また、aaRS と tRNA を改変し、非天然アミノ酸をタンパク質に導入できることが報告され(18,19)、普遍遺伝暗号表には含まれないアミノ酸を遺伝暗号表に導入できることが実験室においても報告された。aaRS と tRNA の改変によってタンパク質に導入された非天然アミノ酸の種類は 70 を超えている(20)。

20 種類のアミノ酸セットからなる普遍遺伝暗号表は、19 種類以下のアミノ酸セットからなる単純な形から進化してきたと考えられている。1.2節で述べたように入手できるアミノ

酸の種類が限られていた時期の遺伝暗号表は、当然アミノ酸の種類に制約を受ける。限られたアミノ酸、tRNA、aaRS からなる遺伝暗号表は、生合成経路の発達によって入手できるアミノ酸の種類が増えると、aaRS と tRNA の多様化によってアミノ酸セットの数を増やしたと考えられる。実際に、aaRS の配列をもとにした系統樹(21)からは、現在 20 種類存在する aaRS が徐々に分岐してきたことが明確に分かる。Ikehara らは、遺伝暗号表が GUC (バリン)、GCC (アラニン)、GAC (アスパラギン酸)、GGC (グリシン)の 4 種類のみでコードされる GNC コードで始まり、中間段階として 16 種類のアミノ酸セットからなる SNS コードを経由して、普遍遺伝暗号表に至ったとする仮説を提唱している(22)。

トリプトファンは、20 種類のアミノ酸セットのうち、遺伝暗号表に加わった最後のアミノ酸であると考えられている。遺伝暗号表の進化に伴ってアミノ酸がどのような順番で遺伝暗号表に加わってきたかという問題は、古くから研究者の興味を惹き付けるテーマである(15)。これまでに様々な研究者が多様な視点からアミノ酸が遺伝暗号表に加わった順番についての考察を行ってきたが、様々な観点から行われた 60 個の先行研究をまとめて評価し、20 種類のアミノ酸それぞれが遺伝暗号表に導入された順番を決めた Trifonov による研究(23)が最も包括的な研究と言える。Trifonov の研究の中で、トリプトファンは遺伝暗号表に加わった最後のアミノ酸として結論付けられている(表 1.1.)。用いられた 60 個の先行研究のうち、生合成経路の複雑さや aaRS の系統樹など、60 個の先行研究のうち 15 個がトリプトファンを 19 番目あるいは 20 番目のアミノ酸としていることから、トリプトファンが遺伝暗号表に最後に導入されたアミノ酸であるという推測は現在、高い確度で真であると言えるだろう。

1.4. 過去の生命現象を推測する方法

過去の生命現象を推測するためには、現存する生物の遺伝子配列から系統樹をもとに過去の遺伝子配列を推測する方法や、過去の生命現象と推測される反応を現在まで保持している生物を発見し調べる方法がある。前者の例として、祖先型のヌクレオシド二リン酸キナーゼ (NDK)のアミノ酸配列を推定し、推定された配列から作成したタンパク質の耐熱性を調べた研究(24)がある。Akanuma らは、現存する複数種の生物の NDK 配列から系統樹を作成し、系統樹の根元付近の古細菌祖先に相当するアミノ酸配列を 3 つと、真正細菌祖先に相当するアミノ酸配列を 3 つ推定し、それぞれの遺伝子を合成した。その結果、6 つの祖先型 NDK はすべて、100°C以上まで変性しない高い耐熱性を持つタンパク質であることが明らかとなった。NDK の変性温度は、その NDK を持っていた生物の至適生育温度と正の相関関係があることが分かっていたため、古細菌の祖先、および真正細菌の祖先は、90°C以上の超好熱菌であったことが推定された。後者の例として、現在のアミノアシル化反応が完成する前の古い反応であると考えられている、間接的なアミノアシル化の反応経路が挙げられる。すべての古細菌と真正細菌の一部は 20 種類のアミノ酸セットを遺伝暗号表に

含みながらも、アスパラギニル tRNA 合成酵素(AsnRS)とグルタミニル tRNA 合成酵素(GlnRS)の両方、あるいはどちらか片方を欠いている。このような生物では、AspRS と GluRS のほか、非識別型の AspRS(ND-AspRS)と GluRS(ND-GluRS)を持つ。ND-AspRS は、tRNA^{Asp} と tRNA^{Asn} の両方にアスパラギン酸を結合させ、ND-GluRS は、tRNA^{Glu} と tRNA^{Gln} の両方にグルタミン酸を結合させる。その後、tRNA とアミノ酸がミスマッチした Asp- tRNA^{Asn} および Glu- tRNA^{Gln} は、tRNA 依存的なアミノ基転移酵素 GatCAB(古細菌の Gln 形成には GatDE)によって Asn- tRNA^{Asn} および Gln- tRNA^{Gln} に変換される(25-28)。同様に、メタン生成古細菌にはシステニル tRNA 合成酵素(CysRS)が存在せず、O-ホスホリル tRNA 合成酵素(SepRS)、O-ホスホセリン-システイン合成酵素(SepCysS)という酵素がシステインの間接的な導入を担っている(29)。このような間接的なアミノアシル化は、tRNA 依存的なアミノ基転移酵素が進化的に古いことが系統樹解析から分かった(30)こともあり、AsnRS、GlnRS、CysRS が関与する直接的なアミノアシル化の経路が完成するまでの進化の途上であると考えられている。

現存する生物を材料に行われる研究とは異なる手法として、現存しない分子や分子システムを実験的に創出する手法が、過去の生命システムについて考察する有効な手段となることがある。系統樹の解析や過去の反応を保持している生物の観察は、どちらも進化研究における重要な手法であるものの、現存する生物に研究対象が絞られるという点に限界がある。なぜなら、非常に時間の離れた過去を推測するための手掛かりとなる情報・材料は、現在に充分残っているとは限らないからである。この限界を補完するためには、現存する情報・材料から得られる仮説を立て、その仮説を反映させた系を実験的に構築することが望ましい。もちろん、仮説を反映させた系が過去に存在した系と完全一致しているかどうかは立証不可能だが、想定する系が物理化学的な現象として機能するかどうかを観察することは、過去の推測のための手掛かりを与えることになるだろう。

1.5. トリプトファンを含まない 19 種類のアミノ酸セットで遺伝暗号表とタンパク質を構築する意義

トリプトファンを含まない 19 種類のアミノ酸セットからなる天然のタンパク質は限られており、また遺伝暗号表にいたっては発見されていないため、過去に存在した可能性のある翻訳系について考察を深めるために、これらを実験的に創ることは有用であると考えた。まず、トリプトファンを含まない 19 種類のアミノ酸セットからなるタンパク質が、トリプトファンなしに変異を蓄積して活性を向上することを確認するために、トリプトファンを含まない緑色蛍光タンパク質(GFP)変異体を作成し、変異と選択を繰り返す Directed Evolution で活性向上するかどうかを確かめる。次に、トリプトファンを含まない 19 種類のアミノ酸セットからなる遺伝暗号表が機能することを確認するために、トリプトファンを含まない遺伝暗号表である「単純化遺伝暗号表」(図 1.3.)を作成する。

本論文は、4章で構成される。第二章では、トリプトファンを含まない19種類のアミノ酸セットからなるタンパク質の人工進化について述べる。トリプトファンが遺伝暗号表に導入される前に存在したであろう19種類のアミノ酸からなるタンパク質が進化できたであろうことは、トリプトファンを含まない野生型タンパク質の人工進化によって確認されている。しかし、トリプトファンを含まない野生型タンパク質は稀であり、人工進化の知見を蓄積するためには進化実験の対象を広げる必要がある。私は、20種類のアミノ酸からなるタンパク質に含まれるトリプトファンを置換し、変異と選択を繰り返す **Directed evolution** を用いることで、トリプトファンを含まない19種類のアミノ酸セットからなるタンパク質の活性が向上することを示す。第三章では、トリプトファンを含まない19種類のアミノ酸セットからなる遺伝暗号表の構築について述べる。トリプトファンを含まない19種類以下のアミノ酸セットからなる遺伝暗号表は、進化の途上で存在したと考えられているが、今のところ現存するものは見つかっておらず、また実験的に構築可能かどうかも確かめられていない。私は、反応系中からトリプトファンを除き、かつトリプトファンコドンに対応するアンチコドンを持つ tRNA^{Ala} 変異体を加えることでトリプトファンを含まない19種類以下のアミノ酸セットからなる遺伝暗号表を構築・機能の確認をし、トリプトファンを含まない遺伝暗号表は成立しえたことを示す。第四章では、第二章と第三章で得られた結果を総合し、トリプトファンが遺伝暗号表に導入された意義について考察する。また、将来展望として、アミノ酸の種類が減少した遺伝暗号表を用いた **Directed evolution** がアミノ酸の種類が減少したタンパク質の作成の効率を高め、タンパク質の進化について考察するための対象を増やすことに寄与することについて述べる。最後に、遺伝暗号表の後期の進化の理解に貢献する、トリプトファンを含まない19種類のアミノ酸セットからなる生物の創出の技術的な課題と提案について述べる。

1.6. 图表

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

图 1.1. 普遍遗传暗号表

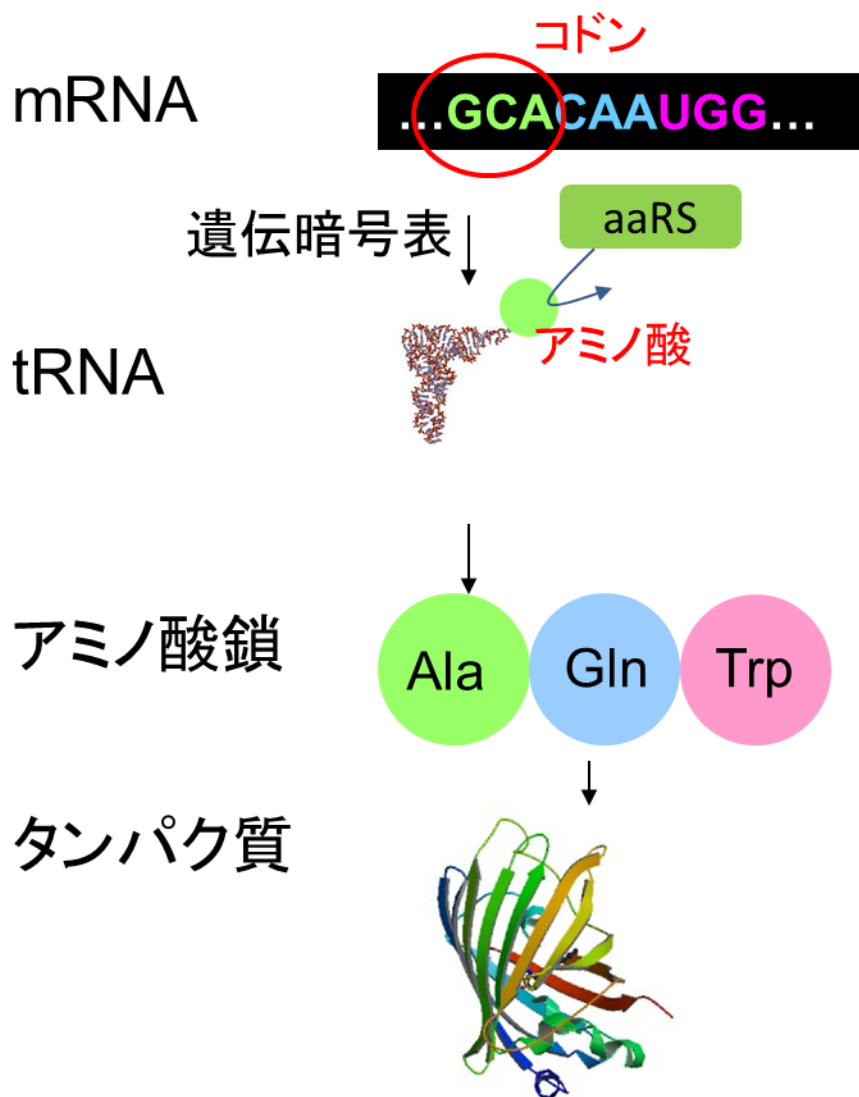


図 1.2. 翻訳系

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

図 1.3. 単純化遺伝暗号表

表 1.1. Trifonov (2004)による、アミノ酸が遺伝暗号表に加わった順番
 左の列にスコアの順番、右の列にアミノ酸の名前を示してある。

1	グリシン
2	アラニン
3	アスパラギン酸
4	バリン
5	プロリン
6	セリン
7	グルタミン酸
8	トレオニン
9	ロイシン
10	アルギニン
11	アスパラギン
12	イソロイシン
13	グルタミン
14	ヒスチジン
15	リジン
16	システイン
17	フェニルアラニン
18	チロシン
19	メチオニン
20	トリプトファン

第二章 トリプトファンを含まない 19 種類のアミノ酸セットからなるタンパク質の人工進化

2.1. 導入

2.1.1. トリプトファンを含まないタンパク質およびその人工進化に関する先行研究

遺伝暗号表を構成するアミノ酸セットにトリプトファンが含まれる以前には、変異と選択によるタンパク質の配列最適化はトリプトファンを含まないアミノ酸セットでなされていたはずである。タンパク質の配列最適化は、遺伝子全体に起こるランダムな変異と、タンパク質として適応度の高いアミノ酸配列が選択される過程によって起こる。20 種類より少ないアミノ酸セットからなるタンパク質を、部位特異的変異導入法である **Site-directed mutagenesis** を用いて作成した研究は存在する(31,32)ものの、遺伝子全体を対象としたランダムな変異を用いて、トリプトファンを含まないタンパク質の配列最適化を行った研究は稀である。

トリプトファンを含まないタンパク質の配列最適化の先行研究として、大腸菌のシャペロンタンパク質である **GroEL** および **GroES** の研究がある。天然の **GroEL** および **GroES** は、アミノ酸配列中にトリプトファンを含まない。**GroEL** および **GroES** の人工進化の先行研究として、**Wang** らはトリプトファンを新たに導入することなく、**GroEL** および **GroES** の、**GFP** のフォールディングを助ける能力を、遺伝子全体への変異導入を伴う **Directed evolution** を用いて高めた(33)。

人工進化ではないものの、**Wang** らはやはりトリプトファンを新たに導入することなく、**GroEL** の **apical domain** に相当するミニシャペロンの安定性を高めた。安定性に寄与する 6 つの変異を **GroEL** のファミリーの配列比較を行うことで特定し、それらを同時に **GroEL** に導入することで安定性を高め、さらにそれらの変異が加算的に安定性を高めることを示した(34)。

トリプトファンを含まないタンパク質の配列最適化を、多くのモデルタンパク質で試みることは、研究対象を広げるために意義がある。トリプトファンを含まないモデルタンパク質は比較的少ない。トリプトファンを含まないタンパク質の配列最適化に関わる現象を、タンパク質の多様な機能と結び付けて確かめるためには、天然に多数存在する 20 種類のアミノ酸セットからなるタンパク質に対してトリプトファンを他のアミノ酸に置換する変異を導入し、トリプトファンを含まない人工的に作成したタンパク質を対象に研究を行えばよい。

2.2. 材料と方法

2.2.1. DNA コンストラクション

DNA コンストラクションに使用したプライマーの配列は表 2.1.に記述した。Site-directed mutagenesis は、Gene Tailor site-directed mutagenesis system (Life Technologies)に基づいて行った。

GFP 変異体をコードしたプラスミド DNA の作成のため、pGFPS1(5)上の GFPS1 のコード配列を、Placq-Bgl_II-RBS-F and Placq-Spe_I-R (Table S1)を用いて増幅した。増幅された DNA 断片は Placq-GFP (6)の Bgl II サイトと Spe I サイトの間に挿入され、Placq-GFP の GFP コード配列は GFPS1 のコード配列に置き換えられた。この DNA コンストラクトを本論文では Placq-GFPS1 と呼ぶ。Placq-S1-W57F および Placq-S1-W57Y を作成するために、Forward プライマーとして W57F-F または W57Y-F を、Reverse プライマーとして W57FY-R を用いて、Placq-GFPS1 への Site-directed mutagenesis を行った。Phe57 に隣接するアミノ酸に変異を導入するため、縮重したコドン (NNK コドン: N = G/A/T/C および K = G/T) を含むオリゴヌクレオチドを用いて、S1-W57F への Saturation mutagenesis を行った。Val29 への変異導入のためには、プライマーペア V29NNK-F および V29NNK-R を用いた。同様に、Phe46、Leu53、Leu60、そして Leu64 への 1 ポイント変異のためには、それぞれのプライマーペア F46NNK-F と F46NNK-R、L53NNK-F と L53NNK-R、L60NNK-F と L60NNK-R、L64NNK-F と L64NNK-R を用いた。連続する 3 つのアミノ酸、Asp216、His217、Met218 への変異導入のためには、プライマーペア D216H217M218NNK-F and D216H217M218NNK-R を用いた。

1 ラウンド目および 2 ラウンド目のプラスミドライブラリの構築のために、過去に報告のあった方法(7)に従って Error-prone PCR を S1-W57F のコード領域に対して行った。通常、50 μ L の反応溶液には 10 mM Tris-HCl [pH 8. 3], 50 mM KCl, 3 mM MgCl₂, 0.5 mM MnCl₂, 0.2 mM of each dATP and dGTP, 1.0 mM of each dTTP and dCTP, 0.2 μ M of each primer set, 1 ng of template DNA, and 1.25 U of rTaq DNA polymerase (TOYOBO)が含まれる。クローニングのために使用したプライマーおよび手順は、Placq-GFPS1 のクローニングのために使用したそれらと同じである。Error-prone PCR の変異導入頻度は、1,000 塩基中に 6.7 変異であることを確認した。ランダムな変異を持った遺伝子からなるプラスミドライブラリは、DH5 α にエレクトロポレーションで導入された。変異体ライブラリは、FACS によって選別された。FACS による選別に用いられたライブラリの通常のサイズは 106 であった。

精製された変異体タンパク質の蛍光強度を測定するために、ヒスチジンタグ配列とリンカー配列を含む以下のアミノ酸配列 YRYEFQLSAASKLAAALEHHHHHH を、それぞれの変異体タンパク質の C 末端へ導入した。ヒスチジンタグを Placq-GFPS1 由来の変異体に導入するため、以下の 3 ステップを私は行った。

ステップ 1 では、GFPS1 の配列の後ろに ClaI 認識配列、そしてヒスチジンタグが続くコード配列、GFPS1-Cla_I-His を作成した。GFPS1-Cla_I-His 作成のために、プライマーペア pGFP-Cla_I-His-F および pGFP-Cla_I-His-R を用いて Site-directed mutagenesis を行い、まず始めに Cla_I 認識サイトを、次にヒスチジンタグ配列を pGFPS1 に導入した。正方向プライマーは GFPS1 のコード配列とヒスチジンタグの間に Cla_I 認識サイトを持ち、最終ステップ

にて望みのクローニングを可能にしている。GFPS1 のコード配列と、読み枠のずれたヒスチジンタグ配列の間に新規のヌクレオチドを挿入することで、正方向プライマーはヒスチジンタグを GFPS1 のコード配列の読み枠と合わせる役割も果たす。

ステップ 2 では、Placq-GFPS1 上の GFPS1 のコード配列は GFPS1-Cla_I-His に置換された。GFPS1-Cla_I-His の PCR 断片はプライマーペア pGFPS1-F および GFP-His Spe_I-R を用いて pGFPS1-Cla_I-His から増幅された。PCR 断片は GFPS1-Cla_I-His の開始コドンに NdeI 認識サイトを、GFPS1-Cla_I-His の終始コドンの隣に Spe I 認識サイトを持つ。PCR 断片は Placq-GFPS1 上の GFPS1 のコード配列を GFPS1-Cla_I-His で置き換えるかたちで Placq-GFPS1 の NdeI 認識サイトと SpeI 認識サイトの間にクローニングされた。

最後の 3 番目のステップとして、EGFP および、セクションによって得られた GFPS1 由来の変異体のコード配列は PCR によって増幅され、GFPS1 のコード配列を変異体の配列で置き換えるかたちで Placq-GFPS1 上の NdeI 認識サイトと SpeI 認識サイトの間にクローニングされた。この操作のため、私は負方向プライマー GFP-cloning-Cla_I-R と、Placq-GFPS1 の作成の時に用いたものと同じ正方向プライマー Placq-Bgl_II-RBS-F を用いた。

2.2.2. フローサイトメーターによる解析とセルソーターによる細胞分取

フローサイトメーターによるすべての解析は、FACSCalibur (Becton-Dickinson)にて、488 nm のレーザーと 515 から 545 nm のエミッションフィルターを用いて行った。細胞のスループットは毎秒 3,000 個に設定した。Saturation mutagenesis によって作成されたライブラリからは、106 個の細胞を解析した。Error prone PCR によって作成されたライブラリからは、 2×10^7 個の細胞を解析した。蛍光の検出チャンネルにゲートを設定し、蛍光強度が強い細胞のみを分取する条件を設定した。また、前方散乱光と側方散乱光の検出チャンネルにもゲートを設定し、コンタミネーションによる粗大な粒子が誤って蛍光強度の強い細胞として分取されるのを防いだ。分取された細胞を遠心分離し、LB 液体培地に回収した。その後細胞を 30 $\mu\text{g}/\mu\text{L}$ のカナマイシンが添加された LB 寒天プレートに播種し、37°C で一晩静置培養した。そして生まれたコロニーをピックアップし、液体培養し、プログラム CellQuest (Becton Dickinson) および WinMDI 2.9 (<http://en.bio-soft.net/other/WinMDI.html>) を用いて解析した。

