Synthetic life

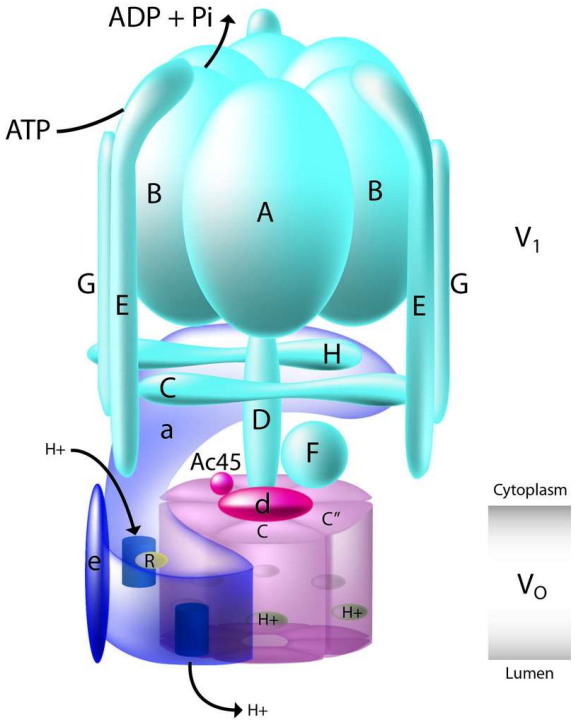
(continuation of „The molecular origins of life" SoSe 2021)

NaturalNews.com

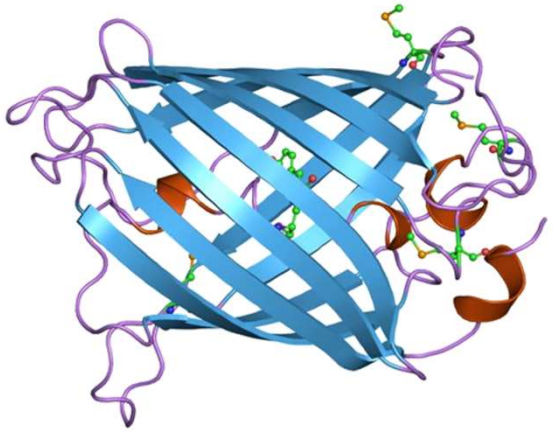
WiSe 2021/22

Zbigniew Pianowski

CHAPTER 2



N.Ochotny

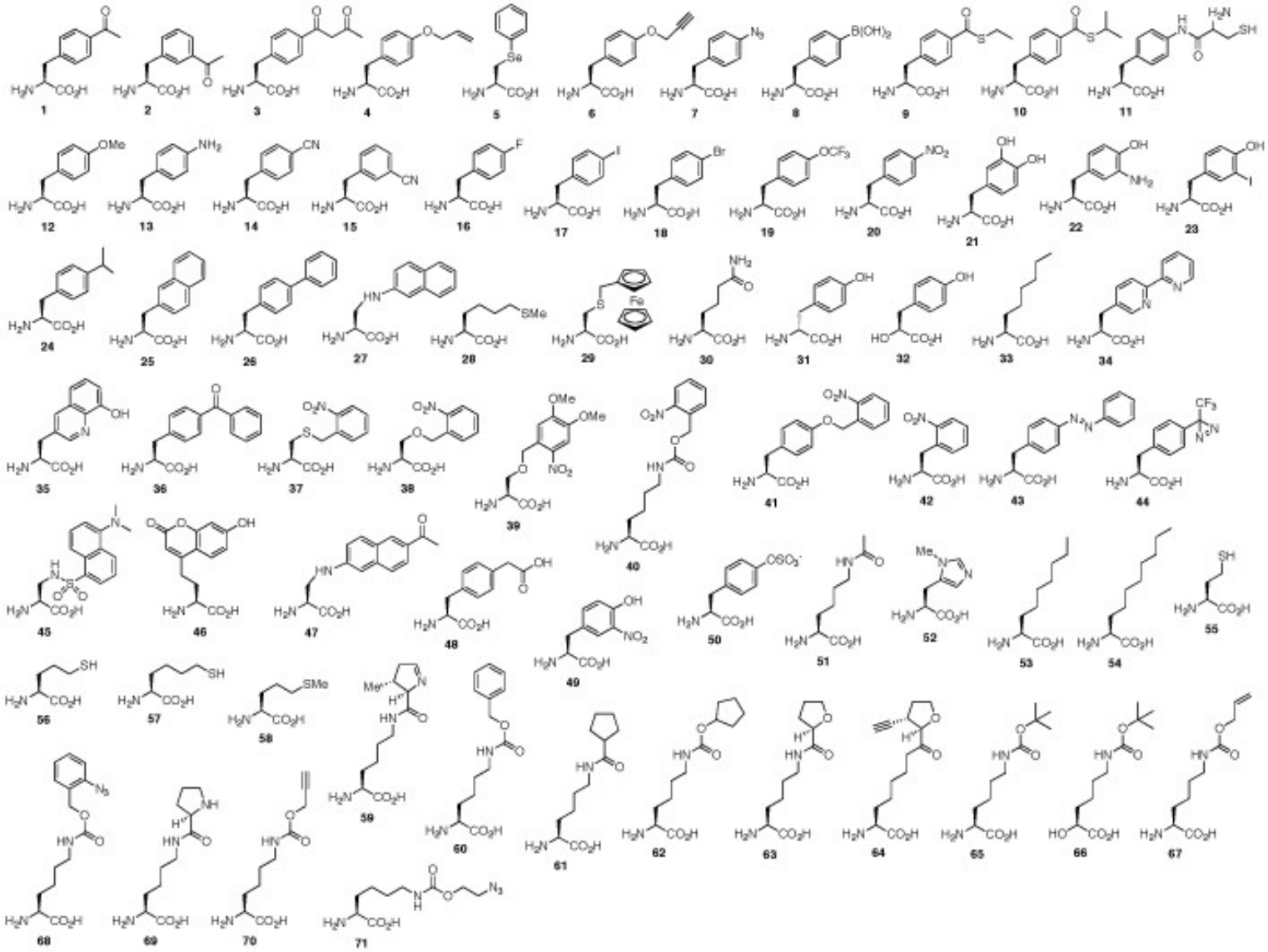


J. Swaminathan

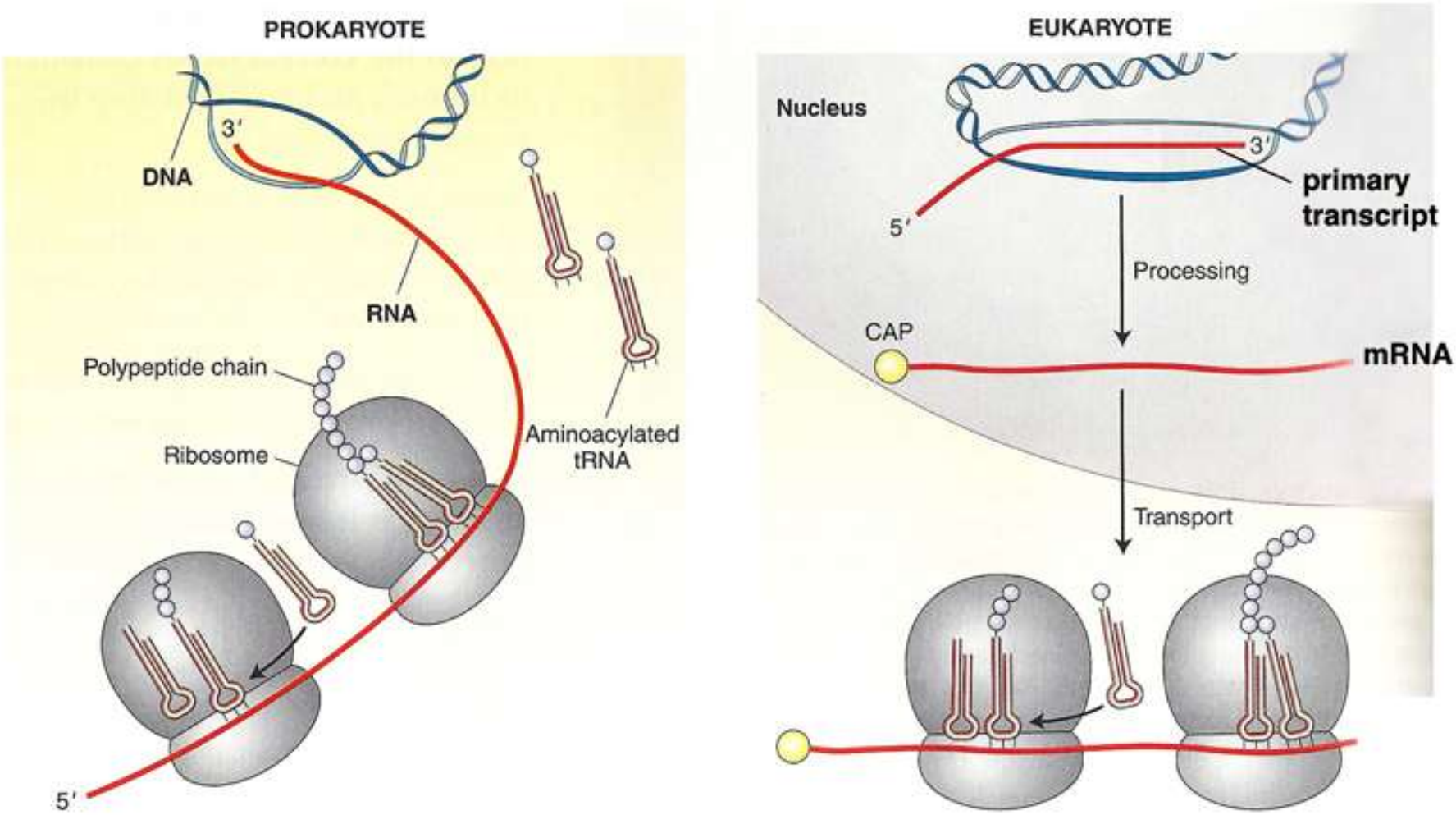
PROTEINS

Genetic encoding of non-standard aminoacids

The expanding genetic code



Translation: RNA → proteins

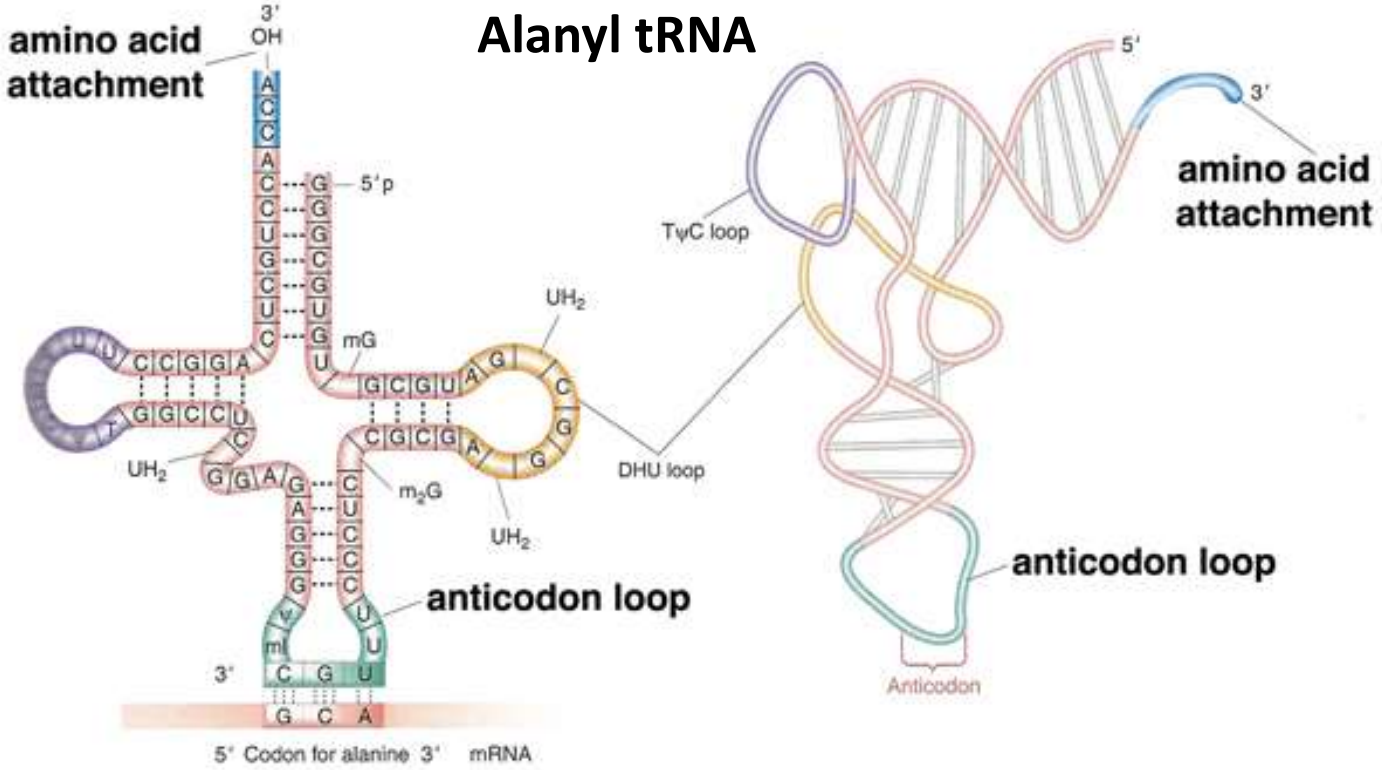


Translation: RNA → proteins

A transfer RNA has a cloverleaf structure with regions of base pairing.
A tRNA has the structure shown here both as a flat cloverleaf and in its folded form.