2.2.3. タンパク質発現・精製

ヒスチジンタグを含む GFP の変異体遺伝子を持ったプラスミドで、大腸菌株 DH5 α を形質転換した。シングルコロニーをピックアップした後、30 $\mu\text{g}/\mu\text{L}$ のカナマイシンが添加された LB 液体培地中にて一晩 37°C で振とう培養した。一晩培養された培地の 1 mL を、30 $\mu\text{g}/\mu\text{L}$ のカナマイシンが添加された 100 mL の LB 液体培地に移し、37°C でさらに振とう培養を続けた。590 nm の吸光度が 0.6 に達した時点で培養温度を 26°C に下げた上でさらに 4 時間の培養を続け、タンパク質発現を継続した。次に、7,000 g、4°C、5 分間の条件で細胞

を遠心分離し、回収した。回収した細胞は 4 mL の非変性バッファーA (40 mM Tris-HCl [pH 7.6]、300 mM NaCl、10 mM イミダゾール、1 mM DTT)に懸濁し、超音波破碎した。破碎された大腸菌を、12,000 g、4°C、10 分の条件で遠心分離した。4 mL の上清を 6 mL の非変性バッファーA と混合し、Millex-GV 0.45- μ m フィルター (Millipore)を用いて濾過した。濾過させた溶液を、500 μ L のベッドボリュームの Talon metal affinity resin (Clontech)と混合し、4°C、1 時間の条件で再度混合した。混合後、30 mL の非変性バッファーA を用いてタンパク質の結合した resin を洗浄した。結合したタンパク質を、非変性溶出バッファー(300mM イミダゾール、その他の組成は非変性バッファーA と同じ)を用いて溶出させた。

2.2.4. 精製されたタンパク質変異体の蛍光強度測定

蛍光強度測定のために、測定対象となる精製タンパク質 15.0 μ g を 500 μ L の PBS に希釈し、FP-6500 spectrofluorometer (日本分光)を用いて室温で測定した。トリプトファン(Trp)の蛍光スペクトルを FP-6500 spectrofluorometer (日本分光)を用いて室温で測定した。励起波長は 488 nm に、蛍光波長の検出範囲は 500 nm から 730 nm に設定した。励起、蛍光ともにバンド幅は 1 nm に設定した。

2.3. 結果

2.3.1. トリプトファンを置換したことによる蛍光強度の低下と、置換部位周辺への Saturation mutagenesis

本章では、緑色蛍光タンパク質(GFP)の蛍光強度と成熟速度が向上した変異体である GFPS1(35)に対して、唯一含まれるトリプトファン(図 2.1.)をフェニルアラニンに変異させた S1-W57F を用い、変異と選択による活性向上実験の起点となる初期変異体とした。Enhanced GFP (EGFP)は、Trp⁵⁷を他の 19 種類のアミノ酸に置換しても(36)、トリプトファンに近い構造の非天然アミノ酸に置換しても(37)、蛍光強度を失うことが報告されている。私は、GFPS1 においても、Trp⁵⁷を構造の似たフェニルアラニンやチロシンに置換すると蛍光強度が失われることを確認した(図 2.2.)。次に私は、S1-W57F の Phe⁵⁷に隣接するアミノ酸部位それぞれに対して saturation mutagenesis を行い、W57F 変異によって引き起こされた構造の乱れを修正しようとした。したがって、以下のいずれかのアミノ酸、Val²⁹、Phe⁴⁶、Leu⁵³、Leu⁶⁰、Leu⁶⁴ (図 2.3.A-E)のコドンが NNK に置換された変異体を作成した。それぞれの変異体タンパク質を発現する 10⁵のコロニーからなる大腸菌ライブラリの蛍光強度を、私は 488 nm の励起波長で FACS によって測定した。その結果、Trp⁵⁷への変異を修正することを期待して導入された変異は、どれも S1-W57F の蛍光強度を回復させなかったことが分かった。さらに私は、やはり Phe⁵⁷に隣接する連続した 3 つのアミノ酸部位である Asp²¹⁶、His²¹⁷、Met²¹⁸(図 2.3.F)のコドンが NNK に置換された変異体を作成した。同様に、生じる変異体のパターン数に対して十分なサイズである 10⁵のコロニーからなる、変異体タンパク質を発現している大腸菌ライブラリの蛍光強度を測定した。しかしこの試みも、S1-W57F の蛍光強度を回

復させなかった。

2.3.2. 1 ラウンド目の Directed evolution から得られた活性変異体の解析

Phe⁵⁷に隣接するアミノ酸に対する saturation mutagenesis では S1-W57F の蛍光強度を向上する変異体を取得することができなかつたため、遺伝子全体へのランダム変異導入法である error-prone PCR (epPCR) を S1-W57F 遺伝子に対して用いて Directed evolution (図 2.4.) を行い、活性向上変異体を得た。S1-W57F の epPCR 産物は、Placq-GFP(38) にクローニングされた。epPCR 産物をコードしたプラスミドを持つ 10⁶ のライブラリサイズからなる大腸菌群から、S1-W57F をコードしたプラスミドを持つ大腸菌よりも強い蛍光強度を示す大腸菌を選択した。私は選択された 10 個のコロニーからプラスミドを抽出し、遺伝子配列を解析した(図 2.5 および表 2.2)。結果、解析した配列の中にはアミノ酸置換を伴う変異が 13 個、S1-W57F の遺伝子全体に散らばって存在することが分かった。いくつかのクローンはまったく同じアミノ酸配列を持っていたため、6 つのクローンを独立の変異体として扱った。6 つの変異体の配列は、57 番目のアミノ酸も含め、トリプトファンを含んでいなかった。さらに、欠失、挿入、そして配列の途中のストップコドンの出現も観察されなかった。最も多く観察された変異は S205T であり、この変異箇所はクロモフォアに隣接し、Phe⁵⁷ には隣接しない位置にある。S205T はフォールディングを促進する変異として報告されている(39)。6 つの変異体をそれぞれ発現する大腸菌の蛍光強度を測定したところ、S205T を配列に含む変異体は、残りの変異体よりも強い蛍光強度を示すことが分かった(表 2.2.)。

2.3.3. 2 ラウンド目の Directed evolution から得られた活性変異体の解析

1 ラウンド目の活性向上変異体である W57F-S205T を R1-1 と名付け、これを鋳型とした 2 ラウンド目の Directed evolution を行ったところ、さらなる活性向上変異体を得ることができた。私は 1 ラウンド目と同様の変異導入と選択を R1-1 に対して行い、4 つの独立した変異体を得た(図 2.6.)。R2-1、R2-2、R2-3 は、1 アミノ酸変異を、R2-4 は 2 アミノ酸変異を持っていた(表 2.3.)。R2-3 に含まれる N105I は、Superfolder 変異体(40)における N105T、superfast 変異体(41)における N105Y と同じ残基であった。得られた変異群のうち、R2-4 に含まれる A206V は R1-1 に含まれる S205T と隣接していたものの、その他の変異はクロモフォアに隣接することもなく、また W57F や S205T とともに隣接していなかった。4 つの変異体はすべて、R1-1 よりも高い蛍光強度を示した(図 2.6.)。

2.3.4. Directed evolution によって得られた変異が与える影響の解析

変異体の蛍光強度を *in vitro* で調べるために、先行研究で報告されている GFP 変異体(EGFP, GFPS1)、初期変異体(S1-W57F)、選択された変異体(R1-1, R2-1, R2-3, R2-4)に His タグを導入し、精製・定量された変異体の蛍光強度を測定した。488 nm の励起波長に対して、本章の変異導入と選択で得られたすべての変異体は EGFP や GFPS1 と同様の蛍光スペクトルを示

した(図 2.7.)。得られた変異体の蛍光強度の強弱の順列は、大腸菌内で発現させて測定した時と、精製して測定した時とで変わらなかった。トリプトファンを含まない変異体間で蛍光強度を比較すると、S1-W57F に対して R1-1 は高く、R1-1 に対して R2-4 は 5.0 倍強い蛍光強度を持っていた。選択された最も強い蛍光強度を示す R2-4 は、20 種類のアミノ酸からなる変異体を比較すると、EGFP と比較して 58%、GFPS1 と比較して 37%の蛍光強度を持っていた。

トリプトファンを含まない GFP 変異体の活性向上に有効だった変異の効果をより深く評価するため、N105I および S205T の単独または組み合わせの変異を、トリプトファンを含まない変異体である S1-W57F および 20 種類のアミノ酸セットからなる GFPS1 に導入し、精製・定量された変異体の蛍光強度を測定した。期待したとおり、N105I と S205T はいずれも単独で S1-W57F の蛍光強度を向上させた(図 2.8.A)。さらに、N105I と S205T が同時に S1-W57F に導入されると、2 つの変異はほぼ加算的に S1-W57F の蛍光強度を向上させることが分かった。しかし、2 つの変異を単独で、あるいは同時に GFPS1 に導入しても、GFPS1 の蛍光強度に対する活性向上効果は明確ではなかった。

2.4. 結果のまとめと考察

本章では、普遍遺伝暗号表に最後に加わったとされるトリプトファンを含まない、19 種類のアミノ酸からなるタンパク質における有用な変異の蓄積に関する追加の実験事例を提供した。まず、GFPS1 における W57F 変異は蛍光活性を失わせた(図 2.2)。W57F 変異のあと、ランダム変異導入と変異による 2 ラウンドの Directed evolution を用い、私は S1-W57F、つまりトリプトファンを含まない 19 種類のアミノ酸からなる初期変異体、は進化し活性向上できるということを示した。Aequorea Victoria 由来の GFP のタンパク質配列に唯一含まれるトリプトファンである Trp⁵⁷ は、緑色蛍光の元となるクロモフォア形成に関わっている。クロモフォア形成の過程は、GFP の 65 番目から 67 番目のアミノ酸残基の自己環化反応を進行させるための疎水的環境を必要とする。GFP のタンパク質配列において、Trp⁵⁷ はプロリンに富む PVPWP モチーフの一部であり、酸素やその他の拡散小分子の衝突による消光からクロモフォアを守るために必要とされる疎水的環境を維持することに貢献している(36)。GFPS1 を作成するために導入された変異(35)が疎水的環境を大きく変えていないと仮定すると、GFPS1 の W57F 変異による蛍光活性の低下は、疎水的環境が乱されたことに起因しているかもしれない。野生型 GFP に W57F 変異が導入されたときに観察されたクロモフォア形成の減少(42)もまた、同様の理由によるものかもしれない。

さらに、トリプトファンを含まない 19 種類のアミノ酸からなるタンパク質における有用な変異の加算性を、遺伝子全体へのランダムな変異導入を用いて私は観察した。N105I、S205T のどちらの単独の変異も、S1-W57F の蛍光活性を向上させた(図 2.8.)。S1-W57F に

N105I および S205T の二重変異を導入すると、蛍光活性はほぼ加算的に向上した。過去の報告では、S205T は GFP のフォールディングを促進する変異として記述されている(Cubitt1999 を引用する)。N105I は、Superfolder 変異体(40)および Superfast 変異体(41)において、それぞれリフォールディングの際に変性剤への耐性を高める変異である N105T および N105Y と同じ位置の変異である。したがって、S205T、そしておそらく N105I も、W57F 変異によって損なわれた S1-W57F のフォールディングを改善することに貢献したのかもしれない。2つの個々の有用な変異は常に加算的に働くとは限らない(43)が、S1-W57F が2つの個々の有用な変異を得て、R2-3 へと進化することを観察することができた。複数の有用な変異が単純な加算的な活性向上効果をもたらすことは、Directed evolution を特徴づける現象の1つである(44)。19種類のアミノ酸からなるタンパク質の Directed evolution における有用な変異の加算性は、タンパク質のファミリーのアミノ酸配列を比較して有用な置換箇所を推定した研究において示唆されていた(34)。さらに、私の研究によって実際に、19種類のアミノ酸からなるタンパク質の遺伝子全体へのランダム変異を伴う Directed evolution による、有用な変異の加算性が示された。

本研究で作成されたトリプトファンを含まない19種類のアミノ酸からなる GFP 変異体は、トリプトファンを任意の場所に導入できる足場として利用できる。トリプトファンを含まない GFP 変異体に対して、変異導入を用いてトリプトファンを新たに導入することは、GFP のフォールディングに新たな知見を提供するだろう。トリプトファンは、特徴的な蛍光特性によって、リフォールディングの過程を追跡するためのレポーターとして利用されてきた。タンパク質が変性すると、トリプトファンがタンパク質内部から溶媒に露出することでトリプトファンの蛍光ピークが長波長側にシフトする。これまで、GFP 変異体のリフォールディング過程は Trp⁵⁷ やクロモフォアの蛍光を観察することで調べられてきた(45)。トリプトファンを含まない変異体を利用して様々な場所に1つのトリプトファンを持つ様々な一置換体を利用したリフォールディング研究は、すでに他のタンパク質を用いた研究で示されている(46)ように、タンパク質のフォールディングに関する有用な知見を提供するかもしれない。トリプトファンを含まない19種類のアミノ酸からなる GFP 変異体へのトリプトファンの導入は、タンパク質のフォールディングの理解をさらに深める助けになるだろう。19種類のアミノ酸からなる GFP 変異体の別の応用例として、システインを含まない19種類のアミノ酸からなる GFP 変異体も報告されている。システインを含まない19種類のアミノ酸からなる GFP 変異体は2つの変異、site-directed mutagenesis によって導入された C48S と saturation mutagenesis によって導入・選択された C70M を組み合わせて作成された(32)。この変異体は酸性溶液中での GFP の拡散性を改善し、タンパク質の分泌経路を解析する際に GFP をタグとして利用する用途を広げた。

タンパク質の活性を向上させる変異は、高活性の変異体上よりも低活性の変異体上にお

いて検出されやすいという報告(47,48)は、トリプトファンを含まない 19 アミノ酸セットからなる変異体においても成立した(図 2.8.B)。私は低活性の S1-W57F を始めに作成し、S1-W57F をもとにランダム変異を用いて作成されたライブラリから、活性型の変異体を選出した。実際に、R2-3に含まれていた変異である N105I と S205T は、それぞれ単独で S1-W57F に導入した場合にも活性向上効果を示した(図 2.8.A)。対照的に、これらの変異は、GFPS1 において明確な活性向上を示さなかった。異なる遺伝的バックグラウンドによって変異の効果が異なる現象は epistasis(49,50)として知られており、低活性の変異体を変異の対象として有効な変異を探索する「suppressor mutation method」(47)の前提となっている。

2.5. 図表

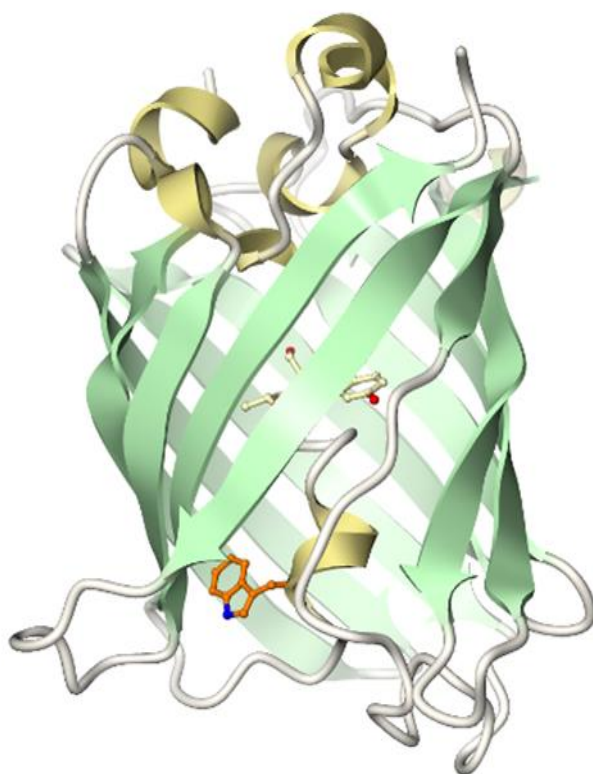


図 2.1. EGFP の立体構造 (PDB entry 2y0g)

トリプトファンを橙色の、クロモフォアを黄色の Ball & Stick モデルで示した。描画はプログラム CueMol 2 (<http://www.cuemol.org>)を用いた。

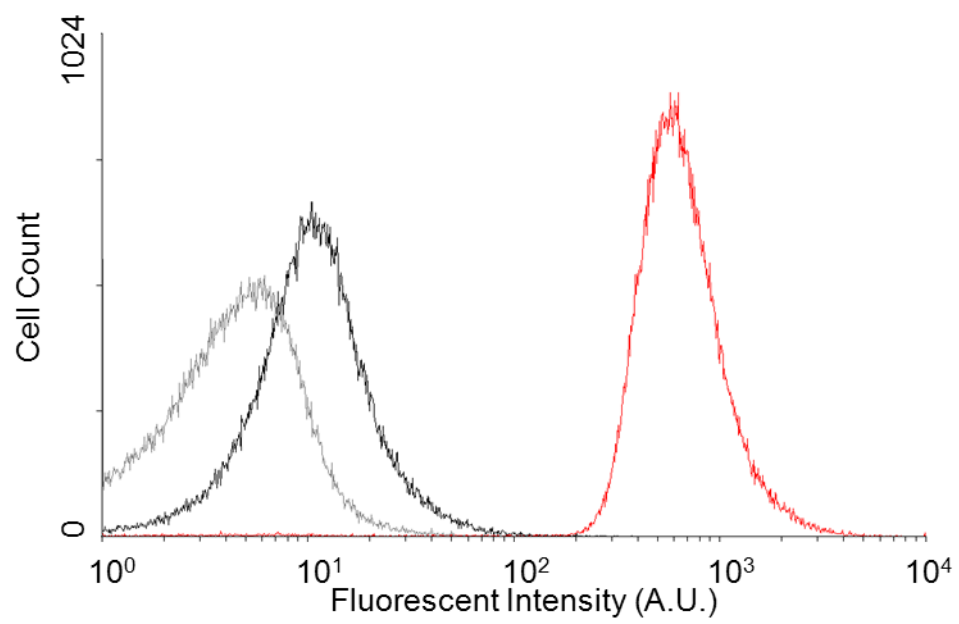
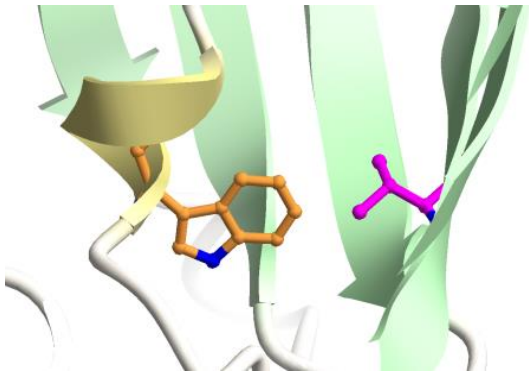
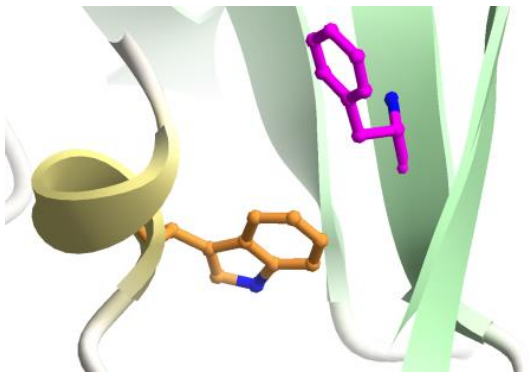


図 2.2. S1-W57Y、S1-W57F、GFPS1 の蛍光強度を表したヒストグラム
S1-W57Y を灰色、S1-W57F を黒、GFPS1 を赤色で示した。

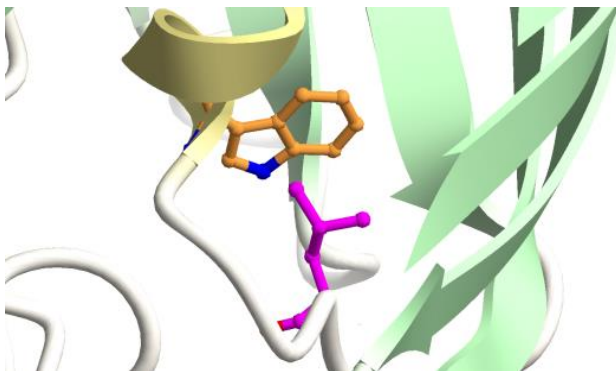
A. Val²⁹



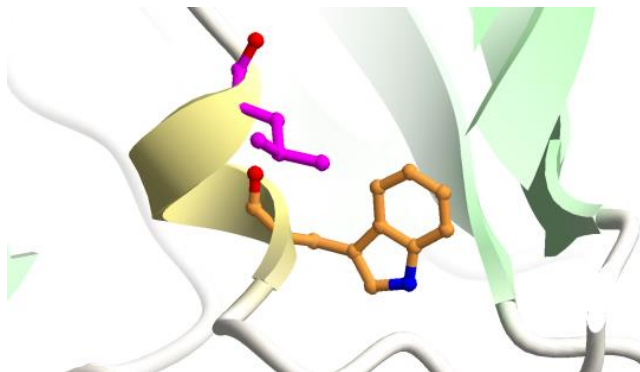
B. Phe⁴⁶



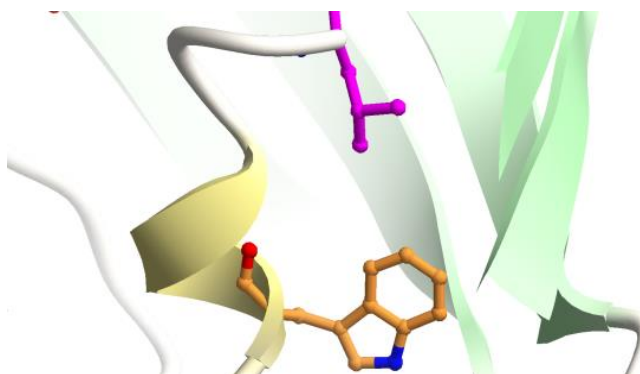
C. Leu⁵³



D. Leu⁶⁰



E. Leu⁶⁴



F. Asp²¹⁶-His²¹⁷-Met²¹⁸

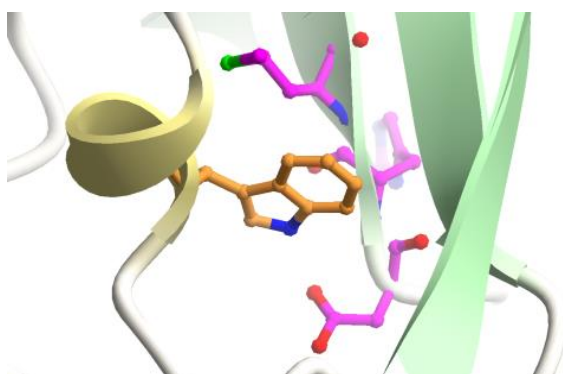


図 2.3. Saturation mutagenesis を行った部位の構造

Saturation mutagenesis を行った以下の部位の構造を、EGFP(PDB entry 2y0g)の構造上にピンク色で示した: A. Val²⁹、B. Phe⁴⁶、C. Leu⁵³、D. Leu⁶⁰、E. Leu⁶⁴、F. Asp²¹⁶-His²¹⁷-Met²¹⁸。Trp⁵⁷ は橙色で示した。

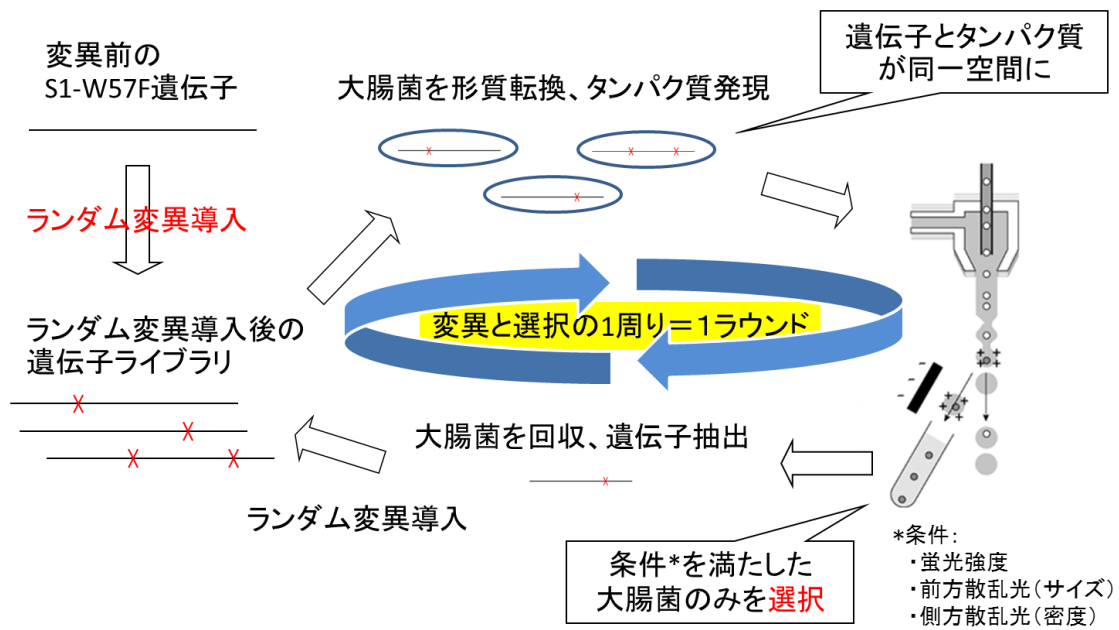


図 2.4. セルソーターを用いた Directed evolution

GFPS1.....	1:MKSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYGLTLKFICTTGKLPVFPPT	60
GFPS1-W57F.....	1:MKSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYGLTLKFICTTGKLPVFPPT	60
R1-1.....	1:MKSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYGLTLKFICTTGKLPVFPPT	60
R1-2.....	1:MKSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYGLTLKFICTTGKLPVFPPT	60
R1-3.....	1:MKSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYGLTLKFICTTGKLPVFPPT	60
R1-4.....	1:MKSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYGLTLKFICTTGKLPVFPPT	60
R1-5.....	1:MKSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYGLTLKFICTTGKLPVFPPT	60
R1-6.....	1:MKSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYGLTLKFICTTGKLPVFPPT	60
EGFP.....	1:MKSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYGLTLKFICTTGKLPVFPPT	59
avGFP.....	1:MKSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYGLTLKFICTTGKLPVFPPT	59
GFPS1.....	61:LVTTLTYGVQCFSTRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEGDTL	120
GFPS1-W57F.....	61:LVTTLTYGVQCFSTRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEGDTL	120
R1-1.....	61:LVTTLTYGVQCFSTRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEGDTL	120
R1-2.....	61:LVTTLTYGVQCFSTRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEGDTL	120
R1-3.....	61:LVTTLTYGVQCFSTRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEGDTL	120
R1-4.....	61:LVTTLTYGVQCFSTRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEGDTL	120
R1-5.....	61:LVTTLTYGVQCFSTRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEGDTL	120
R1-6.....	61:LVTTLTYGVQCFSTRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEGDTL	120
EGFP.....	60:LVTTLTYGVQCFSTRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEGDTL	119
avGFP.....	60:LVTTLTYGVQCFSTRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEGDTL	119
GFPS1.....	121:VNRIELKGIIDFKEDGNILGHKLEYNYNHSHVYITADKQKNGIKANFKIRHNIEDGSVQLA	180
GFPS1-W57F.....	121:VNRIELKGIIDFKEDGNILGHKLEYNYNHSHVYITADKQKNGIKANFKIRHNIEDGSVQLA	180
R1-1.....	121:VNRIELKGIIDFKEDGNILGHKLEYNYNHSHVYITADKQKNGIKANFKIRHNIEDGSVQLA	180
R1-2.....	121:VNRIELKGIIDFKEDGNILGHKLEYNYNHSHVYITADKQKNGIKANFKIRHNIEDGSVQLA	180
R1-3.....	121:VNRIELKGIIDFKEDGNILGHKLEYNYNHSHVYITADKQKNGIKANFKIRHNIEDGSVQLA	180
R1-4.....	121:VNRIELKGIIDFKEDGNILGHKLEYNYNHSHVYITADKQKNGIKANFKIRHNIEDGSVQLA	180
R1-5.....	121:VNRIELKGIIDFKEDGNILGHKLEYNYNHSHVYITADKQKNGIKANFKIRHNIEDGSVQLA	180
R1-6.....	121:VNRIELKGIIDFKEDGNILGHKLEYNYNHSHVYITADKQKNGIKANFKIRHNIEDGSVQLA	180
EGFP.....	120:VNRIELKGIIDFKEDGNILGHKLEYNYNHSHVYITADKQKNGIKANFKIRHNIEDGSVQLA	179
avGFP.....	120:VNRIELKGIIDFKEDGNILGHKLEYNYNHSHVYITADKQKNGIKANFKIRHNIEDGSVQLA	179
GFPS1.....	181:DHYQNTPIGDGPVLLPDNHCLSTQSALS KDPNEKRDMVLEFVTAAGITHGMDELYK	239
GFPS1-W57F.....	181:DHYQNTPIGDGPVLLPDNHCLSTQSALS KDPNEKRDMVLEFVTAAGITHGMDELYK	239
R1-1.....	181:DHYQNTPIGDGPVLLPDNHCLSTQSALS KDPNEKRDMVLEFVTAAGITHGMDELYK	239
R1-2.....	181:DHYQNTPIGDGPVLLPDNHCLSTQSALS KDPNEKRDMVLEFVTAAGITHGMDELYK	239
R1-3.....	181:DHYQNTPIGDGPVLLPDNHCLSTQSALS KDPNEKRDMVLEFVTAAGITHGMDELYK	239
R1-4.....	181:DHYQNTPIGDGPVLLPDNHCLSTQSALS KDPNEKRDMVLEFVTAAGITHGMDELYK	239
R1-5.....	181:DHYQNTPIGDGPVLLPDNHCLSTQSALS KDPNEKRDMVLEFVTAAGITHGMDELYK	239
R1-6.....	181:DHYQNTPIGDGPVLLPDNHCLSTQSALS KDPNEKRDMVLEFVTAAGITHGMDELYK	239
EGFP.....	180:DHYQNTPIGDGPVLLPDNHCLSTQSALS KDPNEKRDMVLEFVTAAGITHGMDELYK	238
avGFP.....	180:DHYQNTPIGDGPVLLPDNHCLSTQSALS KDPNEKRDMVLEFVTAAGITHGMDELYK	238

図 2.5. 第一世代活性変異体のアミノ酸配列

第一世代の活性変異体のアミノ酸配列を、GFPS1、EGFP、オワンクラゲ由来の野生型 GFP 配列である avGFP の配列とともに示した。GFPS1 を基準に、アミノ酸の変化がある箇所を黒で表示した。 α ヘリックスおよび β シートを形成する箇所に下線を引いた。ギャップはハイフンで表示した。

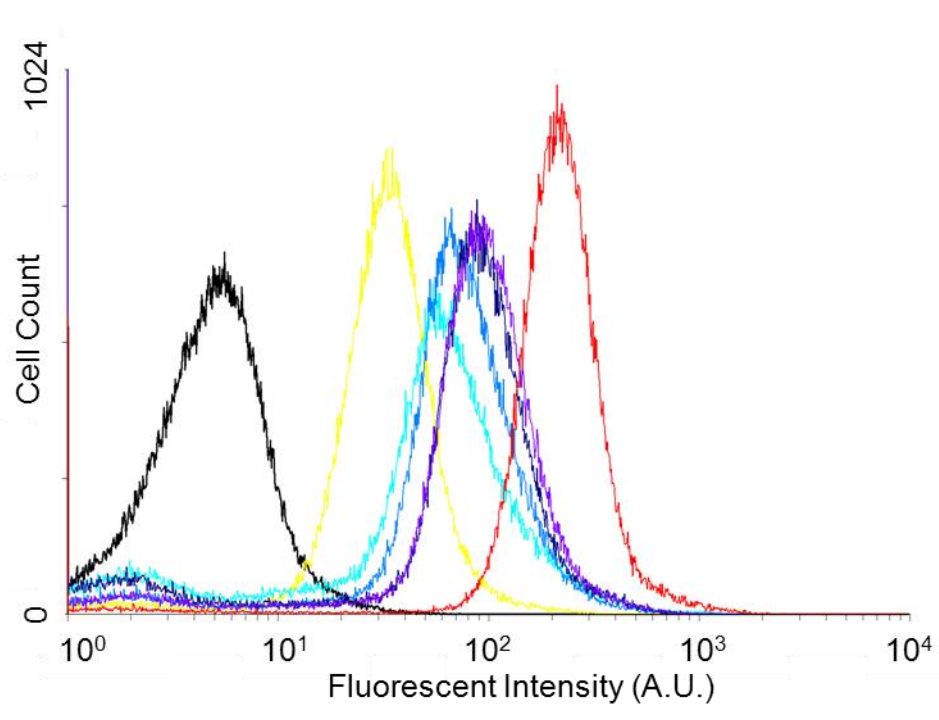


図 2.6. 第二世代活性変異体のヒストグラム
黒 : S1-W57F、黄 : R1-1、水色 : R2-1、青 : R2-2、紺 : R2-3、紫 : R2-4、赤 : GFPS1

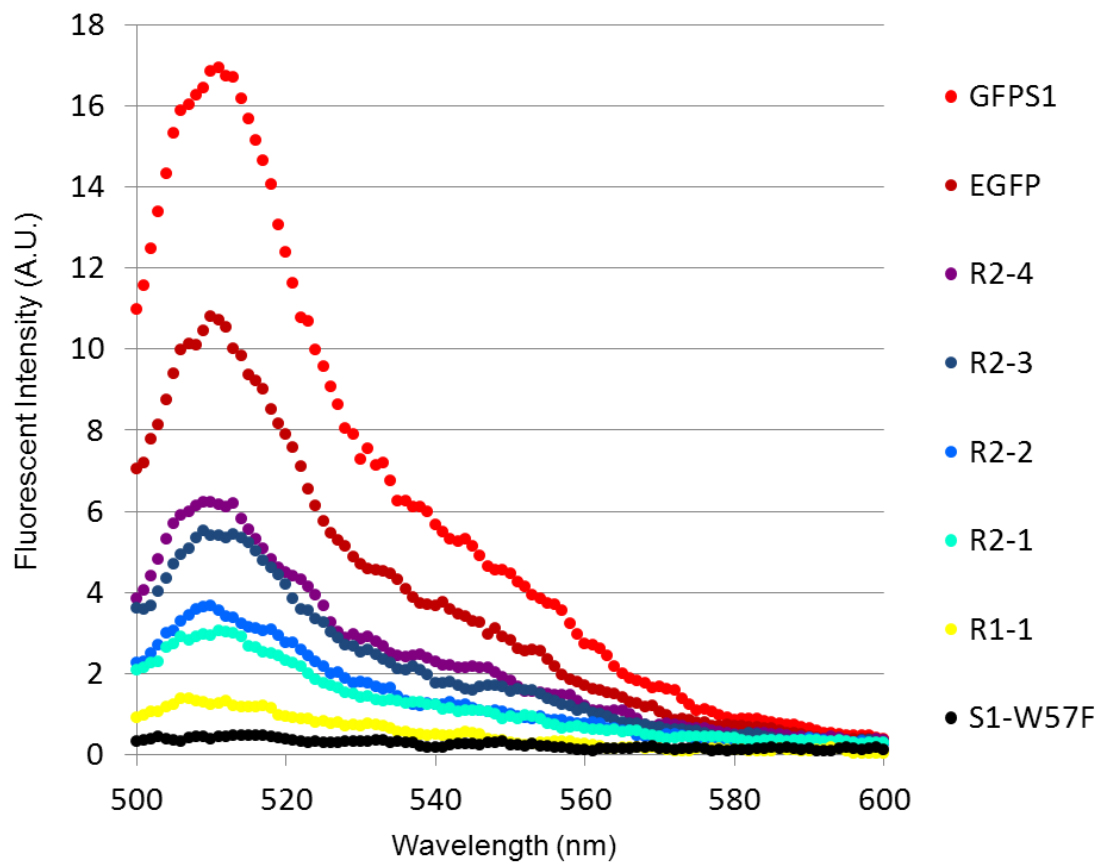
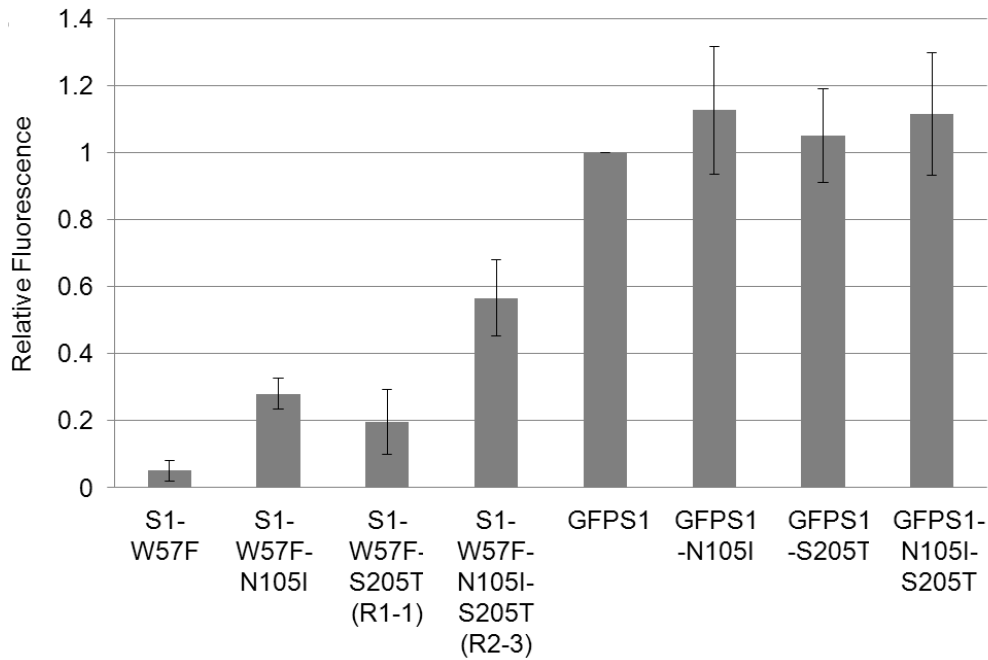


図 2.7. 蛍光スペクトル

A



B.

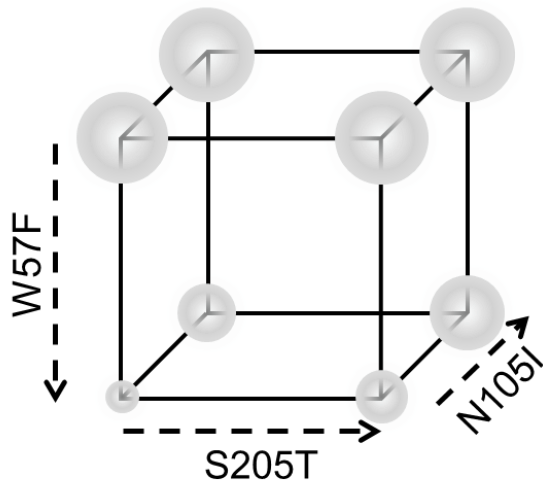


図 2.8. Trp を含まない 19 アミノ酸からなる変異体の活性を向上させた変異の影響の評価

A. GFP 変異体の相対的な蛍光強度。R2-3 に現れた N105I と S205T を単独、あるいは組み合わせて S1-W57F あるいは GFPS1 に導入した。それぞれの変異体を精製し、488 nm の波長で励起し、510 nm の蛍光を検出した。タンパク質濃度は、PBS 中で 7.5 $\mu\text{g/ml}$ に合わせた。測定結果は GFPS1 の蛍光強度を 1 として規格化し、3 回の測定の平均値 \pm 標準偏差で示した。B. 有用な変異は、弱い活性を持った変異体上で検出されやすいことを示した模式図。立方体の頂点は、GFP 変異体の遺伝子型を表している。球の大きさは、蛍光強度の強さを表している。底の面に 19 アミノ酸セットからなる GFP 変異体が、頂上の面に 20 アミノ酸セットからなる GFP 変異体が現れるよう配置した。

表 2.1. プライマーリスト

Primer name	Primer sequence (5' – 3') *
Placq- <i>Bgl</i> _II-RBS-F	GAAGATCTGAAATAATTTTGTTTAACTTTAAGAAGGAG
Placq- <i>Spe</i> _I-R	CCACTAGTTCATTATTTGTAGAGCTCATCCATG
W57F-F	GGAAACTACCTGTTCCATTCCCAACACTTG
W57Y-F	GGAAACTACCTGTTCCATACCCAACACTTG
W57FY-R	TGGAACAGGTAGTTTTCCAGTAGTGC
F57W-F	GAAACTACCTGTTCCATGGCCAACACTTG
F57W-R	ATGGAACAGGTAGTTTTCCAGTAGTGC
N105I-F	TTCAAAGATGACGGGATCTACAAGACGCG
N105I-R	CCCGTCATCTTTGAAAGATATAGTGCGTTC
S205T_F	CTGTTCGACACAAACTGCCCTTTC
S205T_R	ATGGAACAGGTAGTTTTCCAGTAGTGC
V29NNK-F	GGGCACAAATTTTCTNNKAGTGGAGAGG
V29NNK-R	AGAAAATTTGTGCCATTAACATCACC
F46NNK-F	GAAACTTACCCTTAAANNKATTTGCACTAC
F46NNK-R	TTTAAGGGTAAGTTTTCCGTATGTTGC
L53NNK-F	TGCACTACTGGAAAANNKCTGTTCCATG
L53NNK-R	TTTTCCAGTAGTGCAAATAAATTTAAGGG
L60NNK-F	GTTCCATGGCCAACANNKGTCACTACTTTGAC
L60NNK-R	TGTTGGCCATGGAACAGGTAGTTTTCC
L64NNK-F	CAACTTGTCACTACTNNACTTATGGTGTTC
L64NNK-R	AGTAGTGACAAGTGTGGGAATGGAAC
D216H217M218NNK-F	CCCAACGAAAAGCGTNNKNNKNNKGTCTTCTTGAGTTTG
D216H217M218NNK-R	ACGCTTTTCGTTGGGATCTTTCGAAAGG
pGFP- <i>Cla</i> _I-His-F	GAGCTCTACAAATATCGATATGAATTCCAACCTG
pGFP- <i>Cla</i> _I-His-R	TATTTGTAGAGCTCATCCATGCCATG
pGFPS1-F	GTGATGGATATCTGCAGAATTTCGGCTGCTC
pGFP-His <i>Spe</i> _I-R	GGACTAGTTAGTGGTGGTGGTGGTGGTGG
GFP-cloning- <i>Cla</i> _I-R	CCATCGATATTTGTAGAGCTCATCCATGCCATG