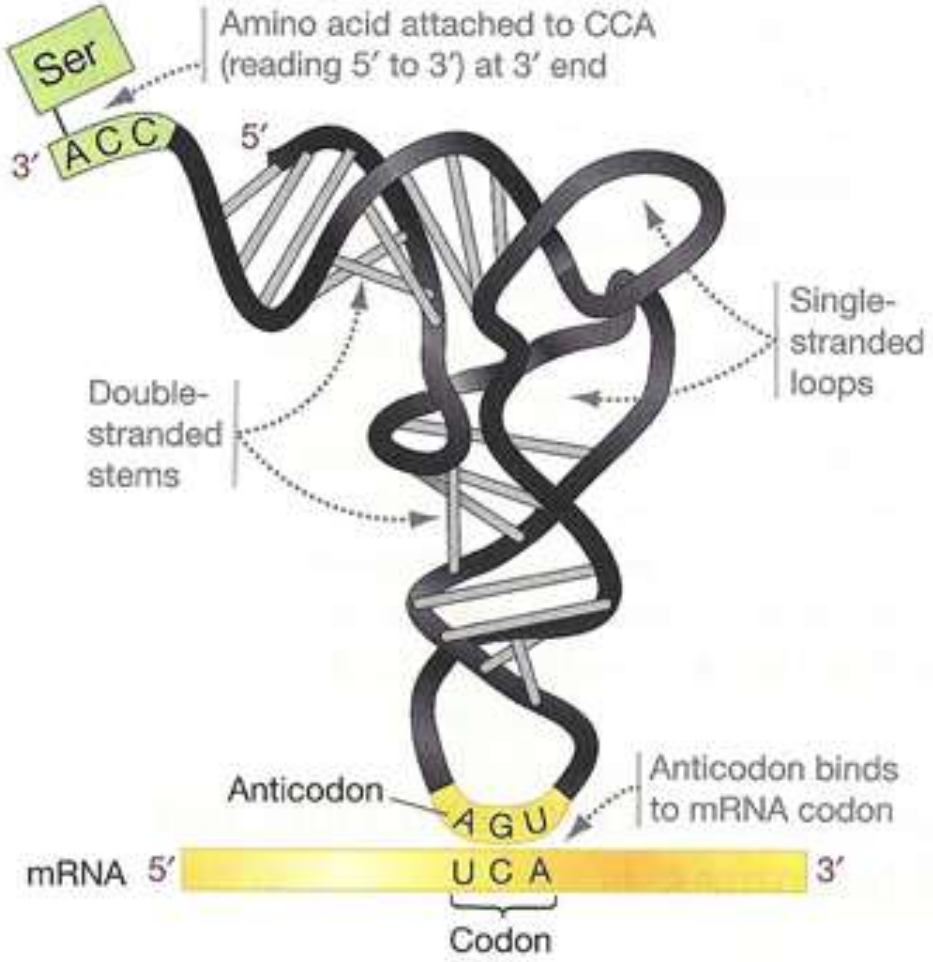
Two important parts of a tRNA:

- the anticodon, which participates in base pairing with a codon in the mRNA
- the site of amino acid attachment at the 3' end of the tRNA



Translation: RNA → proteins

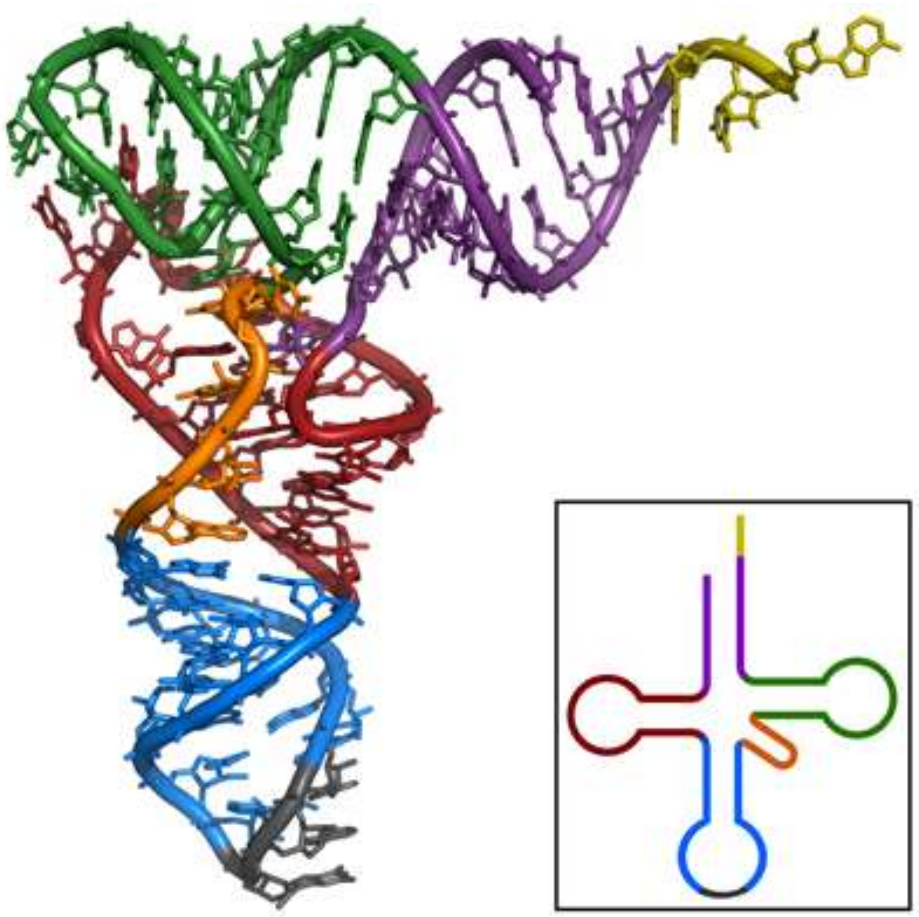
Charged serine tRNA



This shows a "charged" serine tRNA, covalently attached to the amino acid serine at its 3' end, with the anticodon paired to a serine codon

Translation: RNA → proteins

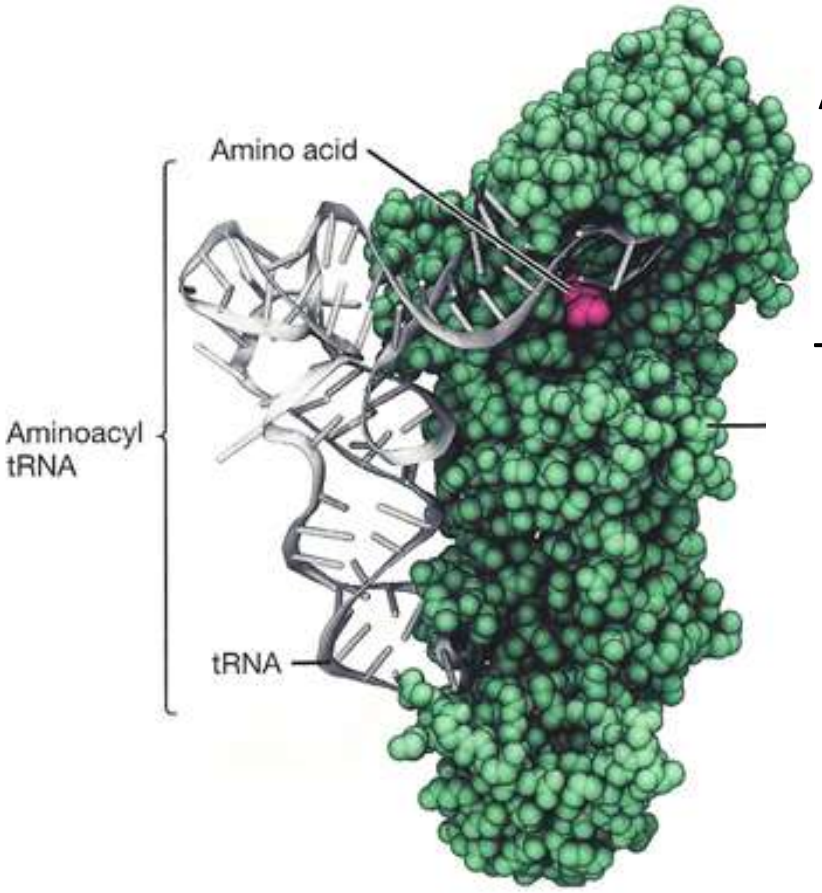
tRNA in 3D



This is a better representation of the 3D structure of a tRNA. The model is color-coded to the flat cloverleaf representation in the lower right

Translation: RNA → proteins

Aminoacyl tRNA synthethase



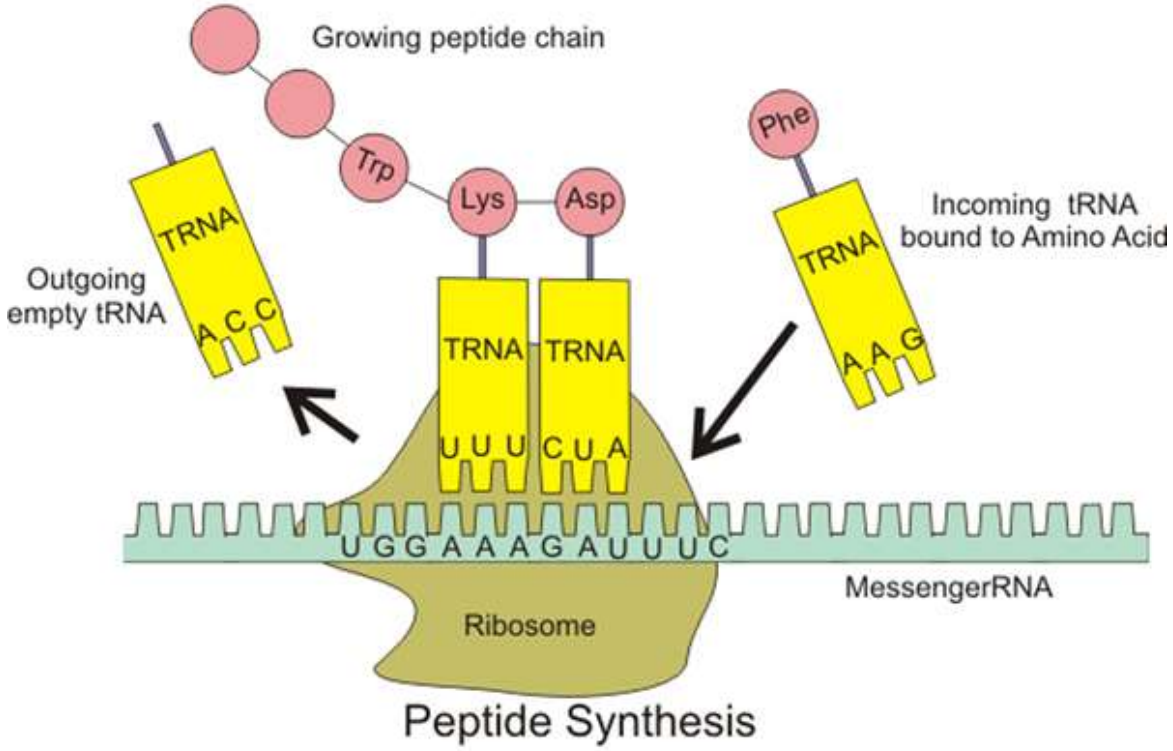
A special set of enzymes "charges" tRNAs, attaching the correct amino acid to particular tRNAs.

A charged tRNA is called an aminoacyl tRNA, so the charging enzymes are more properly called aminoacyl tRNA synthetases.

There is only one aminoacyl tRNA synthetase for each amino acid, even though there can be multiple tRNAs for that amino acid. Each aminoacyl tRNA synthetase is able to recognize all of the tRNAs that need to be charged with the one amino acid that is their specialty.

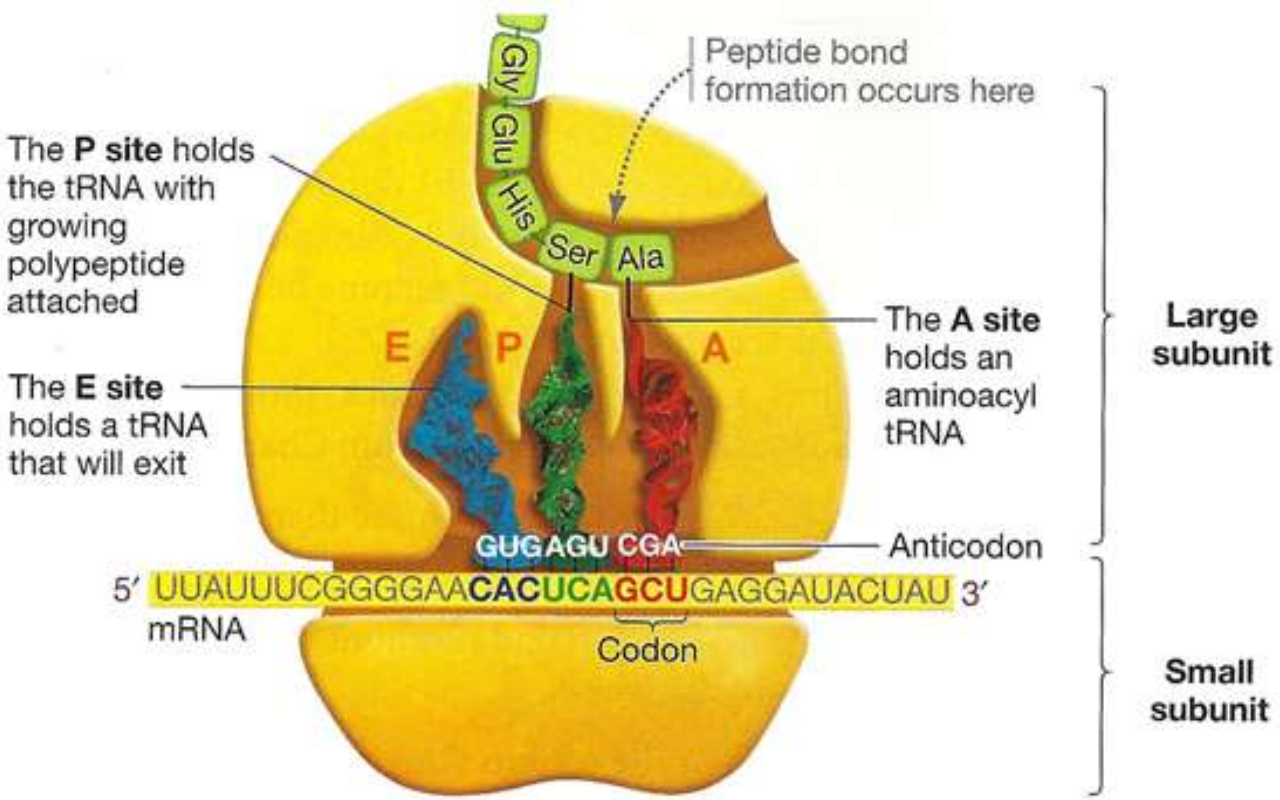
Amino acids are attached to the hydroxyl (-OH) group at the 3' end of the tRNA through their carboxyl (-COOH) group

Translation: RNA → proteins



Proteins are synthesized by ribosomes that read the sequence of mRNA and write it as protein. Translation is accomplished with the help of charged tRNAs that allow individual codons to specify the next amino acid added to the growing polypeptide. The mRNA is read from the 5' end to the 3' end, with the protein being synthesized from the amino terminus to the carboxy terminus