*制限酵素サイトはイタリックで表示した。

表 2.2. 第一世代活性変異体の変異箇所、蛍光強度、得られたクローンの数

Mutant	Amino acid mutations from S1-W57F	Fluorescence Intensity*	Number of clones
S1-W57F	-	20	-
R1-1	S205T	98**	4
R1-2	K1I, S205T	100	1
R1-3	K166E, S205T	96	1
R1-4	Y145F, E235K	54	1
R1-5	Q184L, L236P	52	1
R1-6	K1R, Y145H, C200Y	89***	2

*ヒストグラムデータの中央値を蛍光強度の代表値として表記した

**4 クローンの蛍光強度の平均値

***2 クローンの蛍光強度の平均値

表 2.3. 第二世代活性変異体の変異箇所、蛍光強度、得られたクローンの数

Mutant	Amino acid mutations from R1-1	Fluorescence Intensity ^{*,**,*}	Number of clones
S1-W57F	-	4.8	-
R1-1	-	33	-
R2-1	G232D	56	1
R2-2	E172V	71	1
R2-3	N105I	84	1

*ヒストグラムデータの中央値を蛍光強度の代表値として表記した

**表 2.2. と表 2.3. とでは感度に関わる電圧が異なるため、値が異なることがある。

第三章 トリプトファンを含まない 19 種類のアミノ酸セットからなる遺伝暗号表の構築

3.1. 導入

3.1.1. 天然および実験室内における拡張遺伝暗号表に関する先行研究

遺伝暗号表に含まれるアミノ酸の数の増加は、新たな aaRS と tRNA の導入によって実現されることが、天然で発見された例および実験室内で構築された例で知られている。新たな aaRS と tRNA の導入には、既存の「アミノ酸-aaRS-tRNA-コドン」の経路と交叉しないことが求められる。もし、新たに導入した aaRS や tRNA が既存の経路と交叉すれば、既存のコドンへの非天然アミノ酸の導入、あるいは非天然アミノ酸が導入されるはずのコドンへの既存のアミノ酸の導入が起きてしまい、いずれの場合においても意図したアミノ酸配列が得られなくなってしまう。天然で発見されたピロリジンにおいては、ピロリジル tRNA 合成酵素(PyIRS)によって行われる直接的なアミノアシル化によってピロリジンがピロリジル tRNA(tRNA^{PyI})に結合され、UAG コドンに対応するアミノ酸として交叉反応なくタンパク質に導入される(51)。実験室内で非天然アミノ酸が導入された例においては、*Methanococcus jannaschii* 由来のチロシル tRNA 合成酵素(TyrRS)および tRNA^{Tyr} をそれぞれ改変し、大腸菌内で用いることで、O-methyl-L-tyrosine を UAG コドンに対応するアミノ酸として交叉反応なくタンパク質に導入することに Wang らが成功した(18)。

ここで遺伝暗号表の拡張と単純化に求められる要素を比較すると、拡張・単純化の両者ともにアミノ酸とコドンの新たな対応付けが必要であるという理由から、新規の tRNA 変異体が必要である、という共通点が見られる。いっぽうで、遺伝暗号表の拡張に必要であった、非天然アミノ酸を認識し新規の tRNA 変異体に結合させる aaRS 変異体を作成する、という要求は、遺伝暗号表の単純化には必ずしも必要ではない。代わりに、既存の aaRS の活性を抑制する、ということが遺伝暗号表の単純化には求められる(図 3.1.)。

3.2. 材料と方法

3.2.1. DNA コンストラクションおよび *in vitro* 転写

マルトース結合タンパク質(MBP)、LexA、およびクロラムフェニコールアセチルトランスフェラーゼ(CAT) の遺伝子を、pK7 プラスミド(52)に導入した。GFP の遺伝子を、pGFP プラスミド(35)に導入した。tRNA 変異体をコードする遺伝子を、pUC119 プラスミド(Takara)に導入した。tRNA 変異体を、T7RNA ポリメラーゼを用いた転写反応によって作成した。反応溶液組成や反応条件は過去の報告(19)に従った。基質となる LexA 変異体の遺伝子(53)、すなわち 99 番目のコドンが UAA ストップコドンに置換された LexA L89P-Q92W-Y98K は、pET26b プラスミド(Novagen)に導入した。変異は Site-directed mutagenesis を用いて MBP、LexA、CAT、GFP、および tRNA の DNA コンストラクトに導入した。変異を DNA シークエンシングによって確認した。MBP-W_all_A を化学的に合成し(GeneArt)、pK7 プラスミドに導入した。N 末端のヒスチジンタグを、MBP、LexA、および CAT の遺伝子に導入した。

C末端のヒスチジンタグを、GFPの遺伝子に導入した。PreScission Protease (GE Healthcare)の認識サイトを、LexAのDNAコンストラクトと、結晶化に用いられたGFPのDNAコンストラクトのそれぞれに、ヒスチジンタグと対象タンパク質の配列の間に導入された。

3.2.2. 無細胞翻訳系

本研究では、大腸菌 S30 無細胞翻訳系(以下、単に S30 と表記)を用いた。S30 の組成は、特定の 1 種類のアミノ酸を除外している点、tRNA 変異体を添加している点、および 5.0 μ M の aaSA(aminoacyl adenylate analogs, Integrated DNA Technologies)を加えている点以外は、Kigawa らの先行研究(54) に従った。S30 は大腸菌 BL21(DE3)株から作成した。20 μ L の反応容量に対してはバッチモードを、60 μ L から 3,000 μ L の反応容量に対しては透析モードを用いた。反応時間については、バッチモードに対しては 1 時間、透析モードには 8 時間を採用した。セリンを UGU/UGC コドンに再割り当てする tRNA^{Ser} 変異体を用いた翻訳には、CysRS のシステイン導入に対する阻害効果を持続させるため、最終濃度 5 μ M 相当の Cys-SA を、内液および外液に 1 時間ごとに添加した。

3.2.3. タンパク質精製

GFP、MBP、および CAT タンパク質の精製のため、Talon metal affinity resin (Clontech)を、販売元の推奨条件をもとに、一部改変を加えて使用した。

非変性条件での GFP の精製のため、反応終了後の S30 を 12,000 g、4 $^{\circ}$ C、10 分間の条件で遠心分離した。3 mL の上清を 7 mL の非変性バッファーA(40 mM Tris-HCl [pH 7.6]、300 mM NaCl、10 mM イミダゾール、1 mM DTT)と混合し、Millex-GV 0.45 μ m フィルター (Millipore) を用いて濾過した。濾過させた溶液を Talon metal affinity resin と混合し、4 $^{\circ}$ C、1 時間の条件で再度混合した。混合後、非変性バッファーA を用いてタンパク質の結合した resin を洗浄した。洗浄後、タンパク質を非変性溶出バッファー(300mM イミダゾール、その他の組成は非変性バッファーA と同じ)を用いて溶出させた。

変性条件での MBP および CAT の精製のため、1mL の反応終了後の S30 を 9 mL の変性バッファーA(8 M 尿素、50 mM リン酸バッファー[pH 7.8]、300 mM の NaCl、10 mM のイミダゾール)と混合し、37 $^{\circ}$ C、1 時間の条件で振とう混合した。混合後の溶液を Talon metal affinity resin と混合し、室温、1 時間の条件で再度混合した。再混合後、変性バッファーB (40 mM イミダゾール、その他の組成は変性バッファーA と同じ) を用いてタンパク質の結合した resin を洗浄した。洗浄後、タンパク質を変性溶出バッファー(150mM イミダゾール、その他の組成は変性バッファーA と同じ) を用いて溶出させた。

3.2.4. 放射標識されたタンパク質の検出

S30 を用いた MBP、GFP、および CAT の翻訳を、3.2.2.項で記述した組成に[¹⁴C]ロイシンを加え、20 μ L 容量のバッチモードで 37 $^{\circ}$ C、1 時間の条件で行った。反応を終えた未精製産物を、12%ビストリスゲルと 2-(N-morpholino) ethanesulfonic acid (MES)泳動バッファー(50

mM MES、50 mM Tris-base、3.47 mM SDS、1.0 mM EDTA [pH 7.3])を用いた電気泳動によって解析した。タンパク質の放射能を検出するために、イメージアナライザーFLA-5000(富士フィルム)およびイメージングプレート BAS-IP MS 2040(富士フィルム)を用いた。

3.2.5. アミノ酸組成分析

S30 を用いた MBP および CAT の翻訳を、37°C、8 時間の条件で、中容量の透析モード(内液 1 mL/外液 10 mL) (55)によって行った。N 末端にヒスチジンタグを持つ合成産物を変性条件下で精製し、その後アセトン沈殿を行った。

トリプトファンを検出するため、精製産物を 0.2%の 3-(2-aminoethyl) indole を含む 4 M のメタンスルホン酸(56)中で、110°C、20 時間の条件で加水分解した。加水分解後、MSA を中和するために NaOH を添加した。3-(2-aminoethyl) indole が AQC-アミノ酸化合物のクロマトグラムを乱すため、ニンヒドリン法を用いて加水分解産物を反応させ、反応産物を High Speed Amino Acid Analyzer L-8900 (日立ハイテクノロジー)を用いて検出した。なお、MSA を用いた加水分解によるクロマトグラムは規格化されていない。

アラニン、およびその他のほとんどのアミノ酸の含有量を定量化するため、精製産物を SDS-PAGE によって分離し、PVDF 膜に電氣的に転写し、クマシーブリリアントブルーを用いて染色し、染色されたバンドを切り出した。気層 HCl による加水分解、aminoquinolyl-N-hydroxysuccinimidyl carbamate (AQC)による誘導体化、AQC と反応したアミノ酸のクロマトグラフィーによる検出は、1200 series SL HPLC system (アジレントテクノロジー)を用いたことを除いて、先行研究で報告されている方法(57)に従った。ノルバリンを、内部標準として添加した。サンプルごとに生成されるクロマトグラムは、HCl を用いた加水分解のフェニルアラニンのピークの頂上と底辺を基準に規格化した。

システインを検出するため、精製産物を 1%の DTT を用いて還元し、2.5%のヨード酢酸を用いてアルキル化した。還元、アルキル化の後、反応産物を SDS-PAGE を用いて分離し、PVDF 膜に電氣的に転写し、クマシーブリリアントブルーを用いて染色し、染色されたバンドを切り出した。切り出したバンドを 24、48、72 時間、加水分解した。AQC による誘導体化およびクロマトグラフィーによる検出は上記のとおりである。WT/Univ の 24 時間のデータを除き、得られたデータを用いて 0 時間のデータを線形外挿した。

表 3.1.に用いられている、タンパク質 1 分子あたりのアミノ酸 X の含有量 C_{X1} は、アミノ酸組成分析から得られたアミノ酸 X のピーク強度の実測値 M_x を用いて、以下の式で計算した。

$$C_{X1} = \frac{M_x}{\frac{1}{2} \left(\frac{M_F}{C_{FT}} + \frac{M_L}{C_{LT}} \right)}$$

ここで、 M_F および M_L はそれぞれフェニルアラニンとロイシンのピーク強度の実測値を表している。 C_{FT} および C_{LT} はそれぞれタンパク質 1 分子あたりのフェニルアラニンとロイシ

ンの理論的な含有量を表している。

表 3.3. に用いられているタンパク質 1 分子あたりに含まれるアミノ酸 X の含有量 C_{X3} は、アミノ酸組成分析から得られたアミノ酸 X のピーク強度の実測値 M_x を用いて、以下の式で計算した。

$$C_{X3} = \frac{M_x}{\frac{1}{3} \left(\frac{M_F}{C_{FT}} + \frac{M_L}{C_{LT}} + \frac{M_A}{C_{AT}} \right)}$$

ここで、 M_A はアラニンのピーク強度の実測値を表している。 C_{AT} はタンパク質 1 分子あたりのアラニンの理論的な含有量を表している。

規格化にフェニルアラニン、ロイシン、アラニンの値を用いたのは、これらのアミノ酸が酸による加水分解に対して比較的安定な値を出すことが知られていたためである。

3.2.6. トリプトファンの蛍光スペクトル測定

S30 を用いた MBP の翻訳を、37°C、8 時間の条件で、小容量の透析モード(内液 60 μ L / 外液 600 μ L) (58) によって行った。N 末端にヒスチジンタグを持つ合成産物を変性条件下で精製した。測定対象となる精製タンパク質 15.0 μ g を 500 μ L の変性バッファー A に希釈し、FP-6500 spectrofluorometer (日本分光) を用いて室温で測定した。励起波長は 280 nm に、蛍光波長の検出範囲は 290 nm から 500 nm に設定した。励起、蛍光ともにバンド幅は 1 nm に設定した。

3.2.7. 無細胞翻訳系中における GFP の蛍光強度測定

S30 を用いた GFP の翻訳を、37°C、1 時間の条件で、20 μ L 容量のバッチモードによって行った。未精製産物からの蛍光は、Mx3005P (Stratagene) を用いて測定した。励起波長に 515 nm、検出波長に 550 nm を設定した。MilliQ を含んだセルを測定し、それぞれのサンプルから差分を計算するためのバックグラウンドとした。ヒスチジンタグを抗原とするモノクローナル抗体(Novagen)を用いて、ウエスタンブロットを行った。

3.2.8. 結晶化、および X 線回折データセットの収集

S30 を用いた GFP の翻訳を、30°C、8 時間の条件で、大容量の透析モード(内液 3mL/外液 30mL)(59)によって行った。C 末端にヒスチジンタグを持つ合成産物を非変性条件下で精製した。C 末端のヒスチジンタグを PreScission Protease (GE Healthcare) を用いて切断した。プロテアーゼを、Glutathione Sepharose 4B resin (GE Healthcare) を用いて除いた。PreScission Protease による切断を免れた GFP、および切断後のポリヒスチジンを Talon metal affinity resin (Clontech) を用いて除いた。切断および残渣除去後の精製産物を、Resource Q ion exchange column (GE Healthcare) を用いたイオン交換クロマトグラフィーによって、さらに精製した。得られた精製産物を、20 mM Tris-HCl [pH 8.5]、1 mM DTT、および 50 mM NaCl の溶液中で

透析した。さらに、透析産物を Amicon Ultra-0.5 10,000 MWCO フィルター(MilliPore)を用いて限外濾過法により 15 mg/mL に濃縮した後、0.22 μm のフィルターを用いて濾過した。結晶化は、ハンギングドロップ蒸気拡散法により行った。R2-4_A110X(UGG)/Sim は、17% PEG 20,000、100 mM MES [pH 6.5]の溶液条件で、4°C、7日間で結晶が成長した。R2-4/Univ は、15% PEG 6,000、5% glycerol の溶液条件で、20°C、7日間での結晶が成長した。

得られた結晶を用いて、X線回折実験を行った。X線による結晶の損傷を最小限に抑えるため、回折像の測定は 100 K の窒素ガスのクライオストリーム中にて行った。結晶化に用いたりザーバー溶液と同じ組成の溶液を用いて結晶を含む溶液を徐々に平衡化した。得られた結晶を 1.2 倍濃度のりザーバー溶液中に移した。回折データセットは、大型放射光施設 SPring-8 (播磨)のビームライン BL41XUにて測定した。得られた回折データセットは、プログラム HKL2000(60)を用いて処理した。

3.2.9. 原子モデルの構築、および構造の精密化

結晶モデルを、GFP S65T の構造モデル(PDB code 1EMA)をサーチモデルとして、プログラム MOLREP(61)を用いて作成した。得られた初期の構造モデルを、電子密度に合うようプログラム COOT(62)を用いて修正した。最終的に、R2-4_A110X(UGG)/Sim および R2-4/Univ の原子モデルを、プログラム PHENIX (63)および REFMAC5 (64)を用いてそれぞれ 2.10 Å および 1.85 Å に精密化した。結晶のデータセットと精密化の統計データを表 3.2. に示した。分子モデルの描画には、CueMol 2 (<http://www.cuemol.org/>)を用いた。

3.2.10. LexA 切断アッセイ

LexA 切断アッセイを、過去の報告(53)に従って行った。

基質となる LexA 変異体を得るため、基質となる LexA 変異体の遺伝子を持つプラスミドを用いて大腸菌株 Rosetta2 (DE3) pLysS (Merck)を形質転換し、37°Cで培養した。590 nm の吸光度が 0.6 に達した時点で 0.5 mM の isopropyl- β -D-thiogalactopyranoside (IPTG)を添加してタンパク質発現を誘導した。誘導開始の 4 時間後、大腸菌を遠心分離を用いて回収した。回収した大腸菌を非変性バッファーB (40mM イミダゾール、その他の組成は非変性バッファーA と同じ)に懸濁し、超音波破碎した。破碎された大腸菌を、12,000 g、4°C、10 分の条件で遠心分離した。酵素として働く LexA タンパク質を得るため、S30 を用いた LexA の翻訳を、30°C、8 時間の条件で、中容量の透析モード(内液 1 mL/外液 10 mL) (55)によって行った。合成を終えた溶液を、12,000 g、4°C、10 分の条件で遠心分離した。N 末端にヒスチジンタグを持つ基質となる LexA 変異体、および酵素として働く LexA タンパク質の合成産物を、上記の GFP の精製と同様の条件で精製した。

LexA 切断アッセイを、20 mM CAPS-NaOH [pH 10.0]、200 mM NaCl、3 μM の酵素として働く LexA、および 30 μM の基質となる LexA を含む溶液組成にて、37°C36 時間の反応条件で行った。反応終了時、反応サンプルに 0.2 倍容量の 6 倍濃度の停止バッファー(10% SDS,

36% グリセロール、0.5 M DTT、0.175 M Tris-HCl [pH 6.8]、0.012% ブロモフェノールブルー) を添加し、反応を停止させた。さらに、反応を停止したサンプルを用いて、ランニングゲル(15% ポリアクリルアミドゲル、6 M 尿素、0.1 M リン酸ナトリウム [pH 7.2]、0.1% SDS)、スタッキングゲル(3.5% ポリアクリルアミドゲル、その他の組成はランニングゲルと同じ)、およびランニングバッファー(0.1 M リン酸ナトリウム [pH 7.2]、0.1% SDS)を使用した電気泳動を行った。最後に、電気泳動後のゲルを、クマシーブリリアントブルーで染色した。

3.3. 結果

3.3.1. トリプトファンを含まない遺伝暗号表による翻訳の確認

本研究では、トリプトファンを含まない 19 種類のアミノ酸セットがセンスコドンに対して割り当てられている「単純化遺伝暗号表」を構築し、機能を確認した。単純化遺伝暗号表中では UGG コドンがアラニンに再割り当てされているため、トリプトファンが含まれない(図 3.2.A)。再割り当ての最初のステップとして、私は S30 からトリプトファンを除き、かつトリプトファン tRNA 合成酵素(TrpRS)に対する強力な阻害剤である Trp-SA を加えることで、UGG コドンを空コドンにした。次に私は、UGG コドンに対応するアンチコドンループを持つ tRNA^{Ala} 変異体(図 3.2.B)を加えることで、UGG コドンをアラニンに再割り当てした。アラニル tRNA 合成酵素(AlaRS)はアンチコドンループを認識しない(65-67)ため、アラニンは AlaRS によって tRNA^{Ala} 変異体に結合させられると考えられる。実際、トリプトファン-tRNA^{Trp} が不足することでできてしまう、割り当てのない UGG コドンによる翻訳の停止(図 3.3、レーン 2、3)および、tRNA^{Ala} 変異体を添加することでの翻訳の再開(図 3.3、レーン 4-6)が観察されたことから、アラニンの UGG コドンへの再割り当ては成功していることが推測された。

3.3.2. 合成されたタンパク質がトリプトファンを含まないことの確認

アラニンの UGG コドンへの再割り当てを確認するために、私は 2 つのタンパク質、すなわち 8 つの UGG コドンを持つ MBP の mRNA を単純化遺伝暗号表で翻訳したタンパク質(MBP-WT/Sim)と、同じ mRNA を普遍遺伝暗号表で翻訳したタンパク質(MBP-WT/Univ)のアミノ酸組成を比較した(図 3.4.A)。明らかなトリプトファンのピークがある MBP-WT/Univ と対照的に、MBP-WT/Sim と、すべての UGG コドンを GCU コドンに置換した MBP の mRNA を普遍遺伝暗号表で翻訳した変異体(W_all_A/Univ)では、バックグラウンドレベルのトリプトファンのピークしか検出されなかった。トリプトファンのピークの消失は、MBP-WT/Sim および MBP-WT/Univ の蛍光スペクトル観察によっても確認された(図 3.6)。いっぽうで、別のアミノ酸組成分析の結果から、MBP-WT/Sim と MBP-WT/Univ のそれぞれのアミノ酸組成を比べると、MBP-WT/Sim におけるアラニンの組成のみが増加していることが分かった(図 3.4.B)。実際に、MBP-WT/Sim におけるアラニンの増加分は MBP-WT の mRNA 中の UGG コドンの数と一致すると予想された(表 3.1.)。

3.3.3. トリプトファンを含まない遺伝暗号表が普遍遺伝暗号表と同等の効率でタンパク質を合成できることの確認

単純化遺伝暗号表の翻訳の効率と正確性が、普遍遺伝暗号表のそれらと同等であることを示すため、以下の3つの遺伝子、すなわち GFPS1、R2-4、および R2-4 の 110 番目のアラニンコドン を UGG に変異させた R2-4_A110X(UGG) を単純化遺伝暗号表と普遍遺伝暗号表それぞれから翻訳させ、タンパク質の蛍光強度を比較した(図 3.7.)。R2-4 の mRNA が翻訳の鋳型として用いられると、単純化遺伝暗号表から翻訳された R2-4/Sim も、普遍遺伝暗号表から翻訳された R2-4/Univ も、同等の蛍光強度を示した。また、R2-4_A110X(UGG) を単純化遺伝暗号表で翻訳した R2-4_A110X(UGG)/Sim もまた、R2-4/Sim および R2-4/Univ と同等の活性を示した。これらの結果から、単純化遺伝暗号表は UGG コドンを含むほとんどのコドンを普遍遺伝暗号表と同等の効率と正確性で翻訳するということが示唆される。単純化遺伝暗号表とは対照的に、普遍遺伝暗号表は R2-4_A110X(UGG) の mRNA から活性を持ったタンパク質を翻訳することができなかった。なぜなら、R2-4_A110X(UGG)/Univ では、R2-4/Univ、そしておそらく R2-4/Sim や R2-4_A110X(UGG)/Sim においてアラニンが配置されている 110 番目のアミノ酸にトリプトファンが配置されるためである。R2-4_A110X(UGG) の mRNA を翻訳したときの単純化遺伝暗号表と普遍遺伝暗号表の関係は、GFPS1 の mRNA を翻訳したときの関係と逆である。GFPS1 の mRNA を単純化遺伝暗号表で翻訳したとき、おそらくアラニンは 57 番目に位置する UGG コドンに対応して導入される。GFPS1 における 57 番目のアミノ酸はタンパク質内部に位置しており、アラニンは gfpS1/Univ の活性に必要なトリプトファンを置換してしまう。

3.3.4. トリプトファンを含まない遺伝暗号表から翻訳されたタンパク質の、全体の構造および UGG コドンへの割り当ての結晶構造解析による確認

単純化遺伝暗号表から翻訳されたタンパク質の全体的な構造が正しいこと、また UGG コドンへの Ala の再割り当てが正しく行われていることを確認するため、単純化遺伝暗号表由来の R2-4_A110X(UGG)/Sim および普遍遺伝暗号表由来の R2-4/Univ の結晶構造を決定し、比較した。2つの構造の R.M.S.D は 0.151 Å であり、2つのタンパク質の構造はほぼ同一であることが示唆された(図 3.8.A)。ここで見られる構造の同一性は、正しくフォールドするタンパク質を単純化遺伝暗号表が翻訳できていることを示している。さらに、R2-4_A110X(UGG)/Sim における 110 番目のアミノ酸残基と、R2-4/Univ における 110 番目のアミノ酸であるアラニンが同一であるように見えることが、2つのタンパク質の電子密度を比べることで分かった(図 3.8.B)。いっぽうで、R2-4_A110X(UGG)/Sim における 110 番目のアミノ酸残基は、かさ高いトリプトファンやフェニルアラニンの残基とは明らかに異なることから、UGG コドンのアラニンへの再割り当てが起こっていることを示している。

3.3.5. 19種類のアミノ酸セットからなる遺伝暗号表の構築原理の一般性の確認

別の再割り当て、つまりシステインコドン UGU/UGC に対してセリンが再割り当てされる単純化遺伝暗号表の構築を、tRNA^{Ser} 変異体を用いて行った。tRNA^{Ser} 変異体は、UGU/UGC コドンに対応するアンチコドンループを持つ(図 3.9.A)。セリル tRNA 合成酵素(SerRS)はアンチコドンループを認識しない(68-70)ため、セリンは SerRS によって tRNA^{Ser} 変異体に結合させられると考えられる。再割り当てを確認するため、5 つの UGU/UGC を含む CAT の mRNA を、システインを含まない S30 中で翻訳した。この再割り当てにおいても、システイン-tRNA^{Cys} が不足することでできてしまう、割り当てのない UGU/UGC コドンによる翻訳の停止および、UGU/UGC コドンに対応するアンチコドンループを持つ tRNA^{Ser} 変異体を添加することでの翻訳の再開が観察されたことから、セリンの UGU/UGC コドンへの再割り当ては成功していることが推測された(図 3.9.B)。さらに再割り当てを確認するために、私は 2 つのタンパク質、すなわち CAT の mRNA を単純化遺伝暗号表で翻訳したタンパク質 (CAT-WT/Sim)と、同じ mRNA を普遍遺伝暗号表で翻訳したタンパク質(CAT-WT/Univ)のアミノ酸組成を比較した。CAT-WT/Univ のアミノ酸組成に対して、CAT-WT/Sim ではシステインの数が 5 つ減り、セリンの数が 5 つ増えていた(図 3.9.C および表 3.3.)。

システインコドン UGU/UGC に対してセリンが再割り当てされる単純化遺伝暗号表の翻訳の効率と正確性もまた、普遍遺伝暗号表のそれらと同等であることを示すため、大腸菌由来のプロテアーゼである LexA の活性評価を行った。野生型の LexA の mRNA は、UGU/UGC コドンを持たず、また野生型の LexA タンパク質は UCG コドンによってコードされるセリンを 119 番目のアミノ酸としてタンパク質の活性中心に持ち、このセリンが基質の加水分解に必要不可欠な役割を果たしている(71)。自己切断反応を防ぐため、私は以下 LexA*と表記される LexA-G85D(53)を作成した。私は LexA*の mRNA に対してさらに変異導入を行い、119 番目のアミノ酸に対応するコドンを UGC に置換した LexA*-S119X(UGC)を作成した。この変異体 mRNA を単純化遺伝暗号表で翻訳したタンパク質 LexA*-S119X(UGC)/Sim(図 3.10.、レーン 1、2)は、UGU/UGC の Ser への再割り当ての想定から期待されたとおり、LexA*-WT/Univ(図 3.10.、レーン 3、4)と同等の効率で基質を切断した。対照的に、LexA*-S119X(UGC)の mRNA を普遍遺伝暗号表で翻訳したタンパク質 LexA*-S119X(UGC)/Univ(図 3.10.、レーン 5、6)は、基質切断活性を全く持たなかった。

3.4. 結果のまとめと考察

本章では、トリプトファンを除去した無細胞翻訳系中に TrpSA および tRNA 変異体を加えることで、トリプトファンを含まない遺伝暗号表が構築できることを示した。最初に私は S30 からトリプトファンを除き、トリプトファンと UGG コドンを結ぶ内在性の翻訳経路を遮断した。次に私は tRNA^{Trp} のアンチコドンループ配列を持った tRNA^{Ala} 変異体を作成し加えることで、空コドンになった UGG コドンに対してアラニンに再割り当てした。この単純化遺伝暗号表の翻訳の効率と正確性は、普遍遺伝暗号表のそれらと同等であった(図 3.7.)。

また、システインを含まない単純化遺伝暗号表も、 tRNA^{Cys} のアンチコドンループ配列を持った tRNA^{Ser} を用いた類似の構築手法で作成できることを示した。これらの結果は、適切な翻訳経路が存在すれば、20 種類以下のアミノ酸セットからなる単純化遺伝暗号表は構築可能であり、そして活性のあるタンパク質を合成することも可能であることを示している。

今回用いた S30 のような無細胞翻訳系を用いるメリットは、系の構成物を容易に加えたり除いたりすることができる点にある。遺伝暗号表からトリプトファンを除くためには、 tRNA^{Trp} のアミノアシル化を阻害する必要がある。もし大腸菌などの生きた細胞の中でアミノアシル化を阻害すれば、必須遺伝子から翻訳されるタンパク質の合成が途中で停止し、宿主に致命的な打撃を与えてしまうことが容易に想像できる。この問題に対処するために、私はトリプトファンを除いた S30 を使用した。私は、条件によっては除去したアミノ酸の生合成経路に関わる酵素の活性も阻害すべき(72)と考えた。実際、システインを含まない単純化遺伝暗号表の構築においては、システインのない条件下において、微弱なバンドが観察された(図 3.9.B、レーン 2)。このバンドは、おそらく新たに合成されたシステインがタンパク質に導入されてできたものであろう。私は Cys-SA の添加によって CysRS の阻害をすることで、新たに合成されたシステインの tRNA^{Cys} へのチャージを阻害した。この操作を行うと、微弱なバンドが期待したとおり消失した(図 3.9.B、レーン 3)。

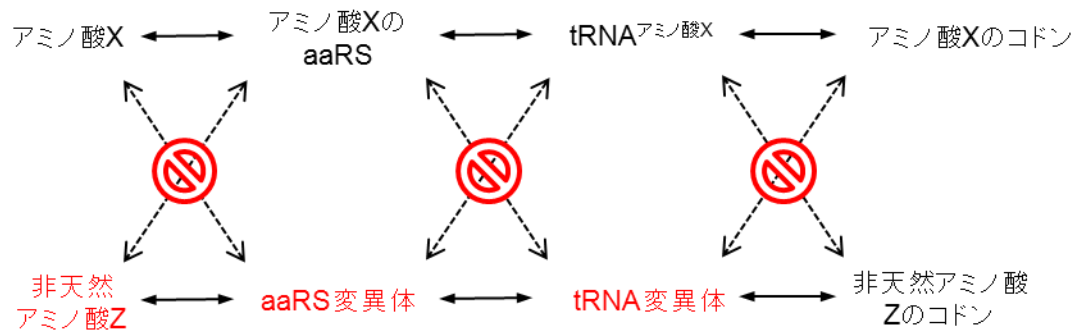
システインコドンに対してセリンを割り当てる単純化遺伝暗号表は、トリプトファンコドンにアラニン割り当てる単純化遺伝暗号表と同様、過去に存在した遺伝暗号表を再現しているかもしれない。アミノ酸の生合成経路の発達に伴って、遺伝暗号表に含まれるアミノ酸が徐々に増加したとする Wong の co-evolution theory(11)によれば、Cys が遺伝暗号表に導入されるまでは UGU/UGC コドンに Ser が割り当てられていたとされている。この想定は、ある種のメタン合成細菌が tRNA^{Cys} にホスホセリンをミスチャージし、その後酵素反応によってシステインに変換する反応を行っている(29)ことから支持される。 tRNA へアミノ酸をミスチャージした後に、酵素的にアミノ酸を別のアミノ酸に変換する反応は、セレノシステインにおいても観察されている(73)。したがって、システインコドン UGU/UGC コドンに Ser を割り当てる単純化遺伝暗号表は、普遍遺伝暗号表が完成する前の過去の遺伝暗号表の 1 つを再現しているかもしれない。

いくつかの天然および改変型の生物は、使用するアミノ酸の種類は 20 種類のままであるものの、普遍遺伝暗号表中で用いられる tRNA とはアンチコドン配列の異なる tRNA 変異体を本研究と同様に用いている。例えば *C. cylindracea* は、ロイシンコドンの 1 つである CUG コドンを、CAG アンチコドンを持った tRNA^{Ser} ($\text{tRNA}^{\text{Ser}}\text{CAG}$) を用いてセリンに割り当てている(74,75)。 $\text{tRNA}^{\text{Ser}}\text{CAG}$ は CUG コドンをめぐって野生型の tRNA^{Leu} と競合し、1 億年ほどの間、CUG コドンをロイシンとセリンの曖昧指定にしていた(76)と考えられている。競合の結果として $\text{tRNA}^{\text{Ser}}\text{CAG}$ が選ばれ、*C. cylindracea* において CUG コドンをセリンが獲得したのである(図 3.11.)。この *C. cylindracea* の CUG コドンをめぐる競合に似た *in vivo* での実験において、 tRNA^{Ala} 変異体や tRNA^{Ser} 変異体が用いられている(77,78)。これらの研究では、

tRNA 変異体と、センスコドンに割り当てられていた内在性の tRNA がやはりコドンをめぐる競争し、コドンの曖昧指定を引き起こしている。*C. cylindracea* において tRNA^{Ser} (tRNA^{Ser}CAG)が野生型の tRNA^{Leu} を排除したように、これらの研究でも競争する内在性の tRNA を tRNA 変異体が排除すれば、*in vivo* におけるコドンの完全な再割り当てが実現するだろう。*in vivo* におけるコドンの再割り当てについては、総合討論 4.3.1.項で再度議論する。

3.5. 図表

A



B

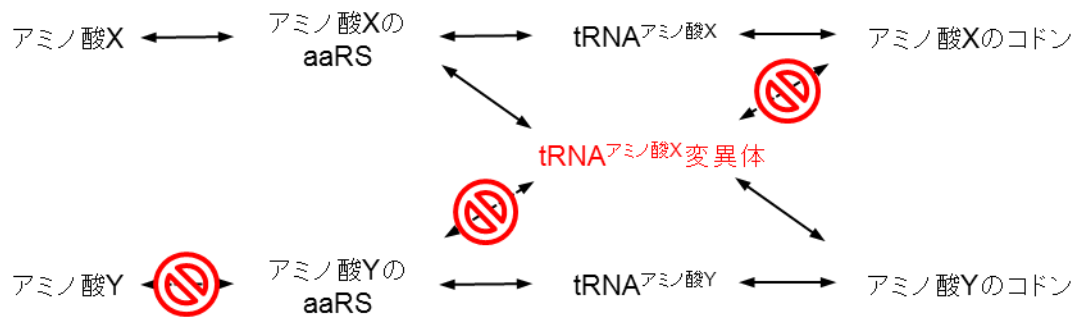


図 3.1.(拡張遺伝暗号表と単純化遺伝暗号表の対比)

tRNA ^{Ala} variant (μM)	-	-	-	0.1	0.25	2.5
Trp	+	-	-	-	-	-
Trp-SA	-	-	+	+	+	+
Lane	1	2	3	4	5	6

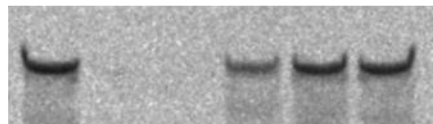


図 3.3. tRNA^{Ala} 依存的なタンパク質合成の再開

レーン 1: 20 種類のアミノ酸を加えた条件での翻訳。レーン 2: トリプトファン以外のアミノ酸を加えた条件での翻訳。レーン 3: トリプトファン以外のアミノ酸を加え、さらに Trp-SA を加えた条件での翻訳。レーン 4-6: レーン 3 の条件に、tRNA^{Ala} 変異体を表記された濃度で加えた時の翻訳。