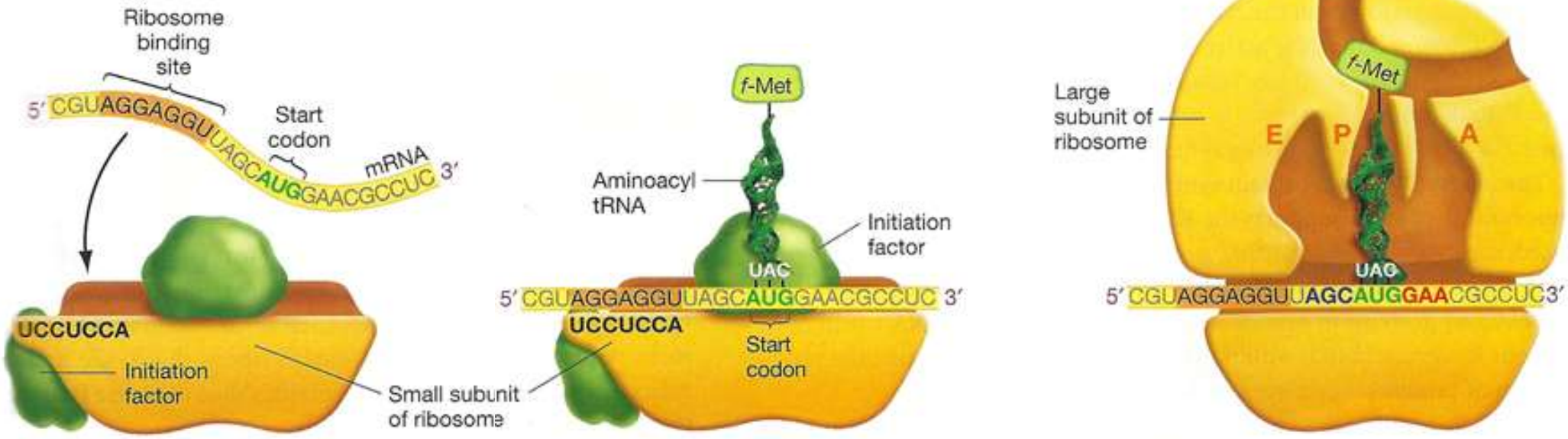
Translation: RNA → proteins



Three sites are associated with tRNAs: the A (aminocyl) site, that accepts a new aminoacyl tRNA; the P (polypeptide) site, that holds a tRNA with the growing polypeptide chain; and the E (exit) site that holds an uncharged tRNA ready to exit the ribosome

Translation: RNA → proteins

Translation initiation



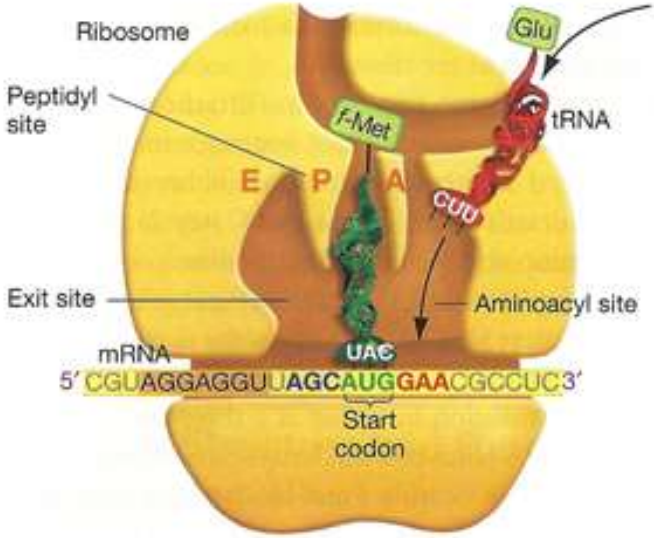
1. mRNA binds to small subunit. Ribosome binding site sequence binds to a complementary sequence in an RNA molecule in the small subunit of the ribosome, with the help of protein initiation factors.

2. Initiator aminoacyl tRNA binds to start codon.

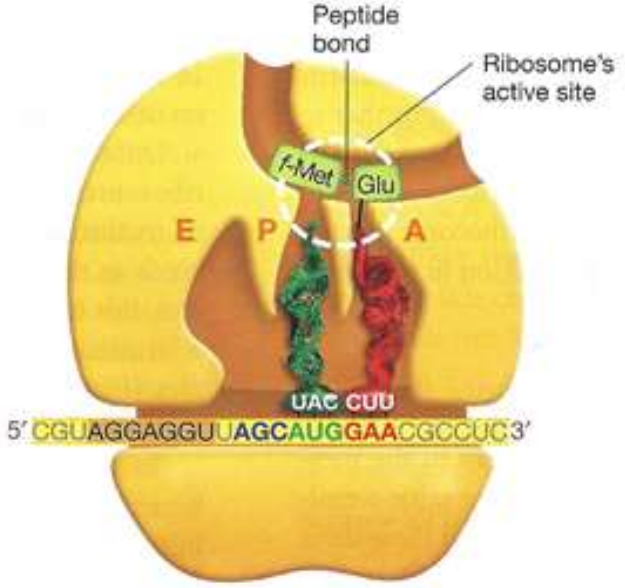
3. Large subunit of ribosome binds, completing ribosome assembly. Translation begins.

Translation: RNA → proteins

Translation elongation



1. Incoming aminoacyl tRNA
New tRNA moves into A site, where its anticodon base pairs with the mRNA codon.

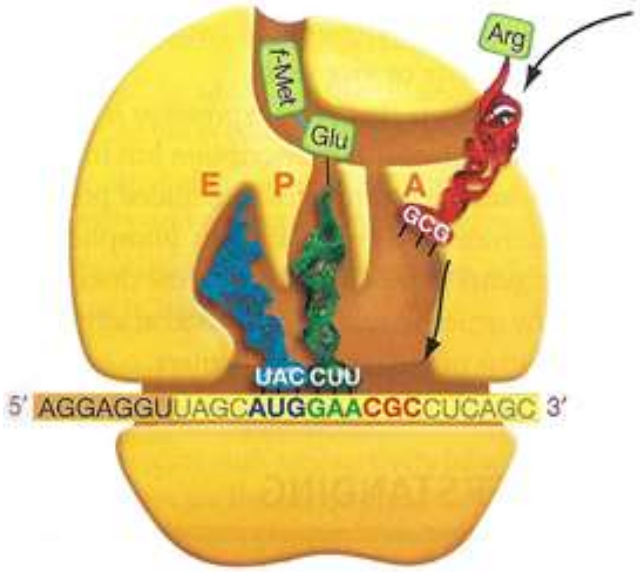


2. Peptide bond formation
The amino acid attached to the tRNA in the P site is transferred to the tRNA in the A site.

3. Translocation
mRNA is ratcheted through the ribosome by elongation factors (not shown). The tRNA attached to the polypeptide chain moves into the P site. The A site is empty.

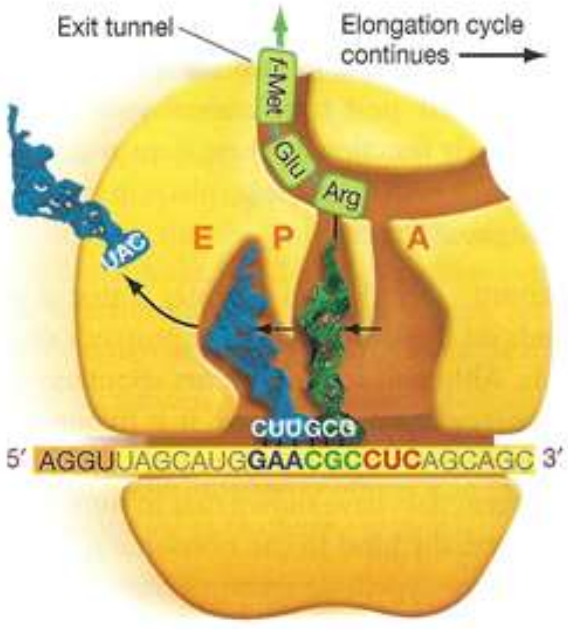
Translation: RNA → proteins

Translation elongation



4. Incoming aminoacyl tRNA
New tRNA moves into A site, where its anticodon base pairs with the mRNA codon.

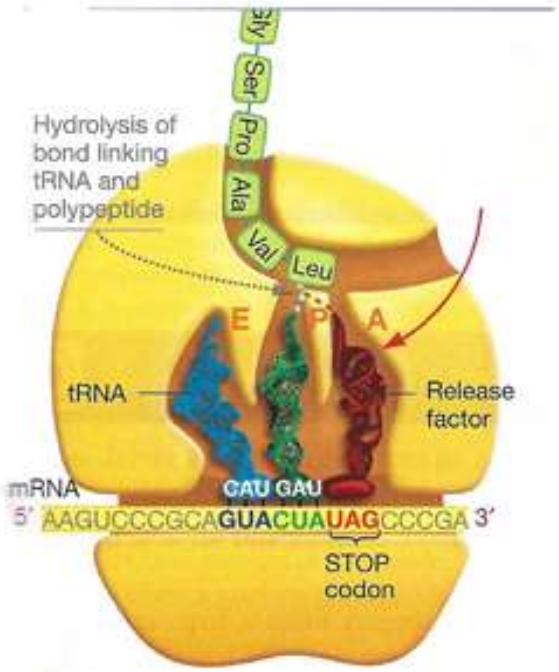
5. Peptide bond formation
The polypeptide chain attached to the tRNA in the P site is transferred to the aminoacyl tRNA in the A site.



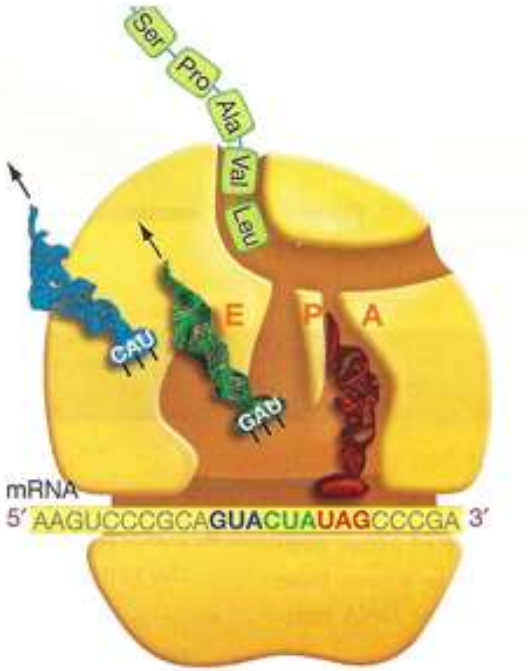
6. Translocation
mRNA is ratcheted through the ribosome again. The tRNA attached to polypeptide chain moves into P site. Empty tRNA from P site moves to E site, where tRNA is ejected. The A site is empty again.

Translation: RNA → proteins

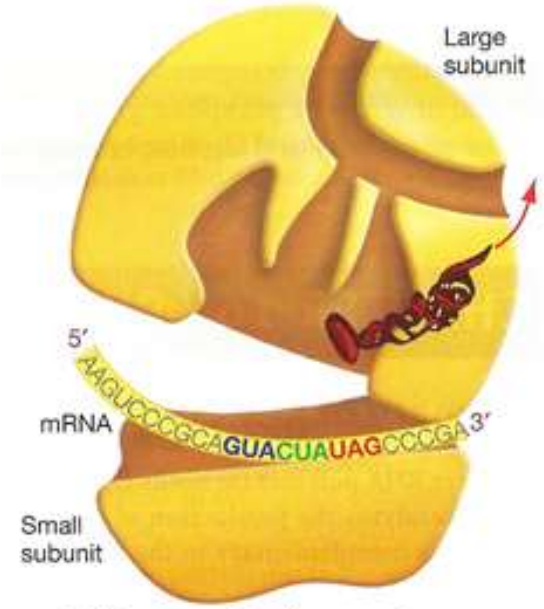
Translation termination



1. Release factor binds to stop codon. When translocation exposes a stop codon, a release factor fills the A site. The release factor breaks the bond linking the tRNA in the P site to the polypeptide chain.



2. Polypeptide is released. The hydrolysis reaction frees the polypeptide, which is released from the ribosome. The empty tRNAs are released either along with the polypeptide or...



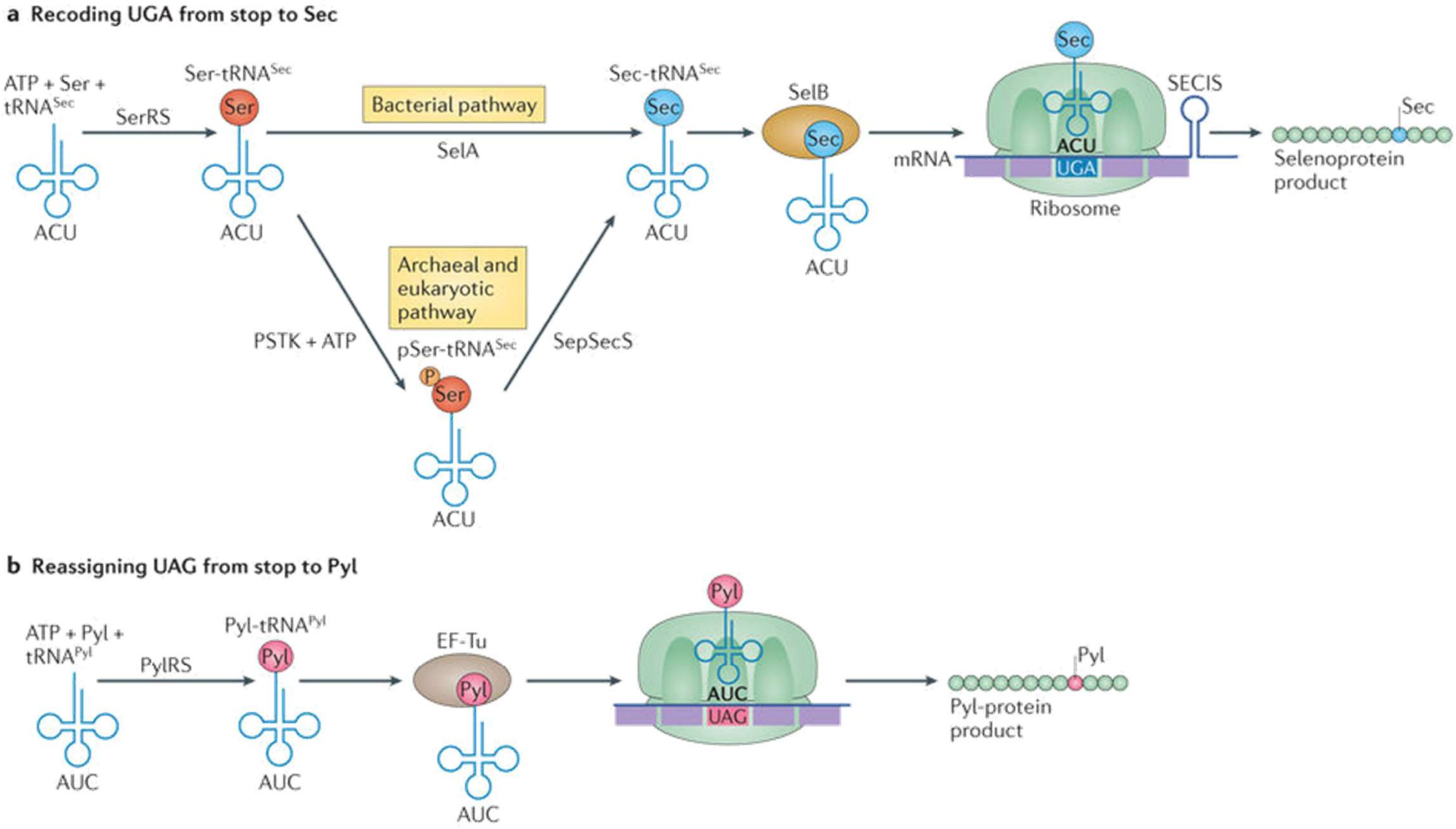
3. Ribosome subunits separate. ...when the ribosome separates from the mRNA, and the two ribosomal subunits dissociate. The subunits are ready to attach to the start codon of another message and start translation anew.