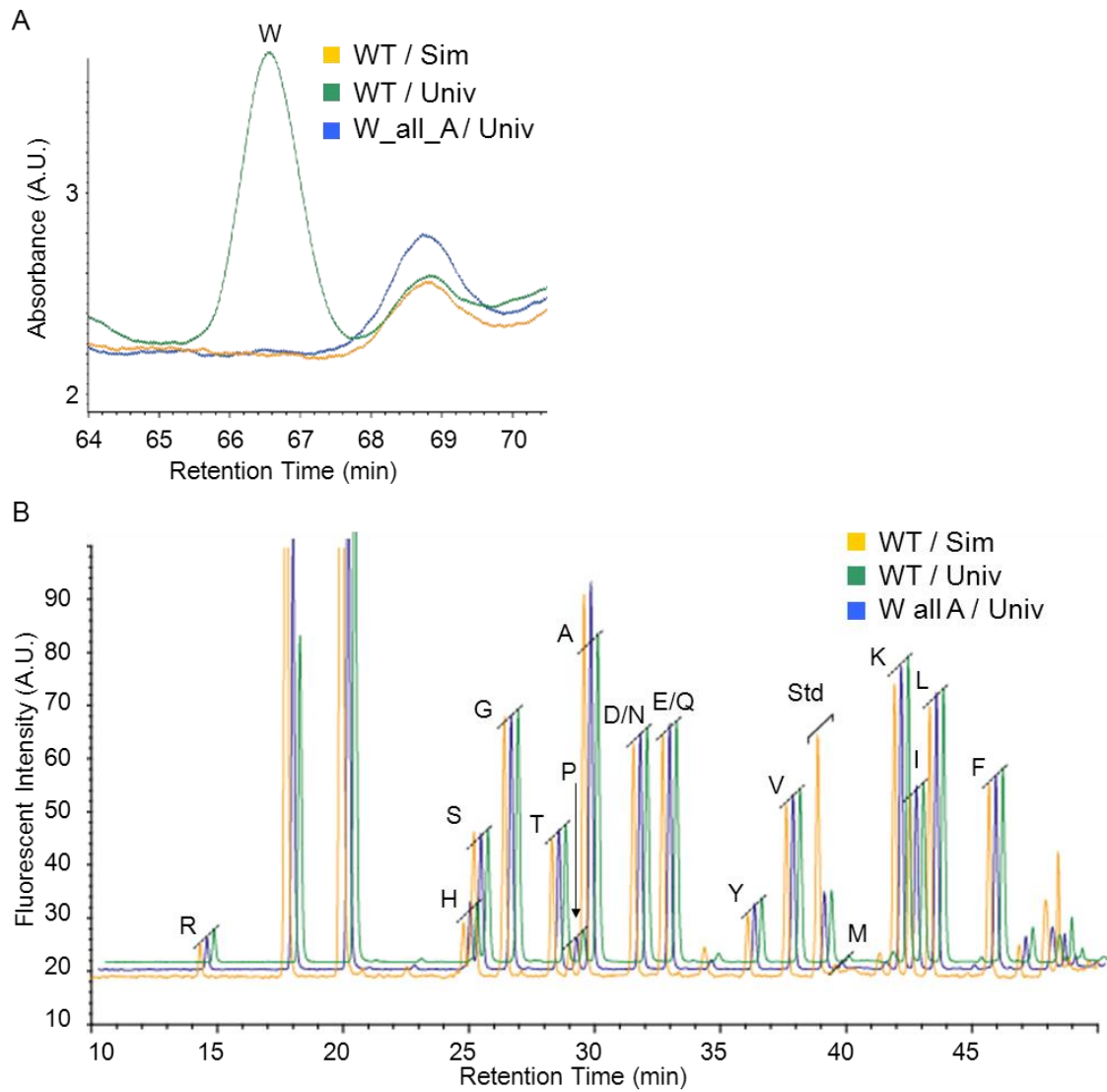


図 3.4. アミノ酸組成分析

A. トリプトファンをピークを検出するための MSA による加水分解産物のクロマトグラム。クロマトグラムの全体図は図 3.5 に示した。B. HCl による加水分解産物のクロマトグラム。

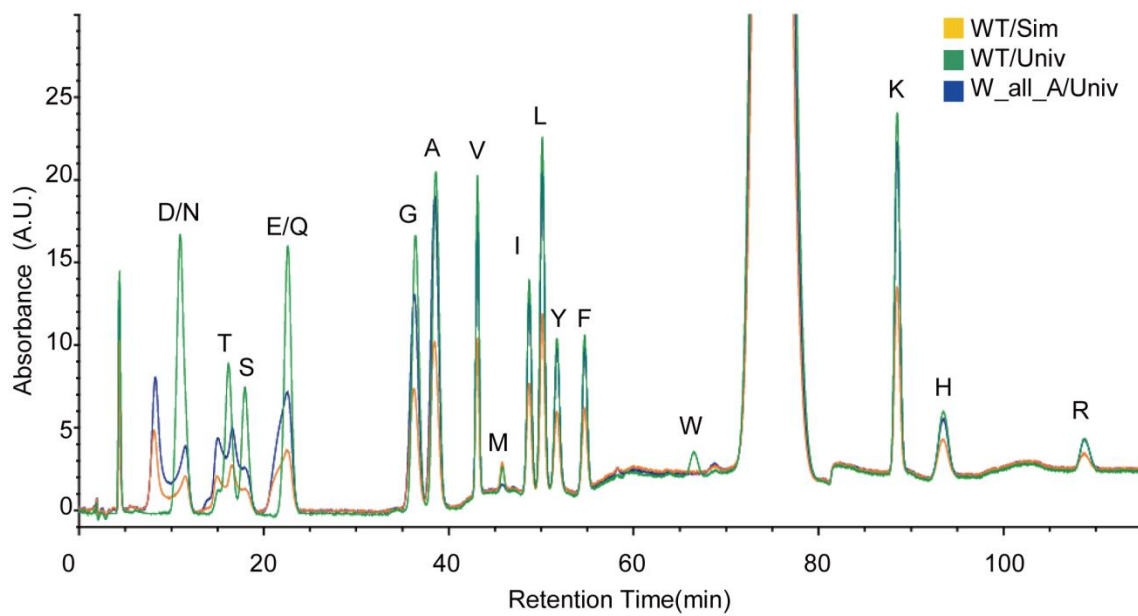


図 3.5. MSA による加水分解産物のクロマトグラムの全体像

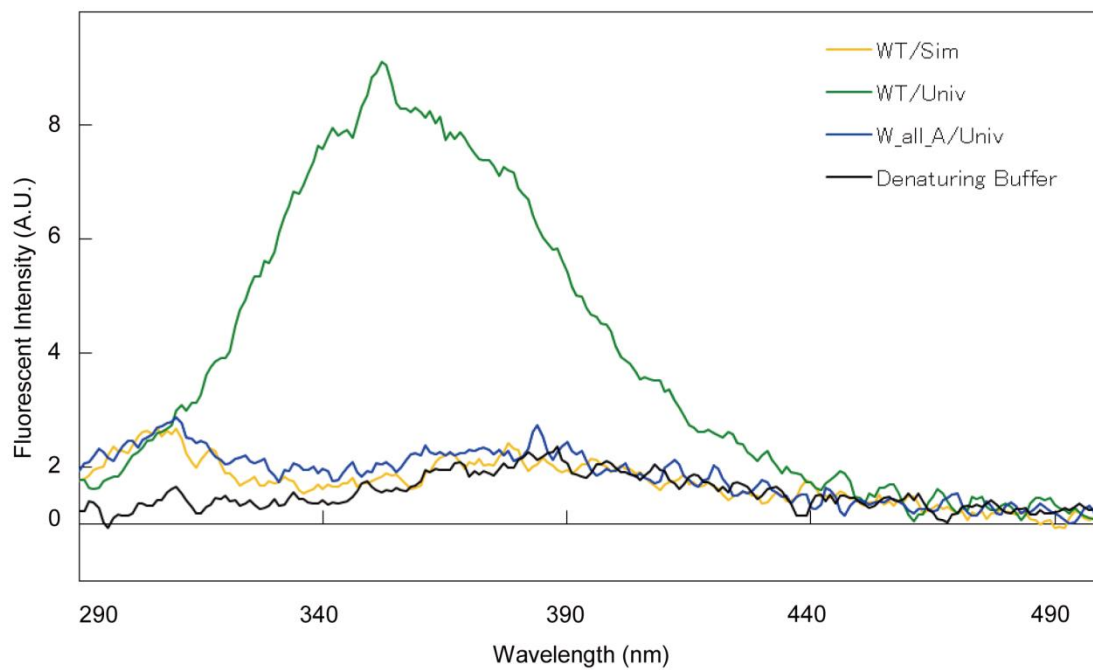
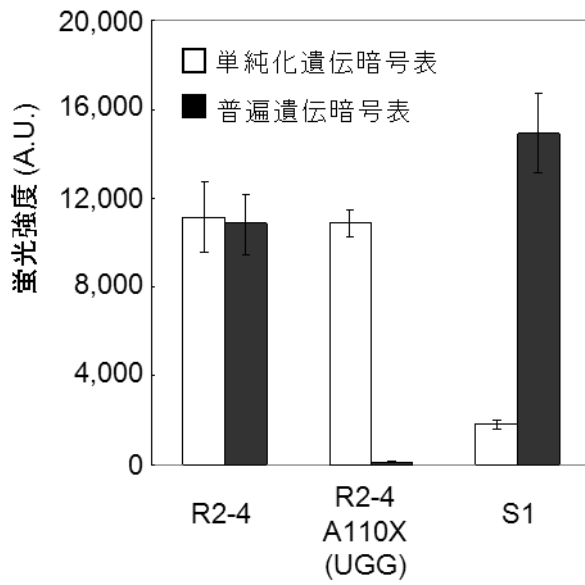


図 3.6. トリプトファンの蛍光ピーク

A.



B.

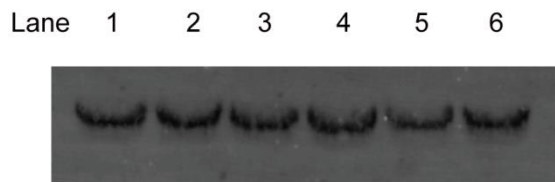


図 3.7. 単純化遺伝暗号表、普遍遺伝暗号表から翻訳された GFP の蛍光強度の比較
A. GFP の蛍光強度の比較。B. GFP の蛍光強度の比較に用いたタンパク質の合成量の確認

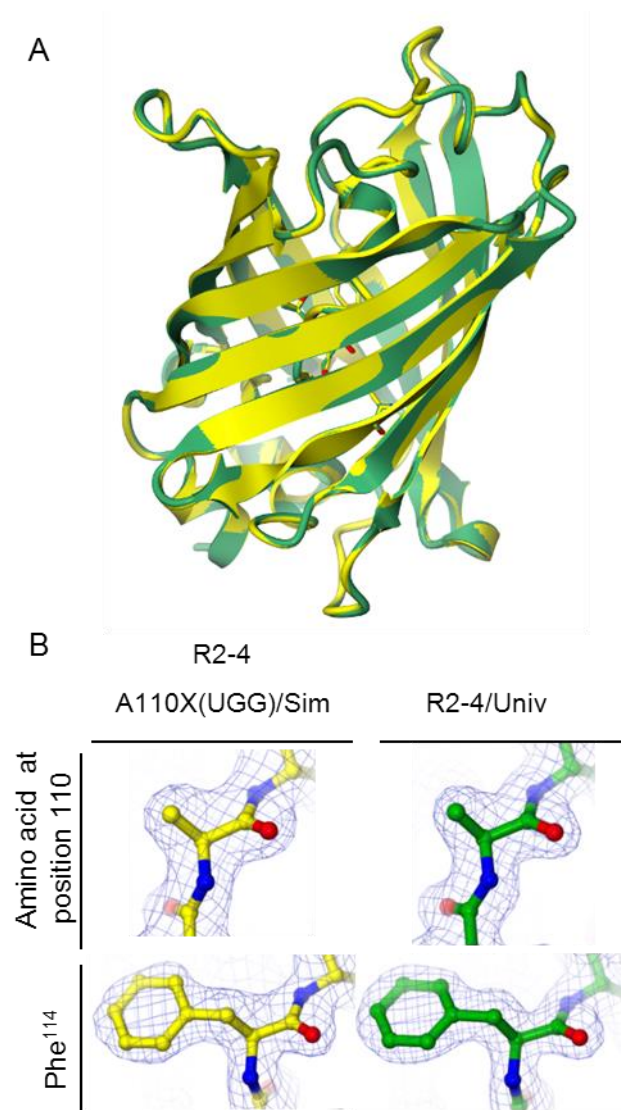


図 3.8. 単純化遺伝暗号表、普遍遺伝暗号表から翻訳された GFP の立体構造

A. 全体の構造の重ね合わせ。B. 110 番目のアミノ酸の比較と、114 番目のフェニルアラニンとの比較。

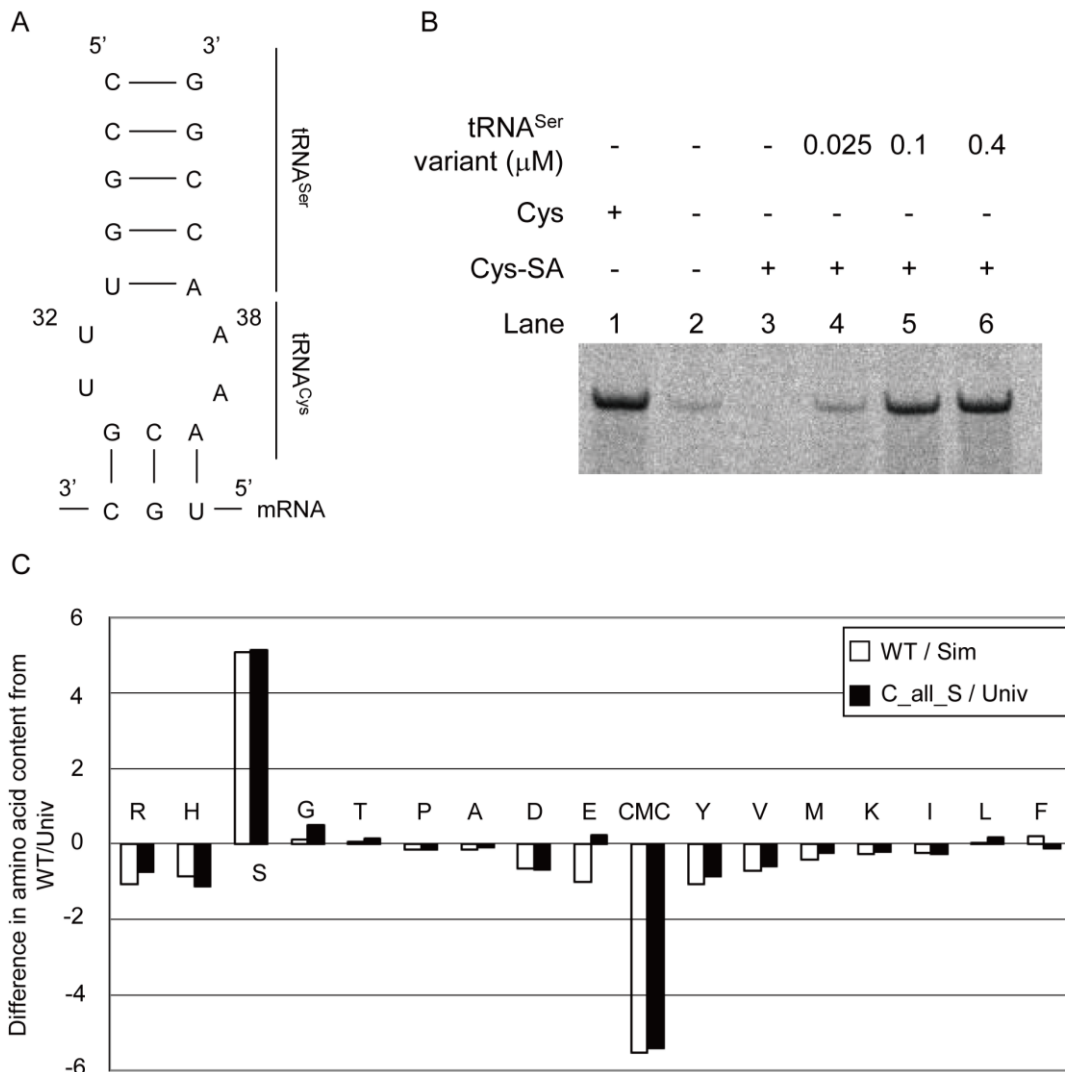


図 3.9. システインを含まない単純化遺伝暗号表

A. tRNA^{Ser}変異体のアンチコドンステムループと mRNA B. レーン1: 20種類のアミノ酸を加えた条件での翻訳。レーン2: システイン以外のアミノ酸を加えた条件での翻訳。レーン3: システイン以外のアミノ酸を加え、さらに Cys-SA を加えた条件での翻訳。レーン4-6: レーン3の条件に、tRNA^{Ser}変異体を表記された濃度で加えた時の翻訳。C. アミノ酸組成分析の結果を、WT/Univ との差で表示したもの。

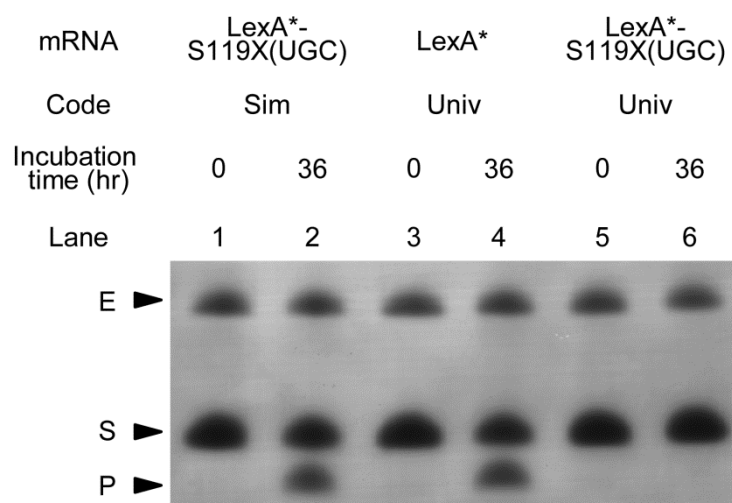


図 3.10. LexA による基質の切断アッセイ

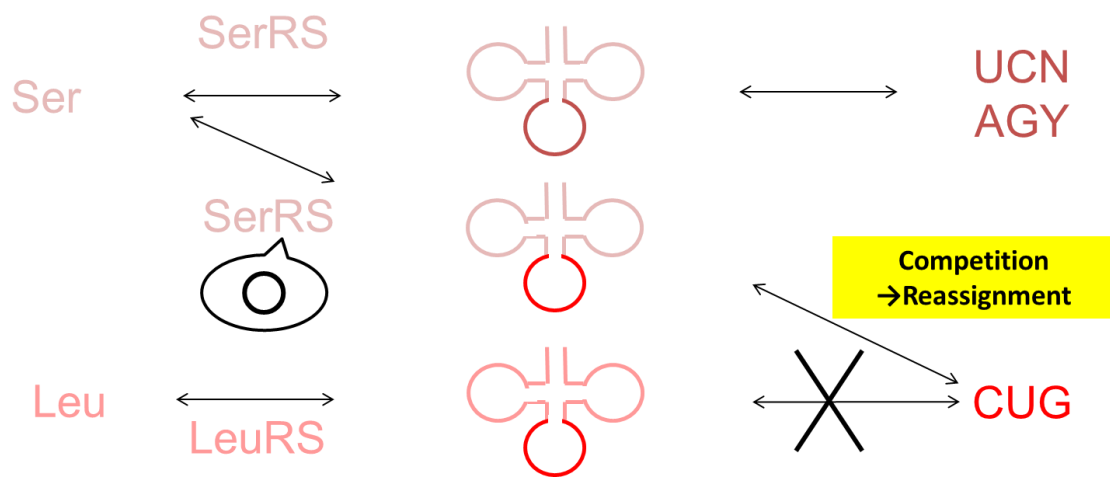


図 3.11. tRNA^{Ser} が Leu コドンを奪った例

表 3.1. HCl による加水分解を行った MBP のアミノ酸組成

Amino acid content (residues / mol)

Residue	Theoretical	WT/Univ	W_all_A /Univ	WT/Sim
Ala (A)	45	45.7	53.3	53.8
Arg (R)	6	5.9	5.7	6.0
His (H)	13	9.8	11.2	9.4
Ser (S)	20	17.7	17.8	19.5
Gly (G)	35	34.5	34.2	35.7
Thr (T)	19	17.9	18.2	18.4
Pro (P)	23	19.3	19.5	20.4
Asp/Asn (D/N)	46	45.0	45.0	44.7
Glu/Gln (E/Q)	38	40.5	41.1	41.4
Tyr (Y)	16	13.8	14.1	13.5
Val (V)	21	20.0	20.0	19.9
Lys (K)	36	35.0	34.6	34.0
Ile (I)	22	19.7	19.9	19.9
Leu(L)	32	32.2	32.3	32.0
Phe (F)	16	15.9	15.9	16.0

Asp/Asn (D/N) は Asp (D) と Asn (N) の和

Glu/Gln (E/Q) は Glu (E) と Gln (Q) の和

表 3.2. GFP の結晶テーブル

データコレクションと精緻化の際の統計データ

	R2-4_A110X(UGG)/Sim	R2-4/Univ
Data collection		
Beamline	SPring-8 BL41XU	SPring-8 BL41XU
Wavelength (Å)	1.000	1.000
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell parameters	a = 51.78 b = 63.05 c = 67.40	a = 51.84 b = 62.96 c = 67.37
	$\alpha = \beta = \gamma = 90^\circ$	$\alpha = \beta = \gamma = 90^\circ$
Resolution (Å)	50.00-2.10 (2.14-2.10)	50.00-1.85 (1.88-1.85)
Unique reflections	13,473	19,367
Redundancy	9.5 (5.7)	10.5 (6.8)
Completeness (%)	99.6 (97.7)	99.6 (99.1)
$I/\sigma(I)$	30.6 (3.7)	51.8 (5.5)
R_{sym}	0.110 (0.387)	0.095 (0.479)
Refinement		
Resolution (Å)	33.70-2.09	46.00-1.85
R_{work}/R_{free}	18.8/24.0	18.8/23.8
No. of atoms		
Protein	1796	1807
Water	74	126
Average B-factor (Å ²)		
Protein	31.5	29.2
Water	44.1	42.9
R.m.s. deviation		
Bond length (Å)	0.022	0.022
Bond angle (°)	2.36	2.30
Ramachandran plot (%)		
Most favored	96.30	96.35
Allowed	3.70	3.65
Generously allowed	0.00	0.00
PDB ID code	3UG0	3UFZ

Values in parentheses are for the highest resolution shell

表 3.3. HCl による加水分解を行った CAT のアミノ酸組成

Residue	Amino acid content (residues / mol)				Difference in amino acid content from WT/Univ (residues / mol)	
	Theoretical	WT/Univ	C_all_S/Univ	WT/Sim	(C_all_S/Univ) - (WT/Univ)	(WT/Sim) - (WT/Univ)
CMC	5	6.0	0.6	0.5	-5.4	-5.5
Ser (S)	15	12.3	17.5	17.4	5.1	5.1
Arg (R)	6	6.6	5.8	5.5	-0.8	-1.1
His (H)	19	18.8	17.7	18.0	-1.1	-0.9
Gly (G)	13	12.2	12.7	12.3	0.5	0.1
Thr (T)	13	12.8	13.0	12.9	0.1	0.1
Pro (P)	8	7.9	7.8	7.8	-0.1	-0.1
Ala (A)	15	15.3	15.2	15.2	-0.1	-0.1
Asp/Asn (D/N)	22	22.3	21.6	21.6	-0.7	-0.7
Glu/Gln (E/Q)	25	26.7	26.9	25.7	0.2	-1.0
Tyr (Y)	11	10.7	9.9	9.7	-0.9	-1.0
Val (V)	17	16.9	16.3	16.2	-0.6	-0.7
Lys (K)	12	11.8	11.6	11.5	-0.2	-0.3
Ile (I)	9	8.5	8.2	8.3	-0.3	-0.2
Leu (L)	14	14.1	14.2	14.1	0.2	0.0
Phe (F)	20	19.5	19.4	19.7	-0.1	0.2

Asp/Asn (D/N) は Asp (D) と Asn (N) の和

Glu/Gln (E/Q) は Glu (E) と Gln (Q) の和

第四章 総合討論

4.1. 結果のまとめ

本研究は、普遍遺伝暗号表が成立する前に存在したと考えられている、トリプトファンを含まないタンパク質の配列最適化およびトリプトファンを含まない遺伝暗号表の構築を行った。第二章では、20種類のアミノ酸セットからなるタンパク質を、トリプトファンを置換する変異導入によってトリプトファンを除去し、トリプトファンを含まないアミノ酸配列であっても配列最適化が起こることを示した。第三章では、トリプトファンを含まないS30中にtRNA^{Trp}のアンチコドンループ配列を持つtRNA^{Ala}を添加することで、トリプトファンを含まない単純化遺伝暗号表を構築した。さらに、この単純化遺伝暗号表の翻訳の効率と正確性が普遍遺伝暗号表のそれらと同等であることを示し、また類似の構築手法でシステインを含まない遺伝暗号表も構築できることを示した。これらの結果から、適切な翻訳経路が存在すれば、20種類以下のアミノ酸セットからなる単純化遺伝暗号表は構築可能であり、そして活性のあるタンパク質を合成することも可能であることが分かる。本研究の一連の結果から、遺伝暗号表の進化の後期に存在したと推測されている、トリプトファンを含まない遺伝暗号表は翻訳システムとして成立しうるものであり、またそこから翻訳されるタンパク質はトリプトファンなしに進化可能であることが実験的に裏付けられた。

4.2. 考察

4.2.1. トリプトファンが遺伝暗号表に導入された意義

トリプトファンが遺伝暗号表に導入されたのは、タンパク質の疎水性コアを効率的に形成するために、かさ高い疎水性アミノ酸が必要とされたからかもしれない。普遍遺伝暗号表がアミノ酸セットを構成する一員としてトリプトファンを採用しているにも関わらず、タンパク質はトリプトファンなしに人工進化できる。例えば、過去の研究は遺伝子全体にランダム変異を導入することでの19種類のアミノ酸からなるタンパク質の *directed evolution* を示している(33,79)。遺伝子配列の一部のみへの変異導入ではあるものの、他の報告では18種類のアミノ酸からなるタンパク質でさえトリプトファンなしに進化できることを示唆している(80)。トリプトファンを含まない19種類のアミノ酸からなるタンパク質は天然の進化を通じても生まれうる。例えば、大腸菌が持つタンパク質のうち、トリプトファンを含まない19種類のアミノ酸からなるタンパク質の種類は全体の11%であり、システインを含まない19種類のアミノ酸からなるタンパク質の種類は全体の15%に次ぐ2番目に多い割合である(81)。さらに、Pezoらによれば、トリプトファンは大腸菌の酵素の触媒部位に必要不可欠というわけではない(82)。もしこの主張が正しければ、なぜトリプトファンは遺伝暗号表に加わったのだろうか、という問いが生まれる。20種類のアミノ酸のうち、最もかさ高いアミノ酸であるトリプトファンの導入は、他のGroELの研究(83)で示唆されているように、疎水性相互作用を強化したかもしれない。FournierとGogartenは、遺伝暗号表にアミノ酸が加わっていく過程で最も特徴的な傾向の1つは、芳香環を持つアミノ酸の増

加だと述べている(84)。彼らは、芳香環がタンパク質の疎水性コアのパッキングに貢献し、フォールディングを助けたのだと主張している。ここで、タンパク質を半径 r の球体と仮定すると、タンパク質の表面積と体積は以下の式で表現される。

$$S = 4\pi r^2$$

$$V = \frac{4}{3}\pi r^3$$

$$\frac{V}{S} = \frac{r}{3}$$

S と V はそれぞれ表面積と体積を表す。遺伝子の伸長時、 r の増大によって V/S は大きくなる。つまり、遺伝子の伸長によるタンパク質の疎水性コアの体積の増加分は、表面積の増加分に比べて大きい、ということが言える。もし、遺伝子の伸長時にタンパク質を構成する疎水性アミノ酸の数と種類が変わらなければ、疎水性コアの体積が足りずタンパク質はフォールドできなかつたであろう。疎水性コアの体積の不足を防ぐためには、以下の変化が必要となる。すなわち、1) タンパク質中での疎水性アミノ酸の出現頻度を高める、2) タンパク質に大きいサイズの疎水性アミノ酸を導入する、である。遺伝暗号表にトリプトファンが導入されることは、1)と2)を同時に満たす。Wong の co-evolution theory(11)によれば、UGG コドンはもともと親水性のセリンに割り当てられており、そこへ疎水性のトリプトファンが新たに導入されたとされているので、遺伝暗号表にトリプトファンを導入することは1)を満たす。またトリプトファンはアミノ酸の中で最も大きいファンデルワールス半径を持つアミノ酸であるので、2)も同時に満たす。このように、かさ高い疎水性アミノ酸であるトリプトファンは、タンパク質の疎水性コアを効率的に埋めただろうと考えられる(85)。もしタンパク質の長さが一定であると仮定すると、疎水性アミノ酸によるタンパク質コアの効率的な占有は、より多くの親水性のアミノ酸残基を、タンパク質 - タンパク質相互作用を担うタンパク質表面に配置できたであろう(図 4.1.)。この観点から、最もかさ高いアミノ酸であるトリプトファンを遺伝暗号表に導入したことは、現在遺伝暗号表に用いられている他のどのアミノ酸を導入することよりも意義深かつたであろう。

芳香環を持つトリプトファンは、カチオン- π 相互作用を通してタンパク質の機能の多様化とタンパク質の構造安定化に貢献してきた。芳香族アミノ酸の中でも、トリプトファンが最も優れた電子供与能を持つことは、古くから知られている(86)。この電子供与能はカチオン- π 相互作用に用いられ、たとえばニコチンアミドアセチルコリン受容体がアセチルコリンに結合する際に寄与(87)するように、タンパク質の分子認識に貢献している。またより一般的には、タンパク質内部で電荷をもつ他のアミノ酸との相互作用によって立体構造の安定化に寄与している。このように、トリプトファンの電子供与体としての役割が遺伝暗号表に求められ、20番目のアミノ酸として導入されるにいたったのかもしれない。

4.2.2. 遺伝暗号表の違いが水平伝播の障壁になる

単純化遺伝暗号表と普遍遺伝暗号表のような異なる遺伝暗号表の間では、他生物由来の遺伝子の授受である水平伝播が機能しない(図 4.2.)。R2-4_A110X(UGG)の mRNA からは単純化遺伝暗号表を通して蛍光強度を持つ GFP 変異体が翻訳されたが、普遍遺伝暗号表を通しては蛍光強度をほとんど持たない GFP 変異体が翻訳された。対照的に、GFPS1 の mRNA からは普遍遺伝暗号表を通して蛍光強度を持つ別の GFP 変異体が翻訳されたが、単純化遺伝暗号表を通しては蛍光強度をほとんど持たない別の GFP 変異体が翻訳された。トリプトファンが遺伝暗号表に取り入れられる前の生物を想像すると、この遺伝暗号表と mRNA の 1 対 1 の関係は、生物間での遺伝子の水平伝播を妨げる要因になったかもしれない(88)。トリプトファンを遺伝暗号表に導入し、20 種類のアミノ酸セットを使い始めた生物は、これまでにないアミノ酸を使えるようになったことで、何らかの面で生存に有利な形質を獲得したと考えられる(89)。そのいっぽうで、遺伝暗号表にトリプトファンを導入していない周囲の生物から、遺伝子を水平伝播によって得る機会を失ったであろう。20 種類の生物が集団中の多数を占めるほどに広まった時期を想定すれば、遺伝子のやりとりの妨げとなる障壁は一転、遺伝暗号表にトリプトファンをいまだ導入できていない生物にとって、同様に周囲の生物から遺伝子を水平伝播によって得る機会を失わせることになる。事実、20 種類のアミノ酸を用いながら遺伝暗号表中の配置が異なる生物では、近年の進化において水平伝播の頻度が少ないことが、UGA コドンをストックコドンではなくトリプトファンコドンとして翻訳するマイコプラズマにおいて示唆されている(90,91)。工学的観点から見れば、単純化遺伝暗号表を通してのみ機能する人工的な遺伝子は、天然の生物が用いている普遍遺伝暗号表と互換性がなく、安全性の面で有用である。実験室における実験では、単純化遺伝暗号表とそこでのみ機能する人工的な遺伝子を用いるようにすれば、人工的な遺伝子が天然へ万一流出しても、水平伝播による遺伝子の拡散が妨げられる。

4.3. 展望

4.3.1. アミノ酸の種類が減少した遺伝暗号表を用いた Directed evolution

単純化遺伝暗号表は Directed evolution と組み合わせることで、トリプトファンを含まない単純化タンパク質をより効率的に探索させ、より一般性のある知見を得ることを助けるだろう。本研究のように、トリプトファンを含まない低活性変異体に対して Directed evolution によって活性向上を試みる場合、ランダム変異からトリプトファンコドンが再出現する可能性がある(図 4.3.)。類似の事例として、Tumor Necrosis Factor- α (TNF- α)に 6 つ含まれるリジンを Saturation mutagenesis によって置換し、リジンを含まない活性型 TNF- α 変異体を作成しようとした先行研究においては、選別された野生型と同等の活性を持つ 197 クローンのうち、191 クローンにおいて少なくとも 1 つのリジンが再出現していた(92)。変異前の配列に戻る復帰変異は変異体タンパク質の活性を向上させることが多いため、除去したいアミノ酸が含まれているにも関わらず Directed evolution にて高活性変異体として選択されてしまい、除去したいアミノ酸が含まれていることが遺伝子配列を調べるまで分からない、

という非効率を生む。しかし、コドンの再出現によって生まれる非効率は、単純化遺伝暗号表を用いることで回避できる。トリプトファンを含まない単純化遺伝暗号表を用いれば、どのような遺伝子配列に対してもトリプトファンを導入することがなくなるためである。今のところ単純化遺伝暗号表は *in vitro* の技術であり、先述の TNF- α を進化させた研究にて用いられている *phage display* のような *in vivo* の技術には直接応用はできないが、*in vitro* で大きなライブラリサイズを扱うことのできる *ribosome display* のような *Directed evolution* の手法と組み合わせることで、単純化タンパク質の効率的な探索に寄与するだろう。単純化遺伝暗号表と *Directed evolution* の併用による単純化タンパク質の探索の効率化は、原始的なタンパク質がどのような機能を持ちえたのか、複数のサンプルからのより確度の高い知見を得ることを可能にすると期待される。

4.3.1. トリプトファンを含まない 19 種類のアミノ酸セットからなる生物の創出

トリプトファンを含まない 19 種類のアミノ酸セットからなる生物が将来創出されると、遺伝暗号表の後期の進化の理解の観点から意義深い。生物に利用されるアミノ酸セットの数は減らせるのかどうかという問いに答えるため、Pezo ら(82)はトリプトファンコドンのアンチコドンを持つ tRNA^{His} 変異体を用いて、大腸菌に含まれるタンパク質からトリプトファンを除くことを試みた。もし、Pezo らが作成した大腸菌から、さらに TrpRS または tRNA^{Trp} を除去したら、UGG コドンは完全に再割り当てされることになる。しかし、Pezo らの研究がトリプトファンの完全な除去に至らなかったことから分かるように、あるアミノ酸を *in vivo* において完全に除去する実験戦略は、宿主を構成するタンパク質群の全体に深刻な影響を与え、宿主に致命的な打撃を与えることが容易に想像できる。トリプトファンを含まない生物を将来的に創るためには、トリプトファンを含まずに機能するタンパク質配列を複数得て、その後ゲノムの大規模な構築・編集技術(93)を用いて統合することが必要である。トリプトファンを含まずに機能するタンパク質配列を得るためには、多数の必須遺伝子を効率よく改変する必要がある、先述した単純化遺伝暗号表と *Directed evolution* の組み合わせを用いることが必須である。単純化遺伝暗号表と *Directed evolution* を組み合わせた研究において、19 種類のアミノ酸からなる活性型 GFP をターゲットタンパク質の C 末端に融合させれば、フォールドしたタンパク質と凝集したタンパク質を区別するレポーターとして利用できる(94)。19 種類のアミノ酸からなる活性型 GFP のフォールディングレポーターとしての利用は、活性型の 19 種類のアミノ酸からなるタンパク質を得ようとする *Directed evolution* において、初期変異体が凝集してしまう場合や、遺伝子全体へのランダム変異によって作成された変異体の多くが凝集してしまう場合に特に有用である。トリプトファンを含まない 19 種類のアミノ酸セットからなる生物が創出されれば、その生物をもとにした実験が可能になり、遺伝暗号表がなぜトリプトファンを導入したのかについての理解が進むだろう。

4.4. 図表

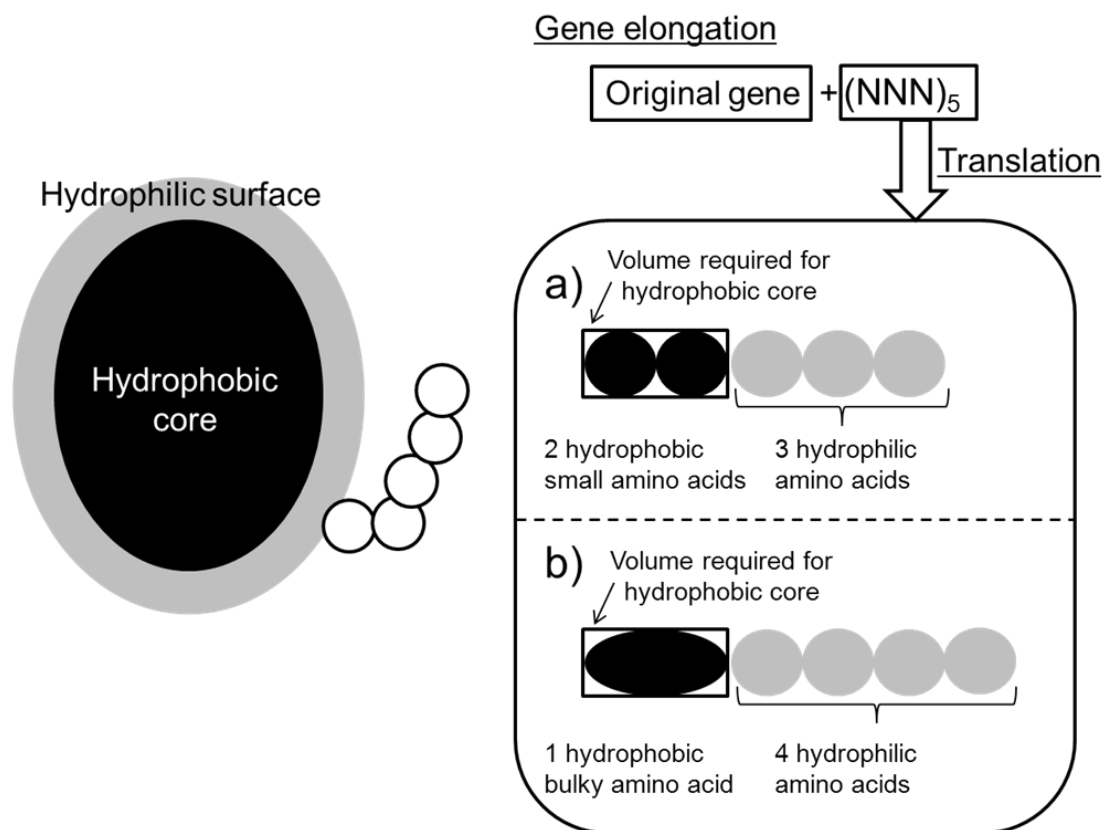


図 4.1. 新たなアミノ酸の導入が進化において新しい機能の獲得にどのように貢献するかを示した模式図

アミノ酸 5 残基が新たに加わったときを想定している。疎水性コアの領域に配置されるアミノ酸の数は、疎水性アミノ酸の数によって変化する。a. 2 つの小さな疎水性アミノ酸が疎水性コアを埋めるために必要とされ、3 つの親水性アミノ酸が親水性領域に配置され b. 1 つの大きな疎水性アミノ酸が疎水性コアを埋めるために必要とされ、4 つの親水性アミノ酸が親水性領域に配置される。親水性領域により多くのアミノ酸を配置できることは、タンパク質-タンパク質相互作用の多様化に貢献する。

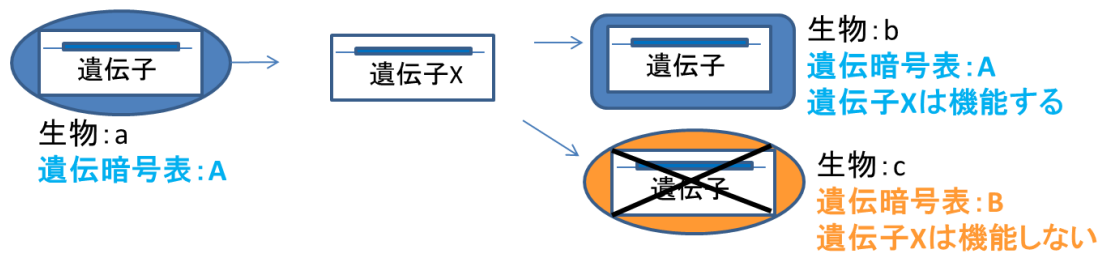
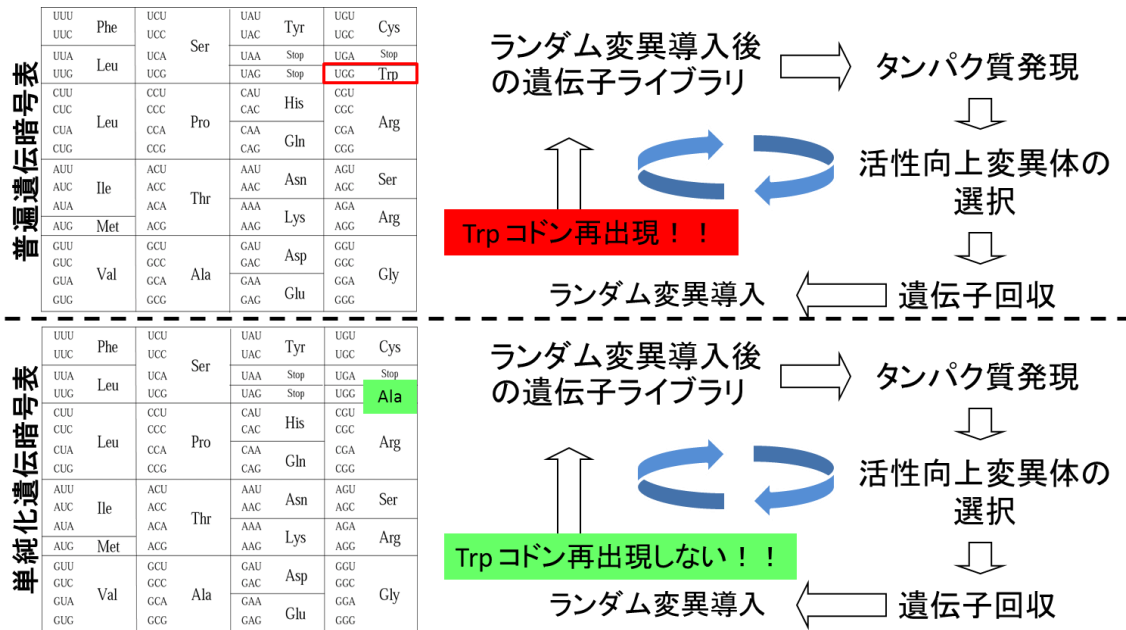


図 4.2. 異なる遺伝暗号表の間では、遺伝子の水平伝播が起こらないことを表した模式図



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. Akanuma, S., Kigawa, T. and Yokoyama, S. (2002) Combinatorial mutagenesis to restrict amino acid usage in an enzyme to a reduced set. *Proc Natl Acad Sci U S A*, **99**, 13549-13553.
4. Jordan, I.K., Kondrashov, F.A., Adzhubei, I.A., Wolf, Y.I., Koonin, E.V., Kondrashov, A.S. and Sunyaev, S. (2005) A universal trend of amino acid gain and loss in protein evolution. *Nature*, **433**, 633-638.
5. Miller, S.L. (1953) A production of amino acids under possible primitive earth conditions. *Science*, **117**, 528-529.
6. Cleaves, H.J., Chalmers, J.H., Lazcano, A., Miller, S.L. and Bada, J.L. (2008) A reassessment of prebiotic organic synthesis in neutral planetary atmospheres. *Orig Life Evol Biosph*, **38**, 105-115.
7. Yanagawa, H. and Kobayashi, K. (1992) An experimental approach to chemical evolution in submarine hydrothermal systems. *Orig Life Evol Biosph*, **22**, 147-159.
8. Huber, C. and Wachtershauser, G. (2006) alpha-Hydroxy and alpha-amino acids under possible Hadean, volcanic origin-of-life conditions. *Science*, **314**, 630-632.
9. Kvenvolden, K., Lawless, J., Pering, K., Peterson, E., Flores, J., Ponnampereuma, C., Kaplan, I.R. and Moore, C. (1970) Evidence for extraterrestrial amino-acids and hydrocarbons in the Murchison meteorite. *Nature*, **228**, 923-926.
10. Walter, K.U., Vamvaca, K. and Hilvert, D. (2005) An active enzyme constructed from a 9-amino acid alphabet. *J Biol Chem*, **280**, 37742-37746.
11. Wong, J.T. (1975) A co-evolution theory of the genetic code. *Proc Natl Acad Sci U S A*, **72**, 1909-1912.
12. Jurgens, C., Strom, A., Wegener, D., Hettwer, S., Wilmanns, M. and Sterner, R. (2000) Directed evolution of a (beta alpha)₈-barrel enzyme to catalyze related reactions in two different metabolic pathways. *Proc Natl Acad Sci U S A*, **97**, 9925-9930.
13. Parker, J. and Holtz, G. (1984) Control of basal-level codon misreading in Escherichia coli. *Biochem Biophys Res Commun*, **121**, 487-492.
14. Laughrea, M., Latulippe, J., Filion, A.M. and Boulet, L. (1987) Mistranslation in twelve Escherichia coli ribosomal proteins. Cysteine misincorporation at neutral amino acid residues other than tryptophan. *Eur J Biochem*, **169**, 59-64.
15. Crick, F.H. (1968) The origin of the genetic code. *J Mol Biol*, **38**, 367-379.
16. Bock, A., Forchhammer, K., Heider, J., Leinfelder, W., Sawers, G., Veprek, B. and Zinoni, F.