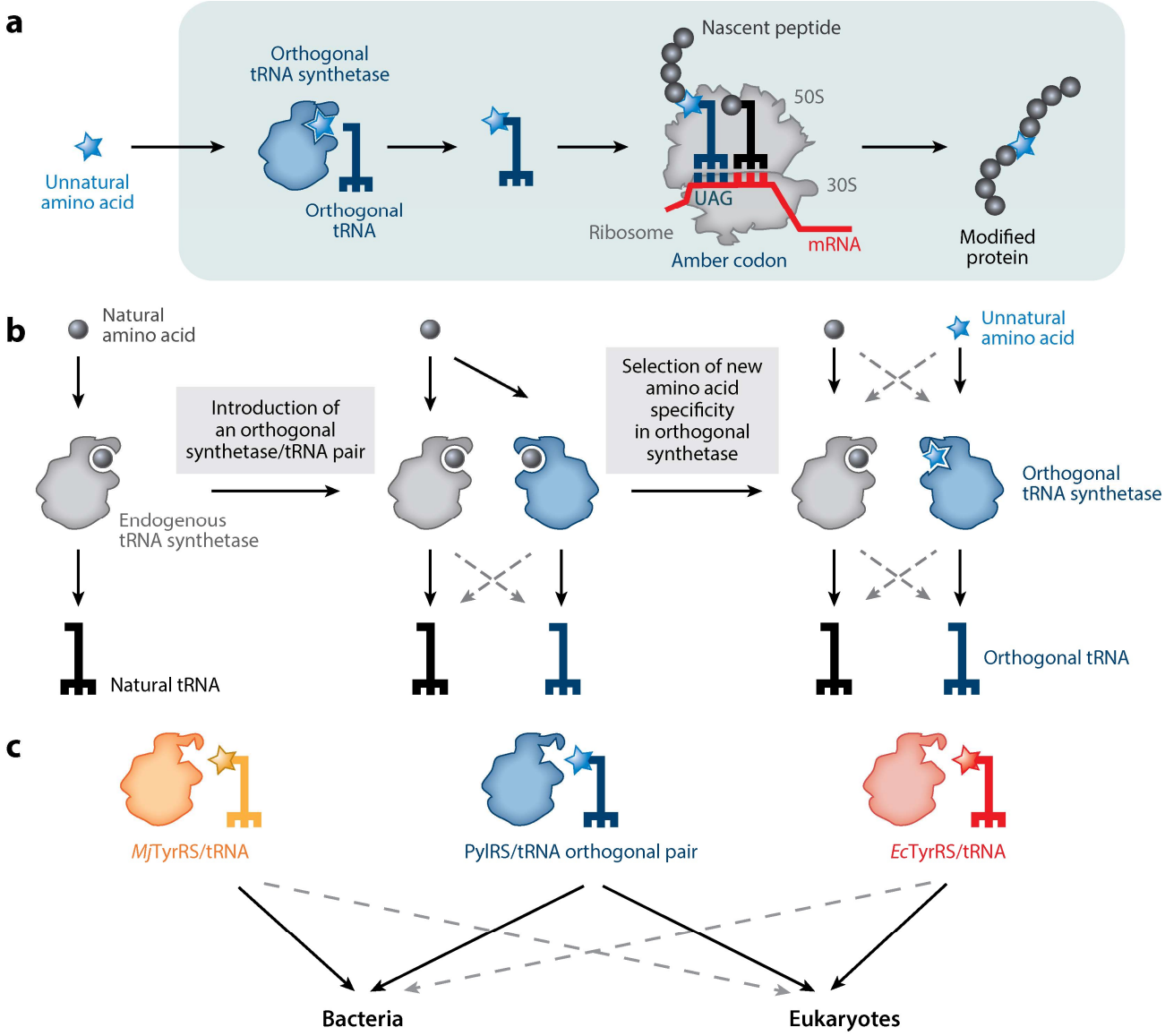
Translation: RNA → proteins – the genetic code

nonpolar polar basic acidic (stop codon)

Standard genetic code

1st base	2nd base								3rd base
	U		C		A		G		
U	UUU	(Phe/F) Phenylalanine	UCU	(Ser/S) Serine	UAU	(Tyr/Y) Tyrosine	UGU	(Cys/C) Cysteine	U
	UUC		UCC		UAC		UGC		C
	UUA	(Leu/L) Leucine	UCA		UAA	Stop (*Ochre*)	UGA	Stop (*Opal*)	A
	UUG		UCG		UAG	Stop (*Amber*)	UGG	(Trp/W) Tryptophan	G
C	CUU		CCU	(Pro/P) Proline	CAU	(His/H) Histidine	CGU	(Arg/R) Arginine	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	(Gln/Q) Glutamine	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	(Ile/I) Isoleucine	ACU	(Thr/T) Threonine	AAU	(Asn/N) Asparagine	AGU	(Ser/S) Serine	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	(Lys/K) Lysine	AGA	(Arg/R) Arginine	A
	AUG[A]	(Met/M) Methionine	ACG		AAG		AGG		G
G	GUU	(Val/V) Valine	GCU	(Ala/A) Alanine	GAU	(Asp/D) Aspartic acid	GGU	(Gly/G) Glycine	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	(Glu/E) Glutamic acid	GGA		A
	GUG		GCG		GAG		GGG		G


a Recoding UGA from stop to Sec
ATP + Ser + tRNA^Sec
SerRS
Ser-tRNA^Sec
Ser
ACU
ACU
Bacterial pathway
SelA
Sec-tRNA^Sec
Sec
ACU
SelB
Sec
ACU
mRNA
Sec
SECIS
ACU
UGA
Ribosome
Sec
Selenoprotein product
Archaeal and eukaryotic pathway
PSTK + ATP
pSer-tRNA^Sec
P
Ser
ACU
SepSecS
b Reassigning UAG from stop to Pyl
ATP + Pyl + tRNA^Pyl
PylRS
Pyl-tRNA^Pyl
Pyl
AUC
AUC
EF-Tu
Pyl
AUC
Pyl
AUC
UAG
Pyl
Pyl-protein product

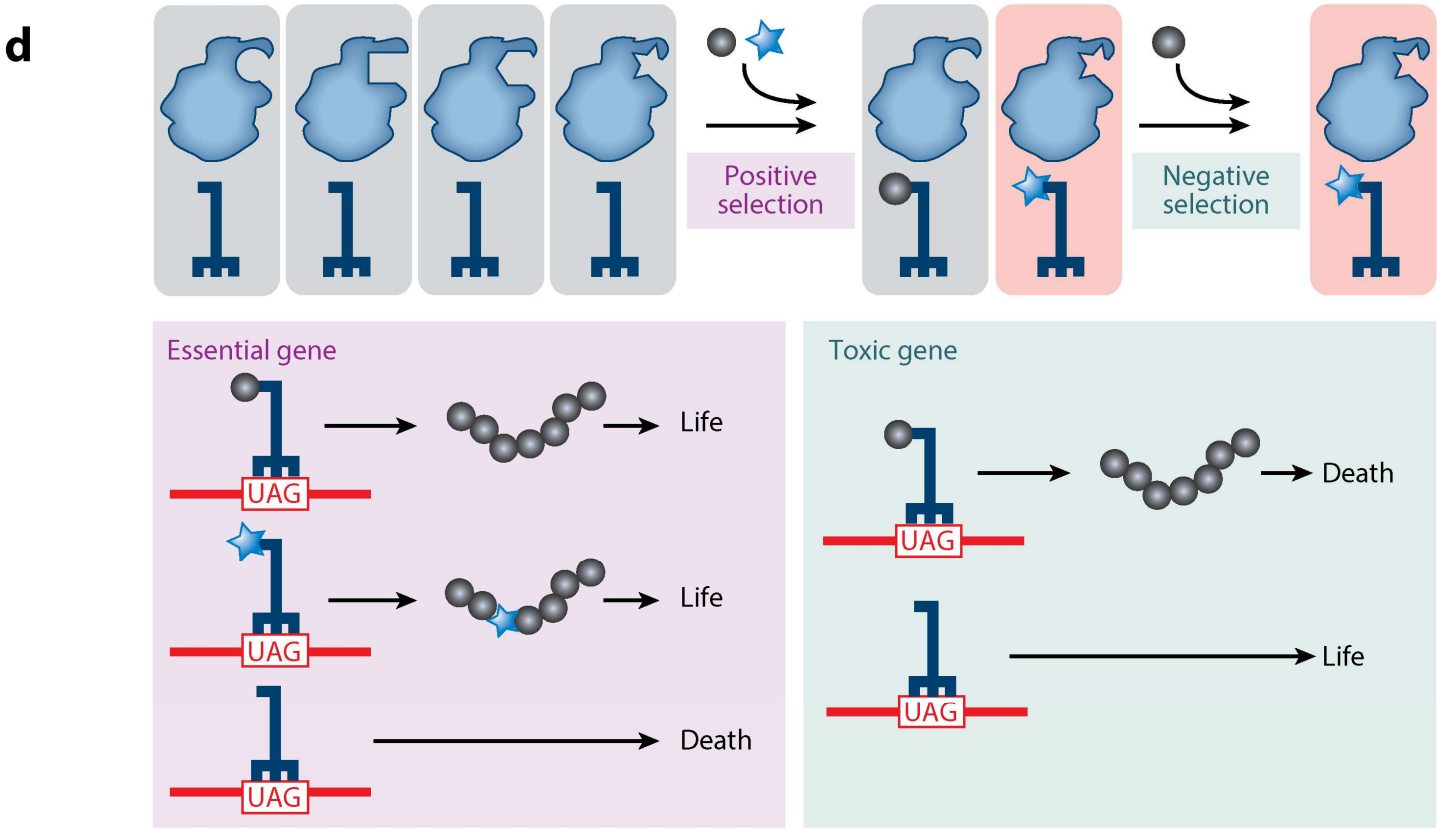


(a) Genetic code expansion enables the site-specific incorporation of an unnatural amino acid into a protein via cellular translation.

(b) The process of discovering orthogonal aminoacyl-tRNA (transfer RNA) synthetases for unnatural amino acids.

(c) Orthogonality of synthetase/tRNA pairs in different hosts. The solid lines indicate that a pair is orthogonal in a host

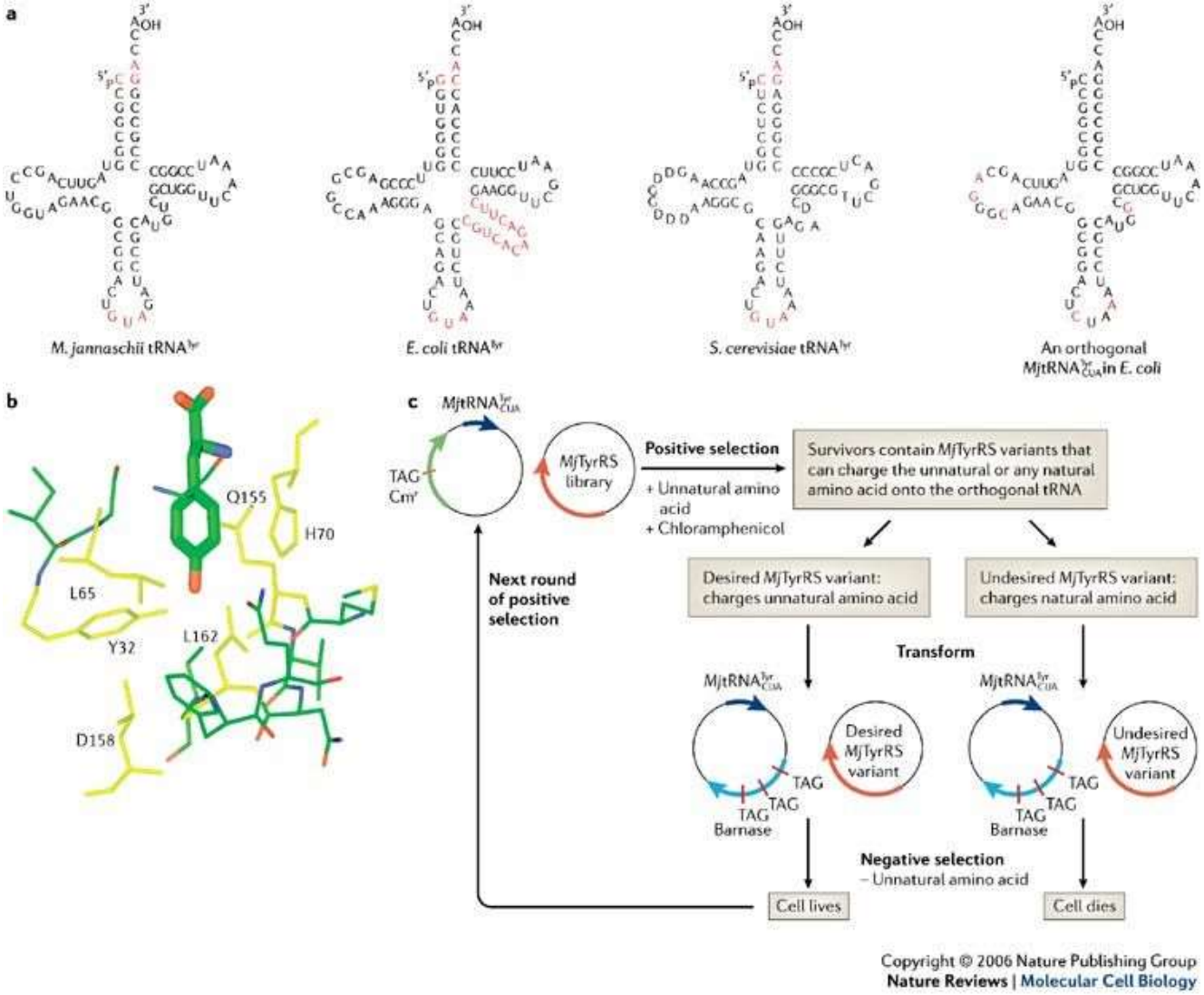
(*d*) Sequential positive and negative selections enable the discovery of synthetase/tRNA pairs that direct the incorporation of unnatural amino acids.



Chin JW. 2014.
Annu. Rev. Biochem. 83:379–408

***Ec*TyrRS,** *Escherichia coli* tyrosyl-tRNA synthetase;
***Mj*TyrRS,** *Methanococcus janaschii* tyrosyl-tRNA synthetase;
mRNA, messenger RNA; **PylRS,** pyrrolysyl-tRNA synthetase.

The expanding genetic code

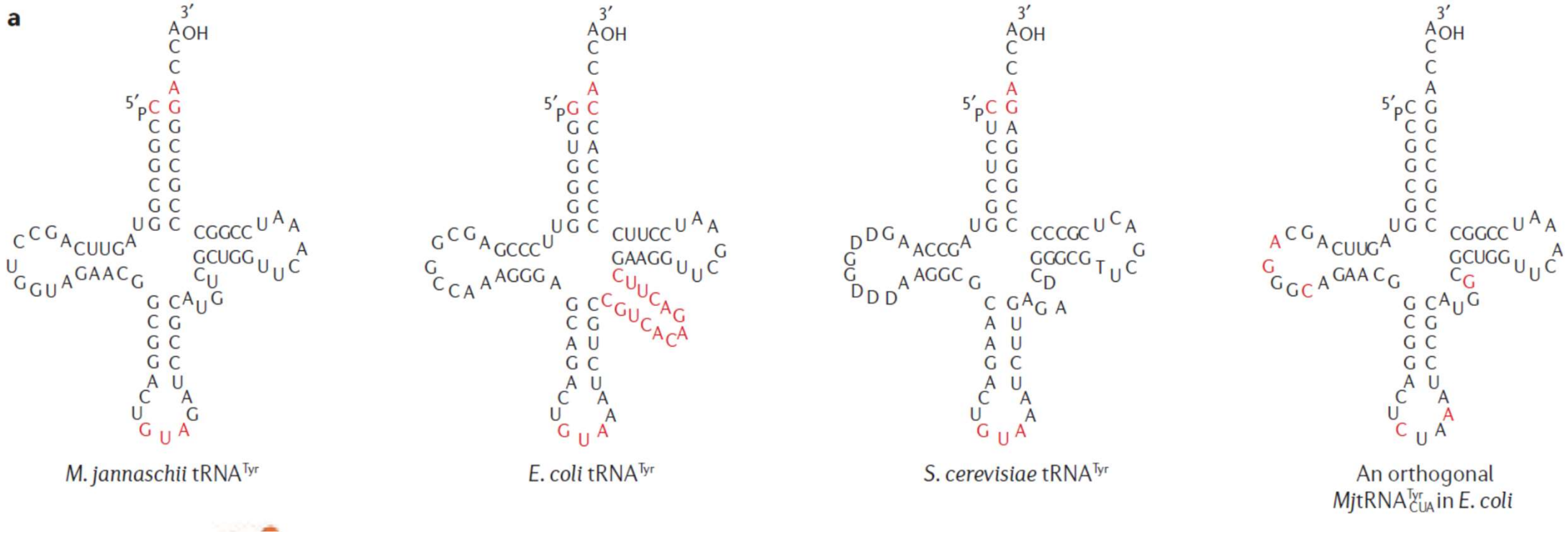


Copyright © 2006 Nature Publishing Group
Nature Reviews | Molecular Cell Biology

J. Xie, P. G. Schultz *Nature Rev. Mol. Cell Biol.* **2006**, *7*, 775-782.

The expanding genetic code

The development of an orthogonal amber suppressor *Methanococcus jannaschii* tyrosyl–transfer-RNA (*Mj* $tRNA_{CUA}^{Tyr}$) in *Escherichia coli* and the modification of the amino-acid specificity of its cognate *M. jannaschii* tyrosyl–tRNA synthetase (*Mj*TyrRS)

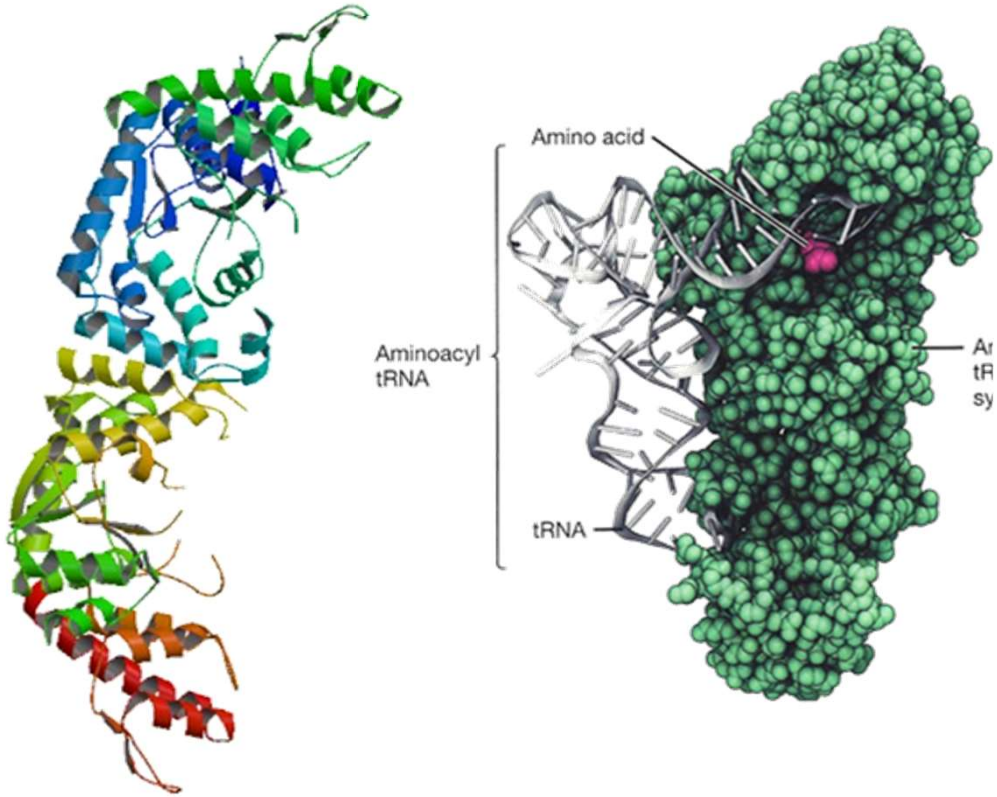


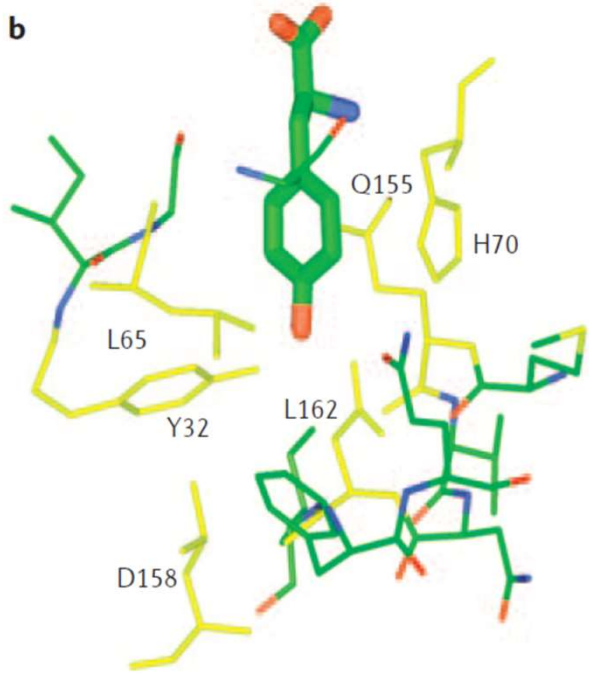
The $tRNA^{Tyr}$ molecules from *M. jannaschii*, *E. coli* and *Saccharomyces cerevisiae* (with the key identity elements that are recognized by the cognate synthetases highlighted in red), and the orthogonal amber suppressor *Mj* $tRNA_{CUA}^{Tyr}$ in *E. coli* (with the modified nucleotides highlighted in red). The D nucleotide is dihydrouridine.

J. Xie, P. G. Schultz *Nature Rev. Mol. Cell Biol.* **2006**, *7*, 775-782.

The expanding genetic code

The development of an orthogonal amber suppressor *Methanococcus jannaschii* tyrosyl–transfer-RNA (*Mj* $tRNA_{CUA}^{Tyr}$) in *Escherichia coli* and the modification of the amino-acid specificity of its cognate *M. jannaschii* tyrosyl–tRNA synthetase (*Mj*TyrRS)





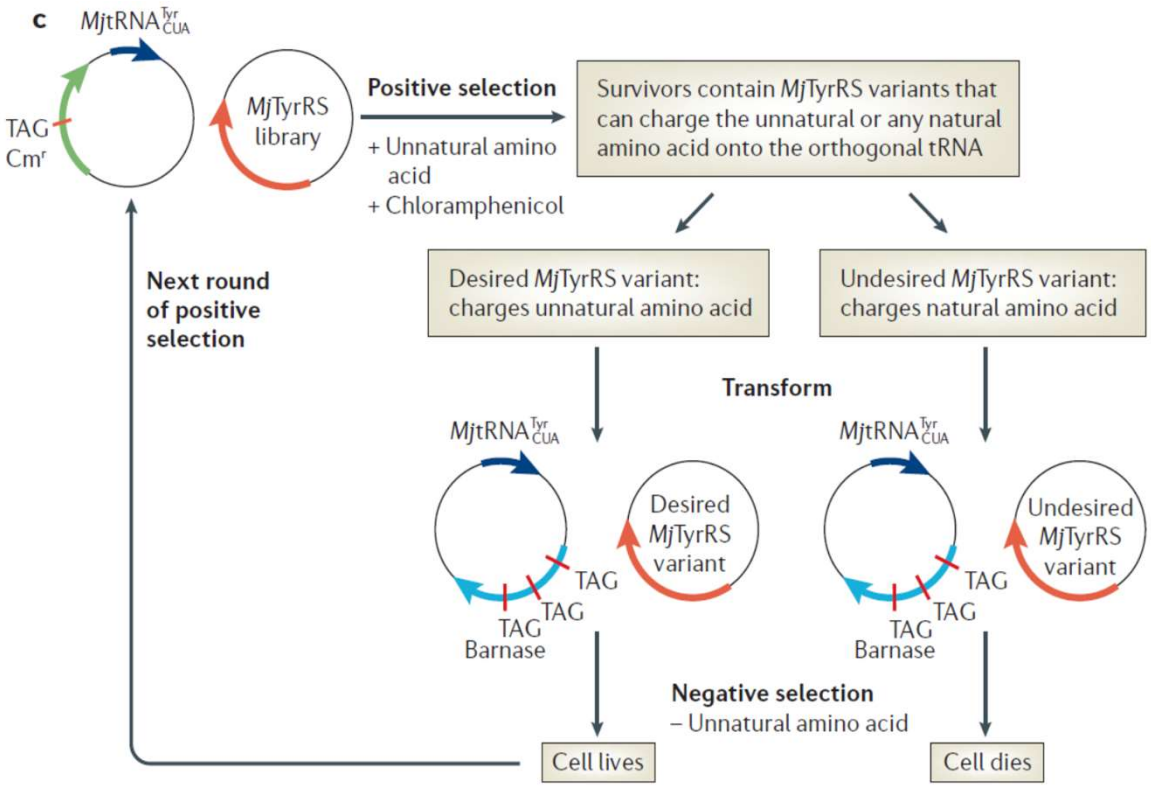
P. G. Schultz *et al.* (2005) *Protein Sci.* **2005**, *14,* 1340-1349

A library of MjTyrRS mutants was generated by randomly mutating 6 residues (shown in yellow) in the Tyr-binding site to all 20 amino acids. Tyr is shown in its binding site using a thicker stick representation.

J. Xie, P. G. Schultz *Nature Rev. Mol. Cell Biol.* **2006**, *7*, 775-782.

The expanding genetic code

A general positive and negative selection scheme for the development of synthetase variants that are specific for an unnatural amino acid in E. coli.

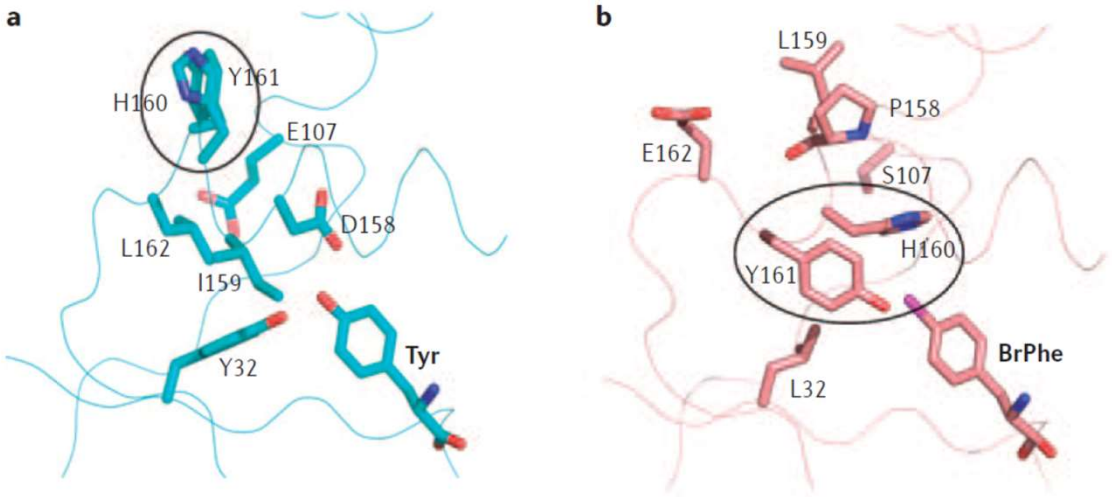


Following the generation of a large library (~10^9 mutants) of, in this case, MjTyrRS active-site mutants, positive and negative selections were carried out. The positive selection was based on resistance to chloramphenicol, which was conferred in the presence of MjTyrRS and the unnatural amino acid (or any natural amino acid that the MjTyrRS could charge onto the orthogonal tRNA) by the suppression of an amber mutation (TAG) at a permissive site in the chloramphenicol acetyltransferase gene (labelled Cm^r). The negative selection used the toxic barnase gene with amber mutations at permissive sites and was carried out in the absence of the unnatural amino acid. Only MjTyrRS variants that could acylate the orthogonal $tRNA^{Tyr}_{CUA}$ with the unnatural amino acid and not with the endogenous amino acids could survive both selections.