- (1991) Selenocysteine: the 21st amino acid. *Mol Microbiol*, **5**, 515-520.
17. Srinivasan, G., James, C.M. and Krzycki, J.A. (2002) Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science*, **296**, 1459-1462.
 18. Wang, L., Brock, A., Herberich, B. and Schultz, P.G. (2001) Expanding the genetic code of *Escherichia coli*. *Science*, **292**, 498-500.
 19. Kiga, D., Sakamoto, K., Kodama, K., Kigawa, T., Matsuda, T., Yabuki, T., Shirouzu, M., Harada, Y., Nakayama, H., Takio, K. *et al.* (2002) An engineered *Escherichia coli* tyrosyl-tRNA synthetase for site-specific incorporation of an unnatural amino acid into proteins in eukaryotic translation and its application in a wheat germ cell-free system. *Proc Natl Acad Sci U S A*, **99**, 9715-9720.
 20. Liu, C.C. and Schultz, P.G. (2010) Adding new chemistries to the genetic code. *Annu Rev Biochem*, **79**, 413-444.
 21. Woese, C.R., Olsen, G.J., Ibba, M. and Soll, D. (2000) Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol Mol Biol Rev*, **64**, 202-236.
 22. Ikehara, K., Omori, Y., Arai, R. and Hirose, A. (2002) A novel theory on the origin of the genetic code: a GNC-SNS hypothesis. *J Mol Evol*, **54**, 530-538.
 23. Trifonov, E.N. (2004) The triplet code from first principles. *J Biomol Struct Dyn*, **22**, 1-11.
 24. Akanuma, S., Nakajima, Y., Yokobori, S., Kimura, M., Nemoto, N., Mase, T., Miyazono, K., Tanokura, M. and Yamagishi, A. (2013) Experimental evidence for the thermophilicity of ancestral life. *Proc Natl Acad Sci U S A*, **110**, 11067-11072.
 25. Ibba, M. and Soll, D. (2000) Aminoacyl-tRNA synthesis. *Annu Rev Biochem*, **69**, 617-650.
 26. Oshikane, H., Sheppard, K., Fukai, S., Nakamura, Y., Ishitani, R., Numata, T., Sherrer, R.L., Feng, L., Schmitt, E., Panvert, M. *et al.* (2006) Structural basis of RNA-dependent recruitment of glutamine to the genetic code. *Science*, **312**, 1950-1954.
 27. Ito, T. and Yokoyama, S. (2010) Two enzymes bound to one transfer RNA assume alternative conformations for consecutive reactions. *Nature*, **467**, 612-616.
 28. Blaise, M., Bailly, M., Frechin, M., Behrens, M.A., Fischer, F., Oliveira, C.L., Becker, H.D., Pedersen, J.S., Thirup, S. and Kern, D. (2010) Crystal structure of a transfer-ribonucleoprotein particle that promotes asparagine formation. *EMBO J*, **29**, 3118-3129.
 29. Sauerwald, A., Zhu, W., Major, T.A., Roy, H., Palioura, S., Jahn, D., Whitman, W.B., Yates, J.R., 3rd, Ibba, M. and Soll, D. (2005) RNA-dependent cysteine biosynthesis in archaea. *Science*, **307**, 1969-1972.
 30. O'Donoghue, P., Sethi, A., Woese, C.R. and Luthey-Schulten, Z.A. (2005) The evolutionary history of Cys-tRNA^{Cys} formation. *Proc Natl Acad Sci U S A*, **102**, 19003-19008.
 31. Goltermann, L., Larsen, M.S., Banerjee, R., Joerger, A.C., Ibba, M. and Bentin, T. (2010) Protein evolution via amino acid and codon elimination. *PLoS One*, **5**, e10104.

32. Suzuki, T., Arai, S., Takeuchi, M., Sakurai, C., Ebana, H., Higashi, T., Hashimoto, H., Hatsuzawa, K. and Wada, I. (2012) Development of cysteine-free fluorescent proteins for the oxidative environment. *PLoS One*, **7**, e37551.
33. Wang, J.D., Herman, C., Tipton, K.A., Gross, C.A. and Weissman, J.S. (2002) Directed evolution of substrate-optimized GroEL/S chaperonins. *Cell*, **111**, 1027-1039.
34. Wang, Q., Buckle, A.M., Foster, N.W., Johnson, C.M. and Fersht, A.R. (1999) Design of highly stable functional GroEL minichaperones. *Protein Sci*, **8**, 2186-2193.
35. Seki, E., Matsuda, N., Yokoyama, S. and Kigawa, T. (2008) Cell-free protein synthesis system from *Escherichia coli* cells cultured at decreased temperatures improves productivity by decreasing DNA template degradation. *Anal Biochem*, **377**, 156-161.
36. Steiner, T., Hess, P., Bae, J.H., Wiltschi, B., Moroder, L. and Budisa, N. (2008) Synthetic biology of proteins: tuning GFPs folding and stability with fluoroproline. *PLoS One*, **3**, e1680.
37. Budisa, N., Pal, P.P., Alefelder, S., Birle, P., Krywcun, T., Rubini, M., Wenger, W., Bae, J.H. and Steiner, T. (2004) Probing the role of tryptophans in *Aequorea victoria* green fluorescent proteins with an expanded genetic code. *Biol Chem*, **385**, 191-202.
38. Ayukawa, S., Kobayashi, A., Nakashima, Y., Takagi, H., Hamada, S., Uchiyama, M., Yugi, K., Murata, S., Sakakibara, Y., Hagiya, M. *et al.* (2010) Construction of a genetic AND gate under a new standard for assembly of genetic parts. *BMC Genomics*, **11 Suppl 4**, S16.
39. Cubitt, A.B., Woollenweber, L.A. and Heim, R. (1999) Understanding structure-function relationships in the *Aequorea victoria* green fluorescent protein. *Methods Cell Biol*, **58**, 19-30.
40. Pedelacq, J.D., Cabantous, S., Tran, T., Terwilliger, T.C. and Waldo, G.S. (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol*, **24**, 79-88.
41. Fisher, A.C. and DeLisa, M.P. (2008) Laboratory evolution of fast-folding green fluorescent protein using secretory pathway quality control. *PLoS One*, **3**, e2351.
42. Bell, A.F., Stoner-Ma, D., Wachter, R.M. and Tonge, P.J. (2003) Light-driven decarboxylation of wild-type green fluorescent protein. *J Am Chem Soc*, **125**, 6919-6926.
43. Weinreich, D.M., Watson, R.A. and Chao, L. (2005) Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution*, **59**, 1165-1174.
44. Tracewell, C.A. and Arnold, F.H. (2009) Directed enzyme evolution: climbing fitness peaks one amino acid at a time. *Curr Opin Chem Biol*, **13**, 3-9.
45. Enoki, S., Saeki, K., Maki, K. and Kuwajima, K. (2004) Acid denaturation and refolding of green fluorescent protein. *Biochemistry*, **43**, 14238-14248.
46. Maki, K., Cheng, H., Dolgikh, D.A. and Roder, H. (2007) Folding kinetics of staphylococcal nuclease studied by tryptophan engineering and rapid mixing methods. *J Mol Biol*, **368**, 244-255.
47. Kotsuka, T., Akanuma, S., Tomuro, M., Yamagishi, A. and Oshima, T. (1996) Further stabilization of 3-isopropylmalate dehydrogenase of an extreme thermophile, *Thermus*

- thermophilus, by a suppressor mutation method. *J Bacteriol*, **178**, 723-727.
48. Zhou, H.X., Hoess, R.H. and DeGrado, W.F. (1996) In vitro evolution of thermodynamically stable turns. *Nat Struct Biol*, **3**, 446-451.
 49. Lehner, B. (2011) Molecular mechanisms of epistasis within and between genes. *Trends Genet*, **27**, 323-331.
 50. de Visser, J.A., Cooper, T.F. and Elena, S.F. (2011) The causes of epistasis. *Proc Biol Sci*, **278**, 3617-3624.
 51. Blight, S.K., Larue, R.C., Mahapatra, A., Longstaff, D.G., Chang, E., Zhao, G., Kang, P.T., Green-Church, K.B., Chan, M.K. and Krzycki, J.A. (2004) Direct charging of tRNA(CUA) with pyrrolysine in vitro and in vivo. *Nature*, **431**, 333-335.
 52. Kim, D.M., Kigawa, T., Choi, C.Y. and Yokoyama, S. (1996) A highly efficient cell-free protein synthesis system from Escherichia coli. *Eur J Biochem*, **239**, 881-886.
 53. Kim, B. and Little, J.W. (1993) LexA and lambda CI repressors as enzymes: specific cleavage in an intermolecular reaction. *Cell*, **73**, 1165-1173.
 54. Kigawa, T., Yabuki, T., Matsuda, N., Matsuda, T., Nakajima, R., Tanaka, A. and Yokoyama, S. (2004) Preparation of Escherichia coli cell extract for highly productive cell-free protein expression. *J Struct Funct Genomics*, **5**, 63-68.
 55. Kigawa, T., Inoue, M., Aoki, M., Matsuda, T., Yabuki, T., Seki, E., Harada, T., Watanabe, S. and Yokoyama, S. (2007) In Spirin, A. S. and Swartz, J. R. (eds.), *CELL-FREE PROTEIN SYNTHESIS : METHODS AND PROTOCOLS*. Wiley-VCH, Weinheim, Germany, pp. 99-109.
 56. Simpson, R.J., Neuberger, M.R. and Liu, T.Y. (1976) Complete amino acid analysis of proteins from a single hydrolysate. *J Biol Chem*, **251**, 1936-1940.
 57. Masuda, A. and Dohmae, N. (2010) Automated Protein Hydrolysis Delivering Sample to a Solid Acid Catalyst for Amino Acid Analysis. *Analytical Chemistry*, **82**, 8939-8945.
 58. Matsuda, T., Kigawa, T., Koshihara, S., Inoue, M., Aoki, M., Yamasaki, K., Seki, M., Shinozaki, K. and Yokoyama, S. (2006) Cell-free synthesis of zinc-binding proteins. *J Struct Funct Genomics*, **7**, 93-100.
 59. Kigawa, T., Matsuda, T., Yabuki, T. and Yokoyama, S. (2007) In Spirin, A. S. and Swartz, J. R. (eds.), *CELL-FREE PROTEIN SYNTHESIS : METHODS AND PROTOCOLS*. Wiley-VCH, Weinheim, Germany, pp. 83-97.
 60. Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology*, **276**, 307-326.
 61. Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.*, **30**, 1022-1025.
 62. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, **60**, 2126-2132.

63. Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K. and Terwilliger, T.C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr*, **58**, 1948-1954.
64. Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*, **53**, 240-255.
65. Hou, Y.M. and Schimmel, P. (1988) A simple structural feature is a major determinant of the identity of a transfer RNA. *Nature*, **333**, 140-145.
66. Francklyn, C. and Schimmel, P. (1989) Aminoacylation of RNA minihelices with alanine. *Nature*, **337**, 478-481.
67. Guo, M., Chong, Y.E., Beebe, K., Shapiro, R., Yang, X.L. and Schimmel, P. (2009) The C-Ala domain brings together editing and aminoacylation functions on one tRNA. *Science*, **325**, 744-747.
68. Sundharadas, G., Katze, J.R., Söll, D., Konigsberg, W. and Lengyel, P. (1968) On the recognition of serine transfer RNA's specific for unrelated codons by the same seryl-transfer RNA synthetase. *Proc Natl Acad Sci U S A*, **61**, 693-700.
69. Sampson, J.R. and Saks, M.E. (1993) Contributions of discrete tRNA(Ser) domains to aminoacylation by E.coli seryl-tRNA synthetase: a kinetic analysis using model RNA substrates. *Nucleic Acids Res*, **21**, 4467-4475.
70. Cusack, S., Yaremchuk, A. and Tukalo, M. (1996) The crystal structure of the ternary complex of T.thermophilus seryl-tRNA synthetase with tRNA(Ser) and a seryl-adenylate analogue reveals a conformational switch in the active site. *EMBO J*, **15**, 2834-2842.
71. Slilaty, S.N. and Little, J.W. (1987) Lysine-156 and serine-119 are required for LexA repressor cleavage: a possible mechanism. *Proc Natl Acad Sci U S A*, **84**, 3987-3991.
72. Yokoyama, J., Matsuda, T., Koshiba, S. and Kigawa, T. (2010) An economical method for producing stable-isotope labeled proteins by the E. coli cell-free system. *J Biomol NMR*, **48**, 193-201.
73. Yoshizawa, S. and Bock, A. (2009) The many levels of control on bacterial selenoprotein synthesis. *Biochim Biophys Acta*, **1790**, 1404-1414.
74. Kawaguchi, Y., Honda, H., Taniguchi-Morimura, J. and Iwasaki, S. (1989) The codon CUG is read as serine in an asporogenic yeast *Candida cylindracea*. *Nature*, **341**, 164-166.
75. Yokogawa, T., Suzuki, T., Ueda, T., Mori, M., Ohama, T., Kuchino, Y., Yoshinari, S., Motoki, I., Nishikawa, K., Osawa, S. *et al.* (1992) Serine tRNA complementary to the nonuniversal serine codon CUG in *Candida cylindracea*: evolutionary implications. *Proc Natl Acad Sci U S A*, **89**, 7408-7411.

76. Massey, S.E., Moura, G., Beltrao, P., Almeida, R., Garey, J.R., Tuite, M.F. and Santos, M.A. (2003) Comparative evolutionary genomics unveils the molecular mechanism of reassignment of the CTG codon in *Candida* spp. *Genome Res*, **13**, 544-557.
77. Dorazi, R., Lingutla, J.J. and Humayun, M.Z. (2002) Expression of mutant alanine tRNAs increases spontaneous mutagenesis in *Escherichia coli*. *Mol Microbiol*, **44**, 131-141.
78. Geslain, R., Cubells, L., Bori-Sanz, T., Alvarez-Medina, R., Rossell, D., Marti, E. and Ribas de Pouplana, L. (2010) Chimeric tRNAs as tools to induce proteome damage and identify components of stress responses. *Nucleic Acids Res*, **38**, e30.
79. Aslan, F.M., Yu, Y., Mohr, S.C. and Cantor, C.R. (2005) Engineered single-chain dimeric streptavidins with an unexpected strong preference for biotin-4-fluorescein. *Proc Natl Acad Sci U S A*, **102**, 8507-8512.
80. Moroz, O.V., Moroz, Y.S., Wu, Y., Olsen, A.B., Cheng, H., Mack, K.L., McLaughlin, J.M., Raymond, E.A., Zhezherya, K., Roder, H. *et al.* (2013) A single mutation in a regulatory protein produces evolvable allosterically regulated catalyst of nonnatural reaction. *Angew Chem Int Ed Engl*, **52**, 6246-6249.
81. Han, M.J. and Lee, S.Y. (2006) The *Escherichia coli* proteome: past, present, and future prospects. *Microbiol Mol Biol Rev*, **70**, 362-439.
82. Pezo, V., Louis, D., Guerineau, V., Le Caer, J.P., Gaillon, L., Mutzel, R. and Marliere, P. (2013) A metabolic prototype for eliminating tryptophan from the genetic code. *Sci Rep*, **3**, 1359.
83. Muller, M.M., Allison, J.R., Hongdilokkul, N., Gaillon, L., Kast, P., van Gunsteren, W.F., Marliere, P. and Hilvert, D. (2013) Directed evolution of a model primordial enzyme provides insights into the development of the genetic code. *PLoS Genet*, **9**, e1003187.
84. Fournier, G.P. and Gogarten, J.P. (2007) Signature of a primitive genetic code in ancient protein lineages. *J Mol Evol*, **65**, 425-436.
85. Alvarez-Carreno, C., Becerra, A. and Lazcano, A. (2013) Norvaline and Norleucine May Have Been More Abundant Protein Components during Early Stages of Cell Evolution. *Orig Life Evol Biosph.*
86. Pullman, B. and Pullman, A. (1958) Electron-Donor and -Acceptor Properties of Biologically Important Purines, Pyrimidines, Pteridines, Flavins, and Aromatic Amino Acids. *Proc Natl Acad Sci U S A*, **44**, 1197-1202.
87. Beene, D.L., Brandt, G.S., Zhong, W., Zacharias, N.M., Lester, H.A. and Dougherty, D.A. (2002) Cation- π interactions in ligand recognition by serotonergic (5-HT_{3A}) and nicotinic acetylcholine receptors: the anomalous binding properties of nicotine. *Biochemistry*, **41**, 10262-10269.
88. Silva, R.M., Paredes, J.A., Moura, G.R., Manadas, B., Lima-Costa, T., Rocha, R., Miranda, I., Gomes, A.C., Koerkamp, M.J., Perrot, M. *et al.* (2007) Critical roles for a genetic code alteration

- in the evolution of the genus *Candida*. *EMBO J*, **26**, 4555-4565.
89. Wiltschi, B. and Budisa, N. (2007) Natural history and experimental evolution of the genetic code. *Appl Microbiol Biotechnol*, **74**, 739-753.
 90. Chambaud, I., Heilig, R., Ferris, S., Barbe, V., Samson, D., Galisson, F., Moszer, I., Dybvig, K., Wroblewski, H., Viari, A. *et al.* (2001) The complete genome sequence of the murine respiratory pathogen *Mycoplasma pulmonis*. *Nucleic Acids Res*, **29**, 2145-2153.
 91. Koonin, E.V., Makarova, K.S. and Aravind, L. (2001) Horizontal gene transfer in prokaryotes: quantification and classification. *Annu Rev Microbiol*, **55**, 709-742.
 92. Yamamoto, Y., Tsutsumi, Y., Yoshioka, Y., Nishibata, T., Kobayashi, K., Okamoto, T., Mukai, Y., Shimizu, T., Nakagawa, S., Nagata, S. *et al.* (2003) Site-specific PEGylation of a lysine-deficient TNF-alpha with full bioactivity. *Nat Biotechnol*, **21**, 546-552.
 93. Isaacs, F.J., Carr, P.A., Wang, H.H., Lajoie, M.J., Sterling, B., Kraal, L., Tolonen, A.C., Gianoulis, T.A., Goodman, D.B., Reppas, N.B. *et al.* (2011) Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science*, **333**, 348-353.
 94. Waldo, G.S., Standish, B.M., Berendzen, J. and Terwilliger, T.C. (1999) Rapid protein-folding assay using green fluorescent protein. *Nat Biotechnol*, **17**, 691-695.

発表論文

Akio Kawahara-Kobayashi, Akiko Masuda, Yuhei Arais, Yoko Sakai, Atsushi Kohda, Masahiko Uchiyama, Shun Asami, Takayoshi Matsuda, Ryuichiro Ishitani, Naoshi Dohmae, Shigeyuki Yokoyama, Takanori Kigawa, Osamu Nureki, and Daisuke Kiga, “Simplification of the genetic code: restricted diversity of genetically encoded amino acids”, *Nucleic Acids Research*, (2012) 40 (20):10576-10584. Featured article.

Akio Kawahara-Kobayashi, Mitsuhiro Hitotsuyanagi, Kazuaki Amikura and Daisuke Kiga, “Experimental Evolution of a Green Fluorescent Protein Composed of 19 Unique Amino Acids without Tryptophan”, *Origins of Life and Evolution of Biospheres*, **Accepted**

謝辞

本研究は、木賀大介先生のご指導と励ましのもとに行われました。先生には、多大なご苦勞をおかけしました。先生に直接ご指導いただいた、実験の進め方、論文執筆や申請書作成にあたっての論理の組み立て方は、私にとっての大きな財産です。また、数多くの先生方との共同研究の機会を与えていただいたことも、貴重な学びの場となりました。心から感謝いたします。

横山茂之先生、濡木理先生、石谷隆一郎先生、堂前直先生、木川隆則先生には、大変お忙しい中、私の研究のためにお力添えをいただき、また惜しみないご助言をいただき、感謝しております。山村雅幸先生、中村清彦先生、瀧ノ上正浩先生、小宮健先生には貴重なご指導とご助言をいただきました。ありがとうございます。

荒磯裕平博士、益田晶子博士、関英子博士、松田貴意博士には、実験面でも論文に関する議論でも、親身にご指導いただきました。皆様の深い知識と洞察に触れ、博士号を取ることへの目標を改めて明確にしました。

研究を進めるにあたって、東京工業大学山村・木賀研究室のメンバーには大変お世話になりました。特に内山正彦博士には、多くの実験技術、作法をご指導いただきました。実験をする姿だけでなく、熱い議論をする姿、またコーヒーを飲みながら1対1で語る姿と、近いところで多くの時間を過ごせたことは、いまでも私の記憶に強く残っています。犬飼直人さん、浅見俊さん、國府田敦さん、後藤玲央さん、藤原久敬さん、網蔵和晃博士、藤本智大さん、一柳光弘さん、酒井洋子さんには、ともに単純化遺伝暗号表に関わる仲間として、実験上の相談から研究の意義、展望についても様々な議論をさせていただきました。石松愛博士、鮎川翔太郎博士、関根亮二博士、畑敬士さん、森谷孟史さん、西田暁史さんは、研究テーマこそ違いましたが、同じ合成生物学ラボのメンバーとして私に刺激を与え続け、また日々励ましの言葉もいただき、感謝しております。吉田美智子さん、西文恵さん、平野明子さんには秘書業務でお世話になりました。ありがとうございます。上記の皆様を含む、現在および過去に研究室で一緒にさせていただきました皆さまには、公私に渡り、本当にお世話になりました。ありがとうございます。

最後に、私の進む道に制限を課すことなく見守り、支えてくれた両親、さらに、いつも私の健康に気を配り、自由な私に理解を示し応援してくれた妻・澄枝に厚く感謝を申し上げます。