J. Xie, P. G. Schultz *Nature Rev. Mol. Cell Biol.* **2006**, *7*, 775-782.

The expanding genetic code

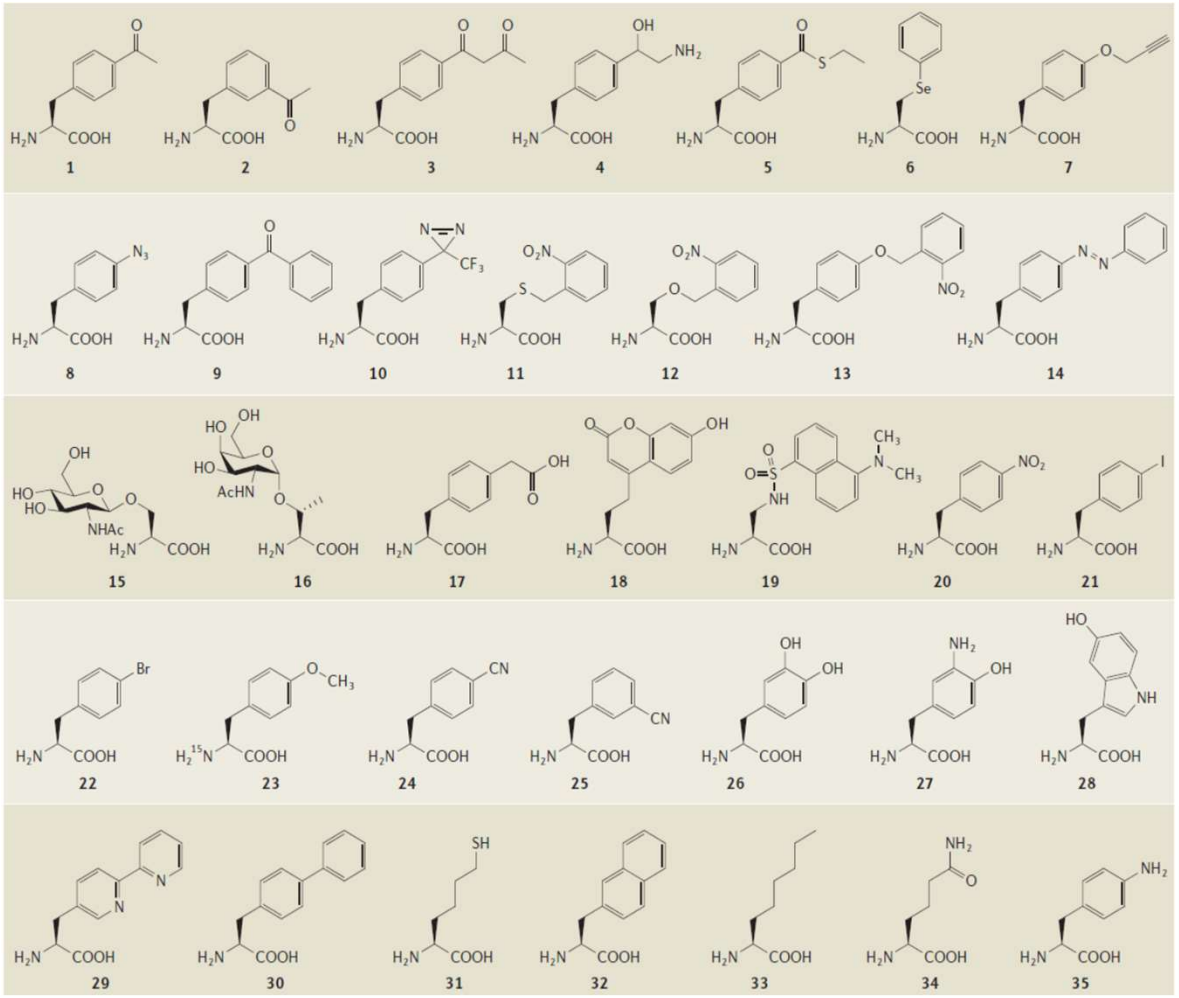
The structures of the wild-type and a mutant Methanococcus jannaschii tyrosyl–tRNA synthetase bound to their cognate amino acids.



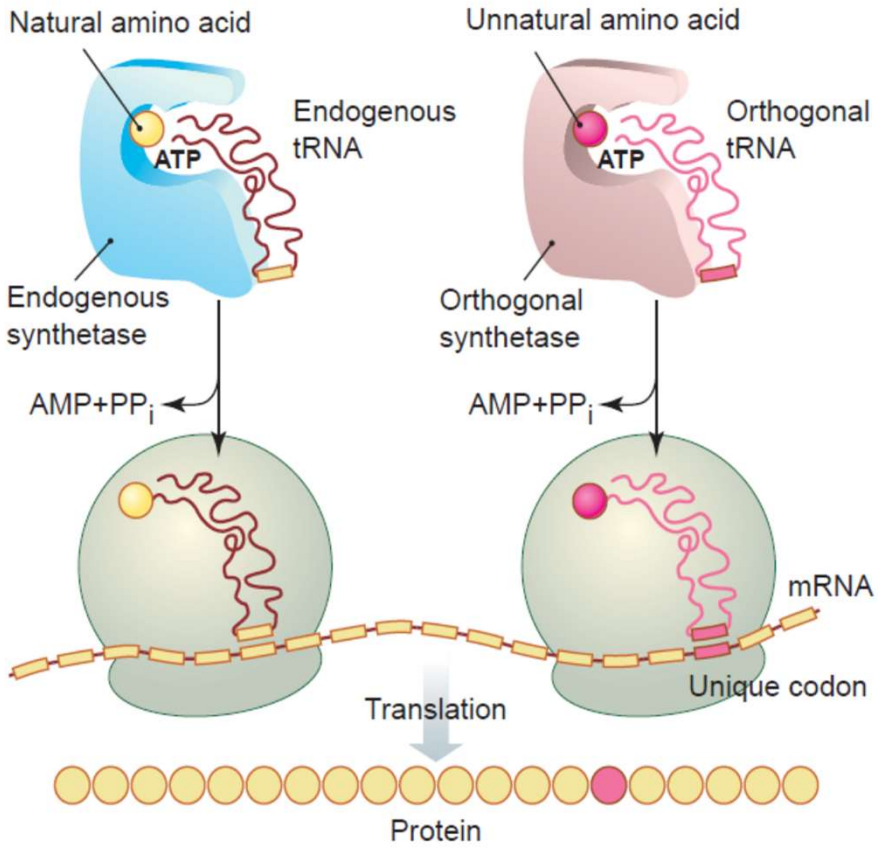
a The active site of wild-type *Methanococcus jannaschii* tyrosyl–transfer-RNA synthetase (MjTyrRS) bound to Tyr.
b The active site of a mutant MjTyrRS that binds to p-bromophenylalanine (labelled BrPhe in the figure). The active site of the mutant contains the mutations Y32L, E107S, D158P, I159L and L162E. The active-site D158P and Y32L mutations remove two hydrogen bonds to the hydroxyl group of the Tyr side chain, which disfavours the binding of the natural substrate. The D158P mutation results in the termination of helix α8 and produces significant translational and rotational movements of several active-site residues. These effects, in conjunction with the effects of the Y32L mutation, lead to an expanded hydrophobic active-site cavity that favours the binding of p-bromophenylalanine. Black frames highlight the different positioning of H160 and Y161 in these structures.

J. Xie, P. G. Schultz *Nature Rev. Mol. Cell Biol.* **2006**, *7*, 775-782.

The expanding genetic code

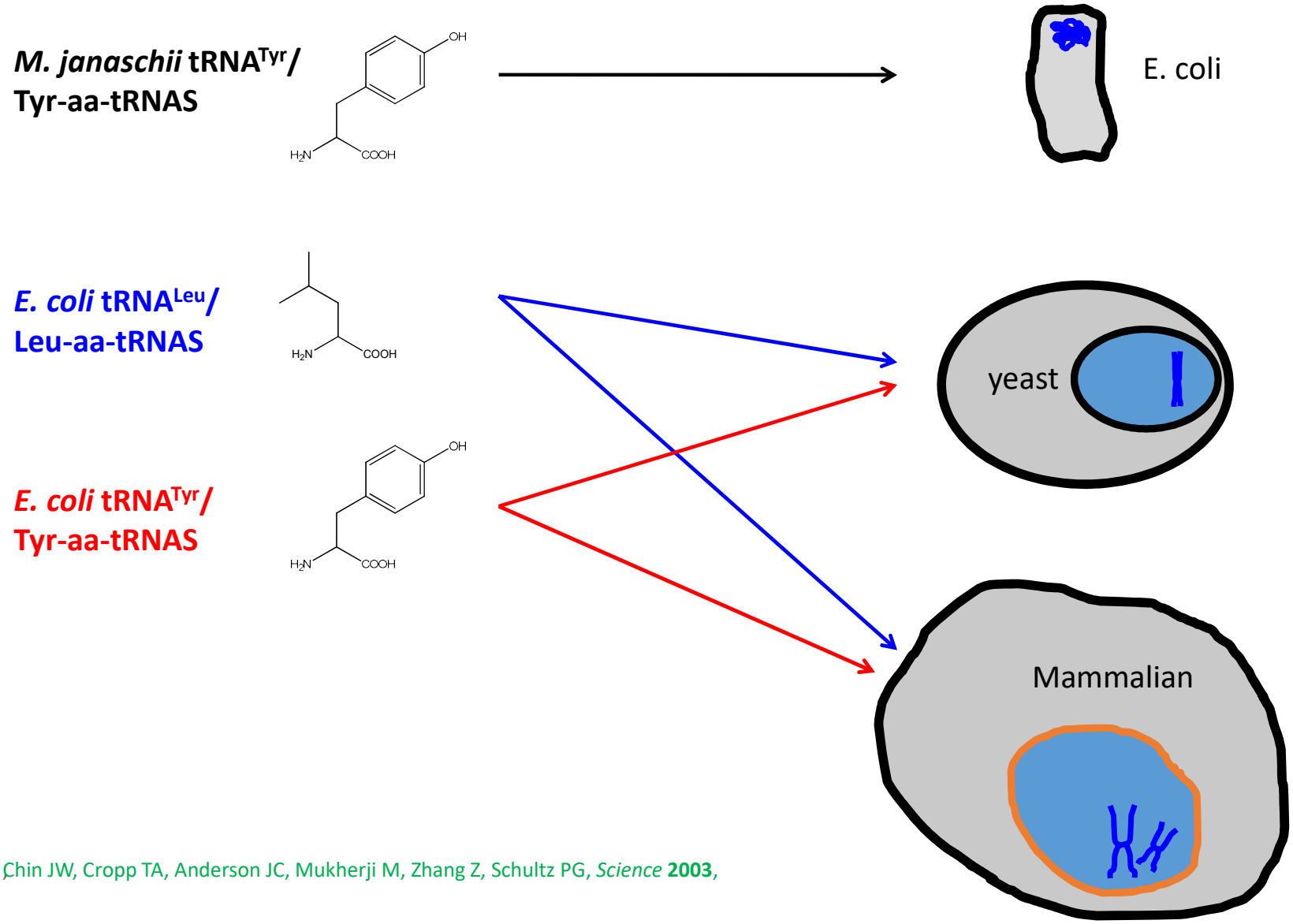


J. Xie, P. G. Schultz *Nature Rev. Mol. Cell Biol.* **2006**, *7*, 775-782.



New building blocks. A general method for genetically encoding unnatural amino acids into proteins.

L. Wang *Science* **2003**, *302*, 584-585.



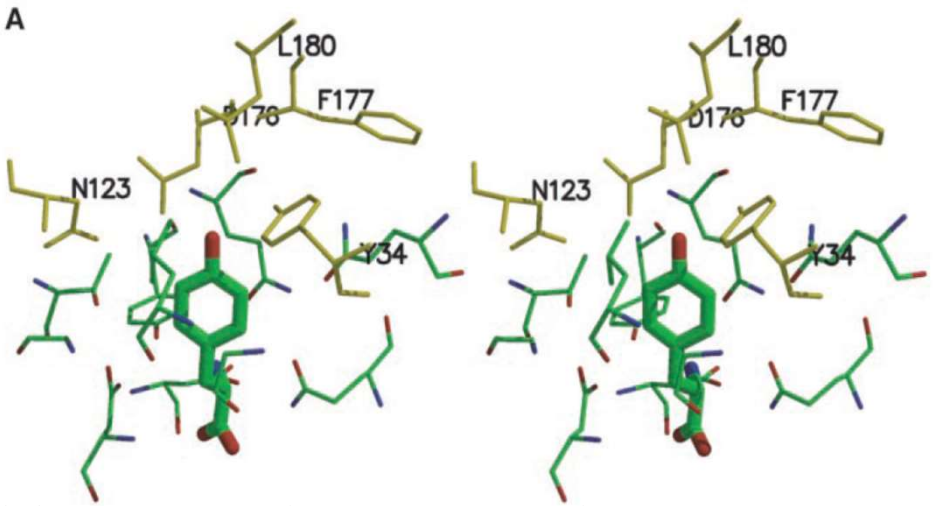
Chin JW, Cropp TA, Anderson JC, Mukherji M, Zhang Z, Schultz PG, *Science* **2003**,

The expanded eucaryotic genetic code

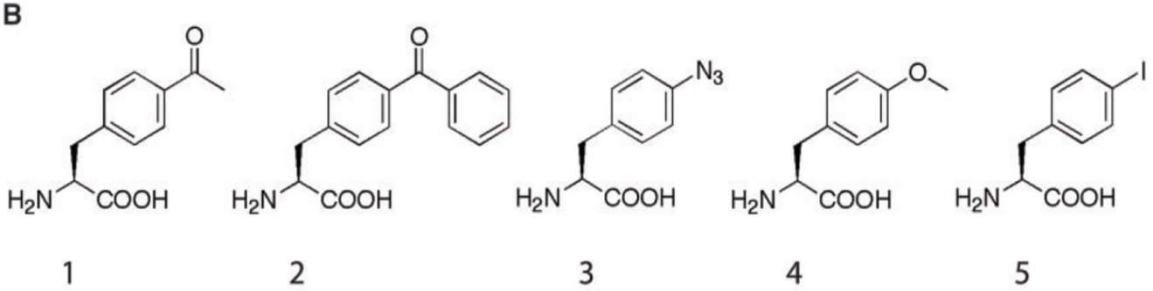
E. coli tyrosyl–tRNA synthetase (TyrRS) efficiently aminoacylates *E. coli* $tRNA_{CUA}$ when both are genetically encoded in *S. cerevisiae* but does not aminoacylate *S. cerevisiae* cytoplasmic tRNAs

In addition, *E. coli* tyrosyl $tRNA_{CUA}$ is a poor substrate for *S. cerevisiae* aminoacyl–tRNA synthetases but is processed and exported from the nucleus to the cytoplasm and functions efficiently in protein translation in *S. cerevisiae*

On the basis of the crystal structure of the homologous TyrRS from *Bacillus stearothermophilus* , five residues (Fig. 1A) in the active site of *E. coli* TyrRS were randomly mutated.



(**A**) Stereoview of the active site of *B. Stearothermophilus* tyrosyl–tRNA synthetase with bound tyrosine. The mutated residues (*E. Coli):* Tyr_{37} (*B. stearothermophilus* TyrRS residue Tyr_{34}), Asn_{126} (Asn_{123}), Asp_{182} (Asp_{176}), Phe_{183} (Phe_{177}), and Leu_{186} (Leu_{180}).

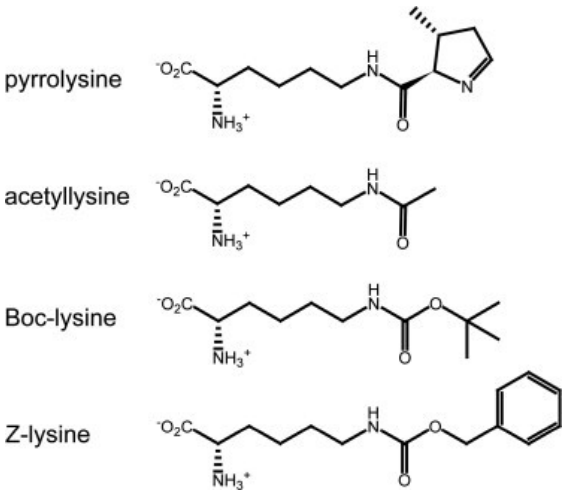


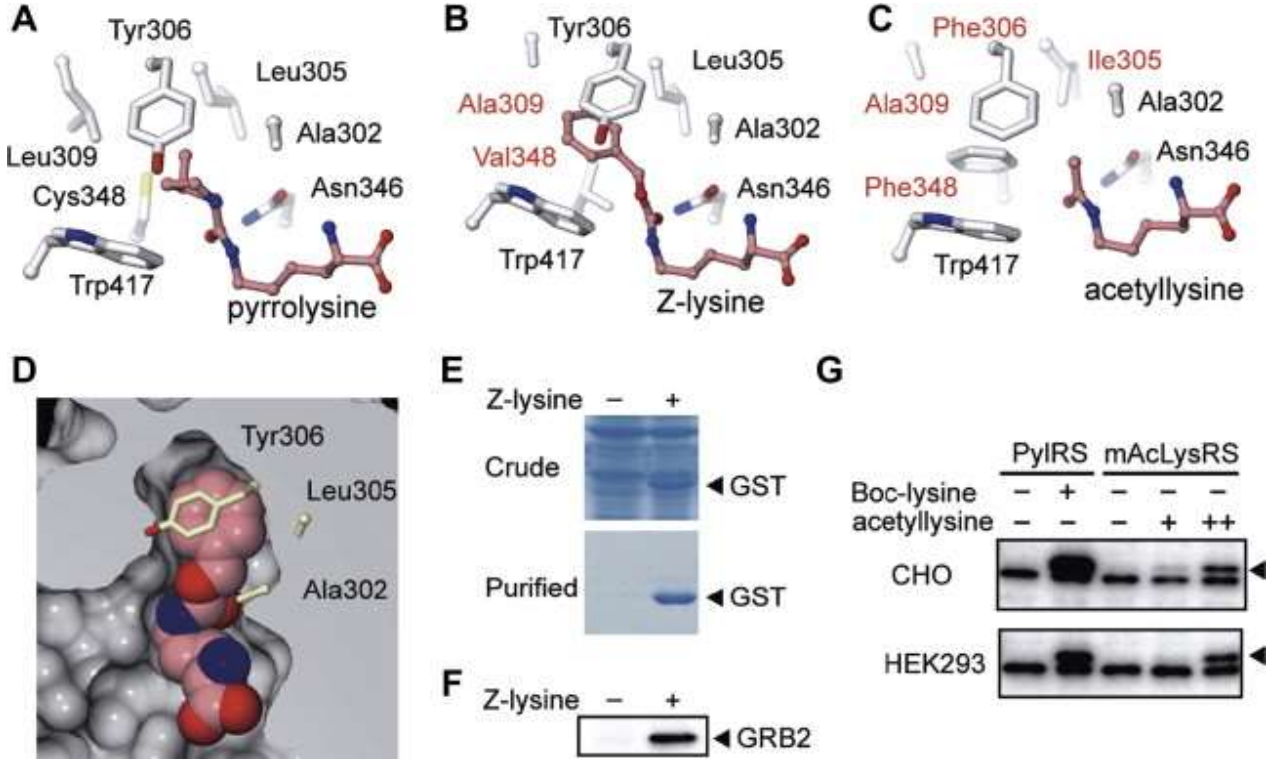
(**B**) Chemical structures of *p*acetyl-L-phenylalanine, **1**; *p*-benzoyl-L-phenylalanine, **2**; *p*-azido-L-phenylalanine, **3**; methyl-L-tyrosine, **4**; and *p*-iodo-L-tyrosine, **5**.

Chin JW, Cropp TA, Anderson JC, Mukherji M, Zhang Z, Schultz PG, *Science* **2003**, *301*, 964-967

The expanded eucaryotic genetic code

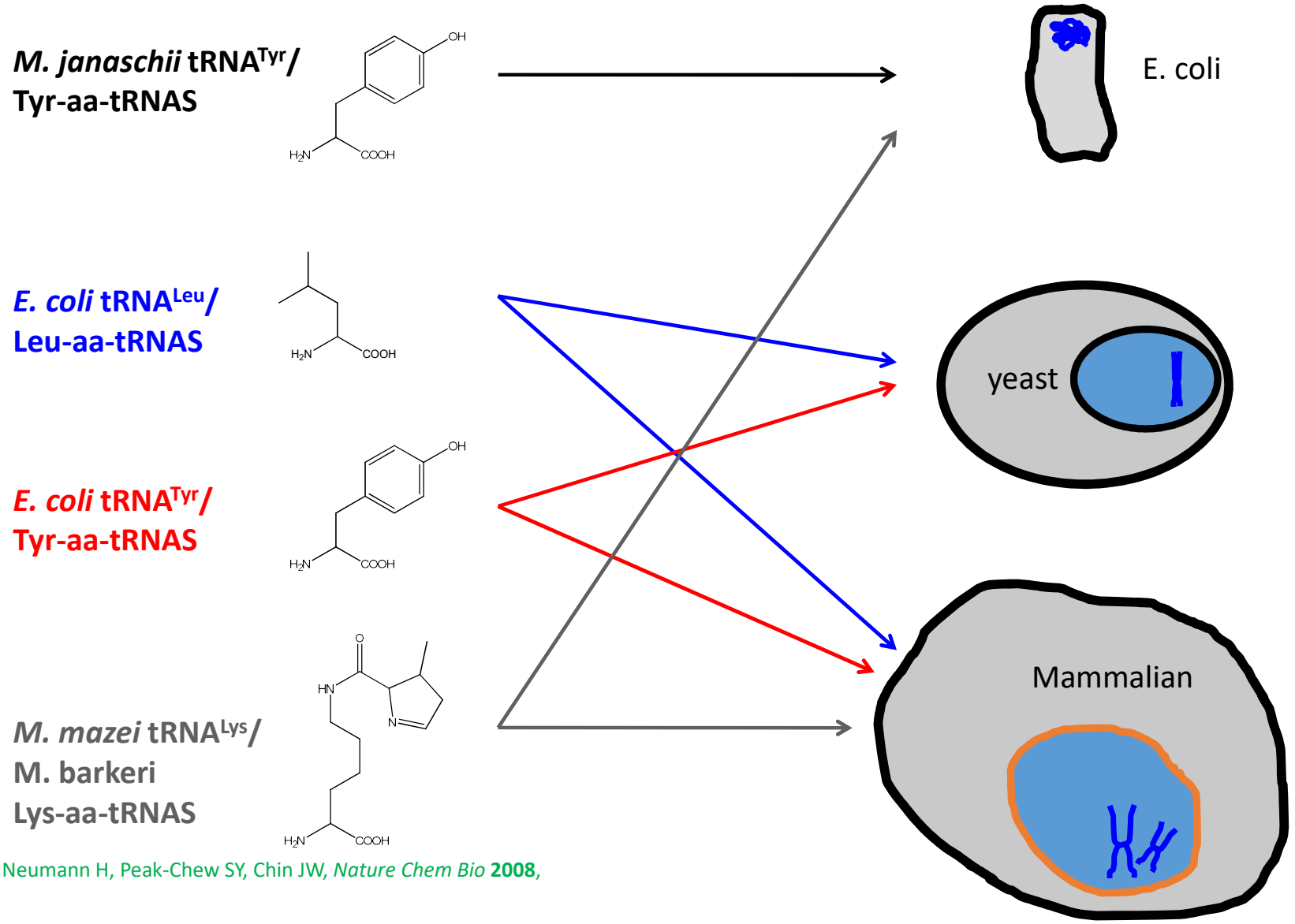
PylRS engineering and the site-specific incorporation of lysine derivatives into proteins in E. coli and mammalian cells.





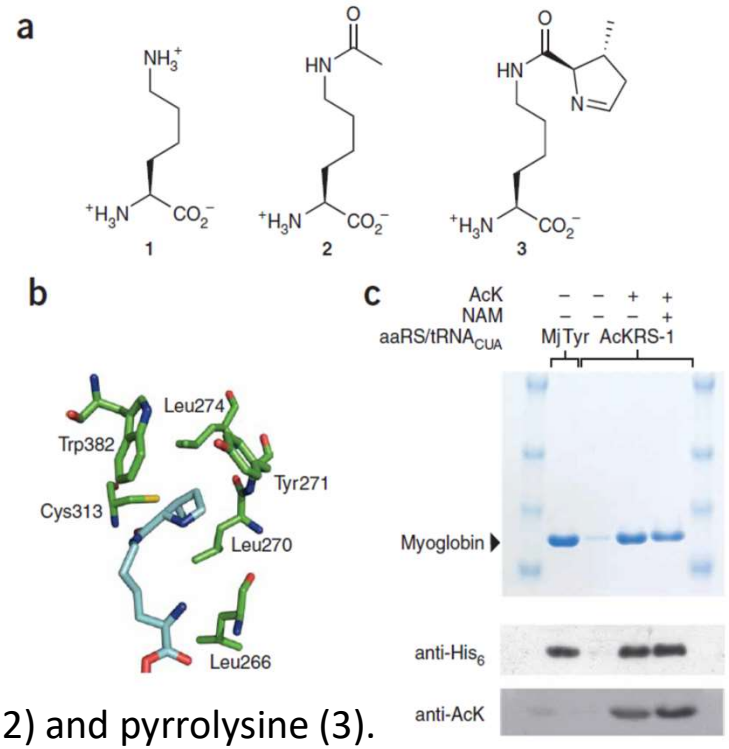
The amino-acid binding pocket with the bound pyrrolysine (X-ray *M. mazei* PylRS•pyrrolysine complex) (**A**), and the structural modeling of the binding pockets of ZLysRS **(B**) and AcLysRS (**C**) bound with Z-lysine and acetyllysine, respectively. (**D**) The space-filling model of Z-lysine in the binding pocket of ZLysRS. (**E**) Production of GST(Am25) containing Z-lysine in *E. coli* cells. (**F**) Production of GRB2(Am111)-FLAG containing Z-lysine in HEK293 c-18 cell. (**G**) The GRB2-FLAG molecules containing Boc-lysine and acetyllysine (CHO and HEK293 c-18 cells). Acetyllysine: 0 mM (−), 1.4 mM (+), and 14 mM (++).

Mukai et al, *Biochem, Biophys Res Com* **2008**, *371(4),* 818-822



Neumann H, Peak-Chew SY, Chin JW, *Nature Chem Bio* **2008**,

Design and evolution of an MbPylRS/tRNA$_{CUA}$ pair for the genetic incorporation of N$^{\varepsilon}$-acetyllysine.

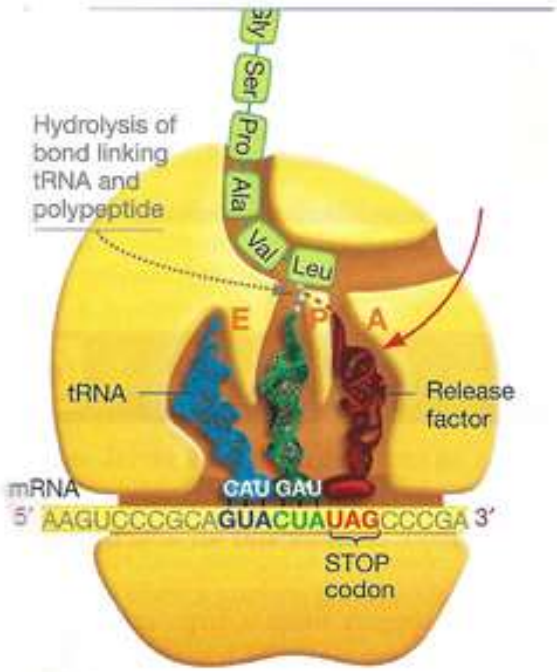


a) Structure of lysine (1),N$^{\varepsilon}$-acetyllysine(2) and pyrrolysine (3).
b) Structure of the active site of *M. Mazei* PylRS bound to pyrrolysine. The active site residues shown are conserved between *M. Mazei* PylRS and *M. Barkeri* PylRS. These residues form the hydrophobic binding pocket of pyrrolysine and are mutated in the library to each of the common 20 amino acids. PDB: 2Q7H.
c) Myoglobin-His 6 produced in the presence of MjTyrRS/MjtRNA$_{CUA}$ (lane 1) or in the presence of AcKRS-1 without or with 1 mM N$^{\varepsilon}$-acetyllysine (AcK, lanes 2 and 3, respectively), or in the presence of 1 mM N$^{\varepsilon}$-acetyllysine and 50 mM NAM (lane 4).

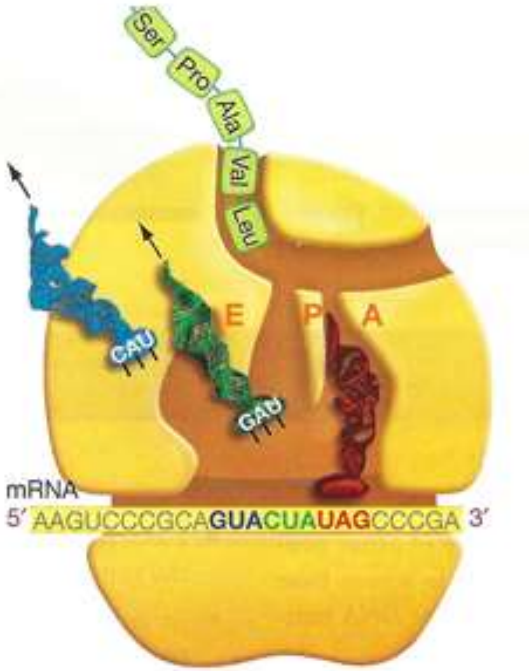
Neumann H, Peak-Chew SY, Chin JW, *Nature Chem Bio* **2008**,

Translation: RNA → proteins

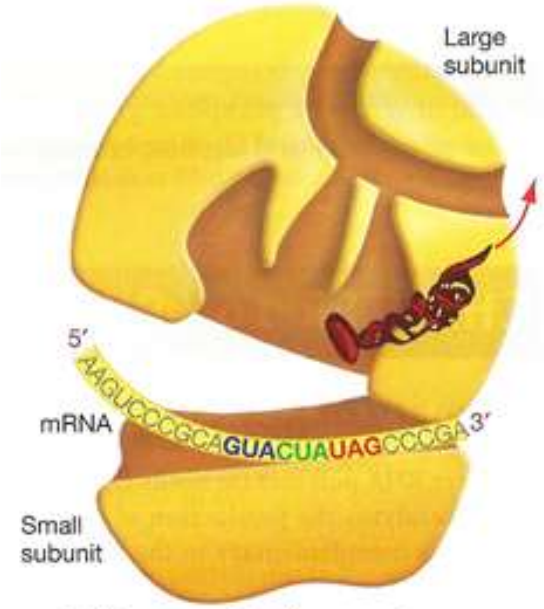
Translation termination



1. Release factor binds to stop codon. When translocation exposes a stop codon, a release factor fills the A site. The release factor breaks the bond linking the tRNA in the P site to the polypeptide chain.



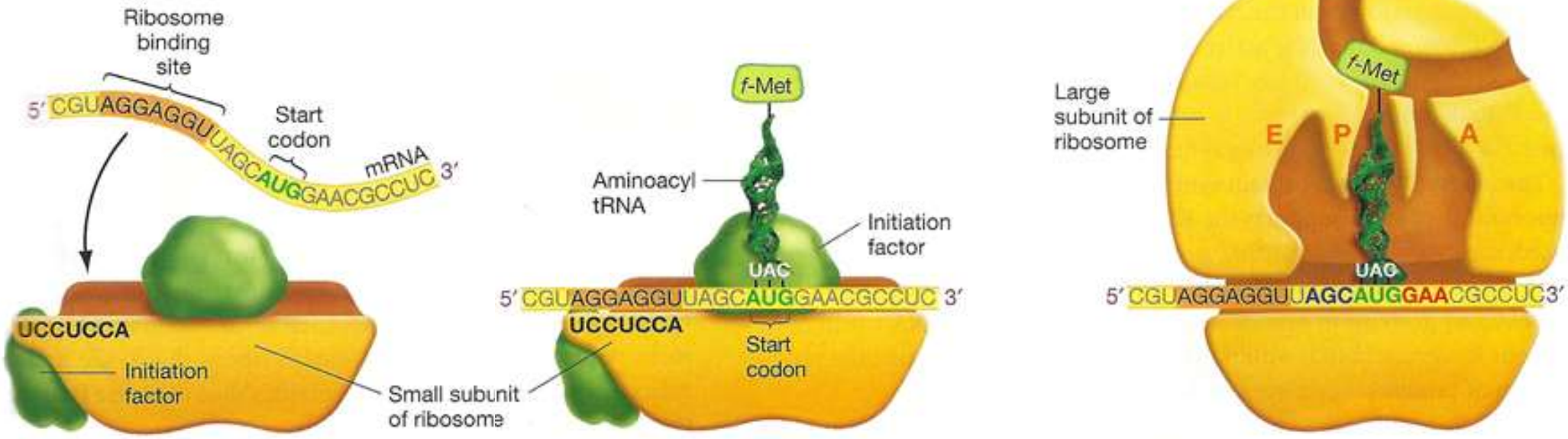
2. Polypeptide is released. The hydrolysis reaction frees the polypeptide, which is released from the ribosome. The empty tRNAs are released either along with the polypeptide or...



3. Ribosome subunits separate. ...when the ribosome separates from the mRNA, and the two ribosomal subunits dissociate. The subunits are ready to attach to the start codon of another message and start translation anew.

Translation: RNA → proteins

Translation initiation

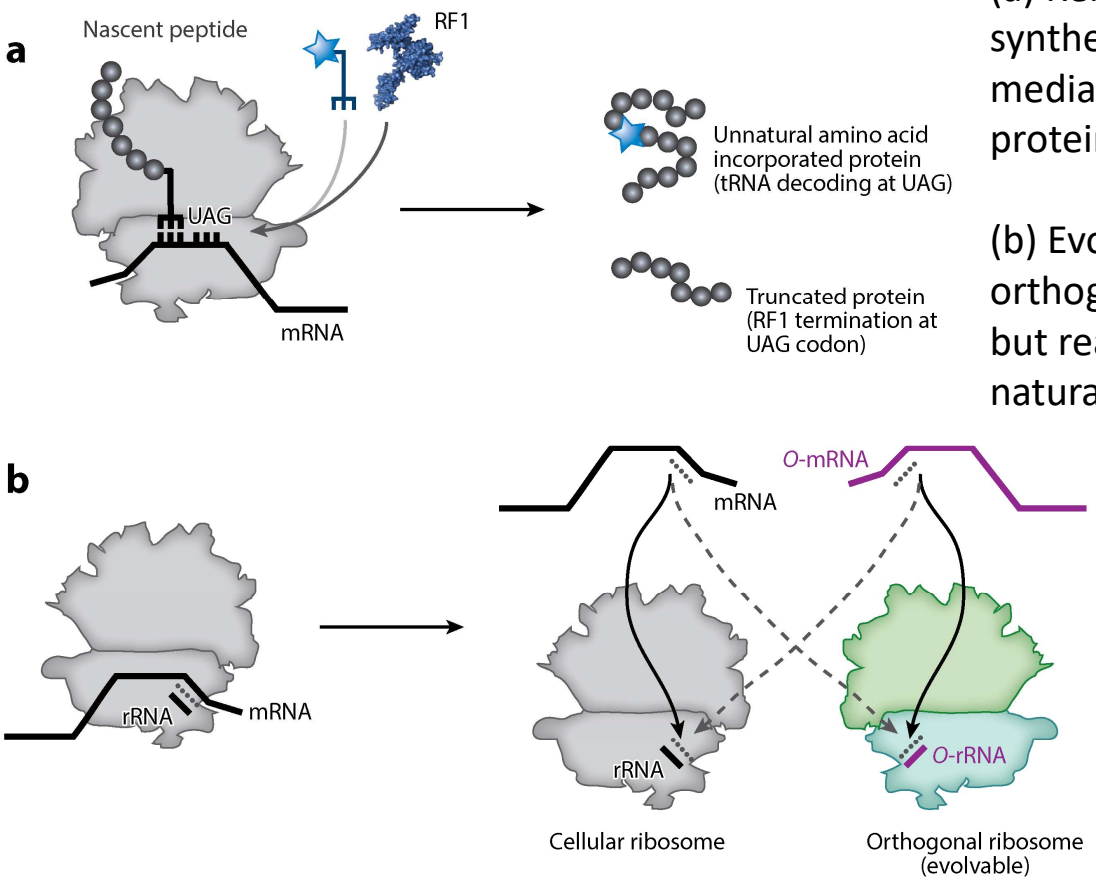


1. mRNA binds to small subunit. Ribosome binding site sequence binds to a complementary sequence in an RNA molecule in the small subunit of the ribosome, with the help of protein initiation factors.

2. Initiator aminoacyl tRNA binds to start codon.

3. Large subunit of ribosome binds, completing ribosome assembly. Translation begins.

Strategies to enhance unnatural amino acid incorporation in response to the amber stop codon in Escherichia coli.

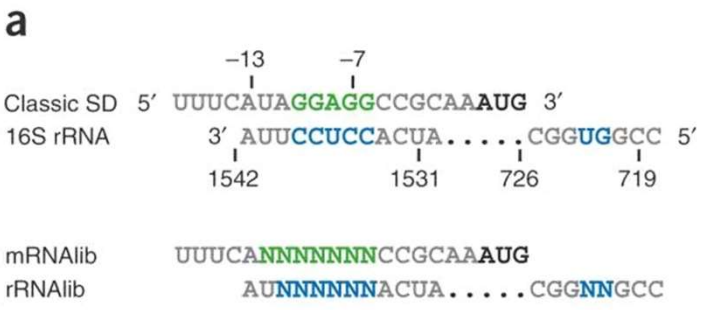


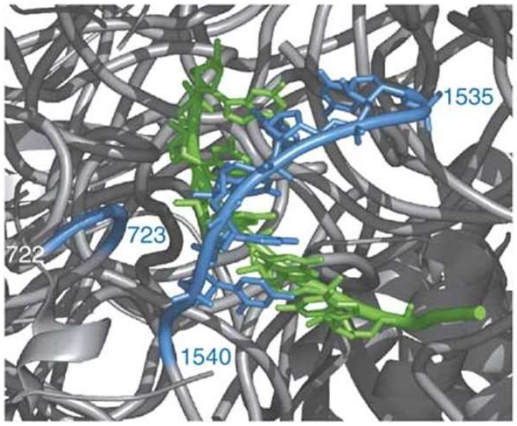
(a) Release factor 1 (RF1)-mediated termination of protein synthesis competes with amber-suppressor transfer RNA (tRNA)-mediated elongation of protein synthesis that yields a full-length protein bearing the unnatural amino acid.

(b) Evolution of an orthogonal ribosome in *E. coli.* The orthogonal ribosome functions alongside the natural ribosome but reads a distinct message that is not a substrate for the natural ribosome.

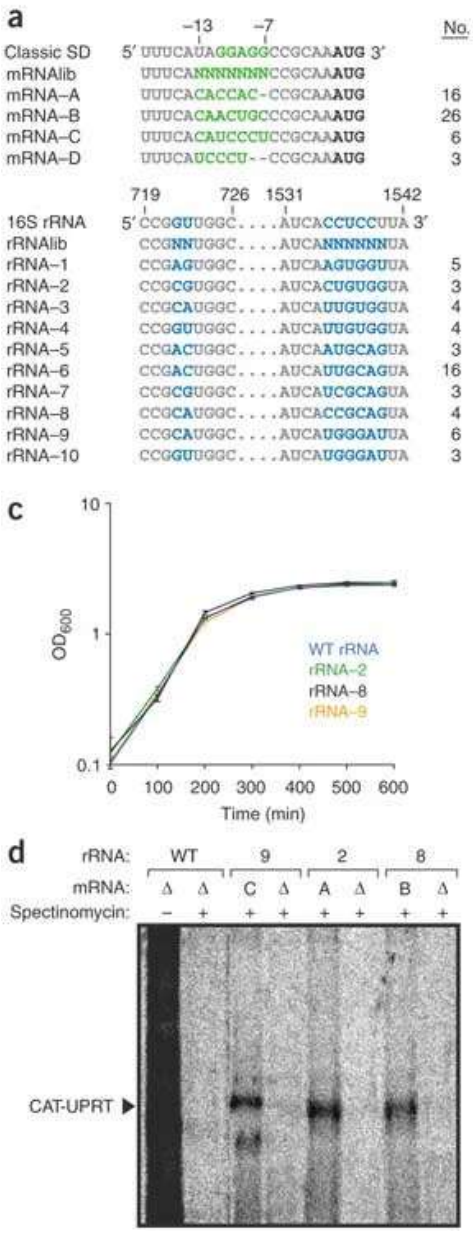
O. Rackham, J. W. Chin *Nature Chem. Biol*, **2005**, *1*, 159-166

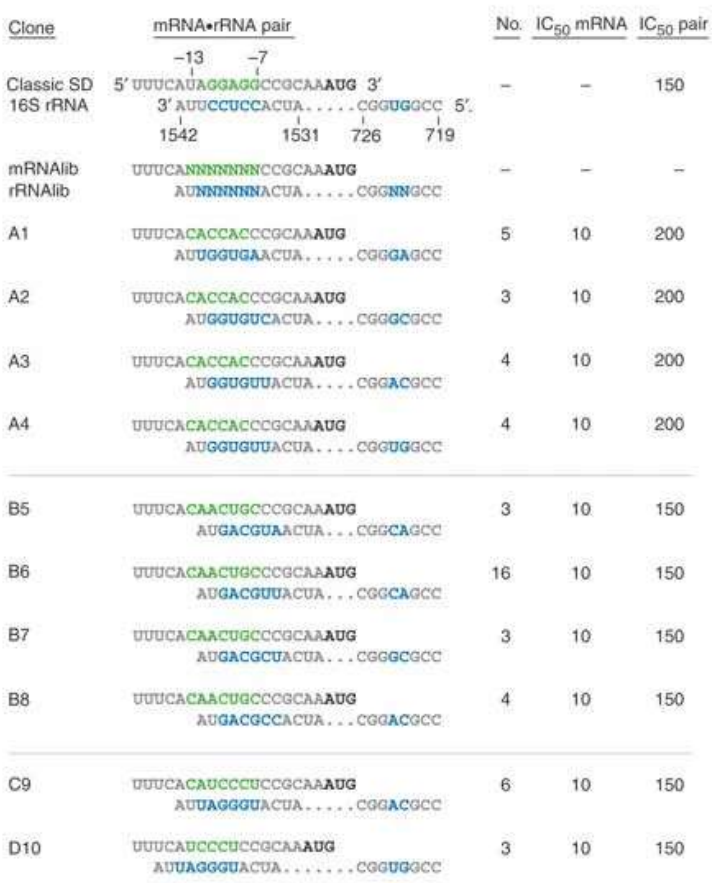
The orthogonal ribosome design





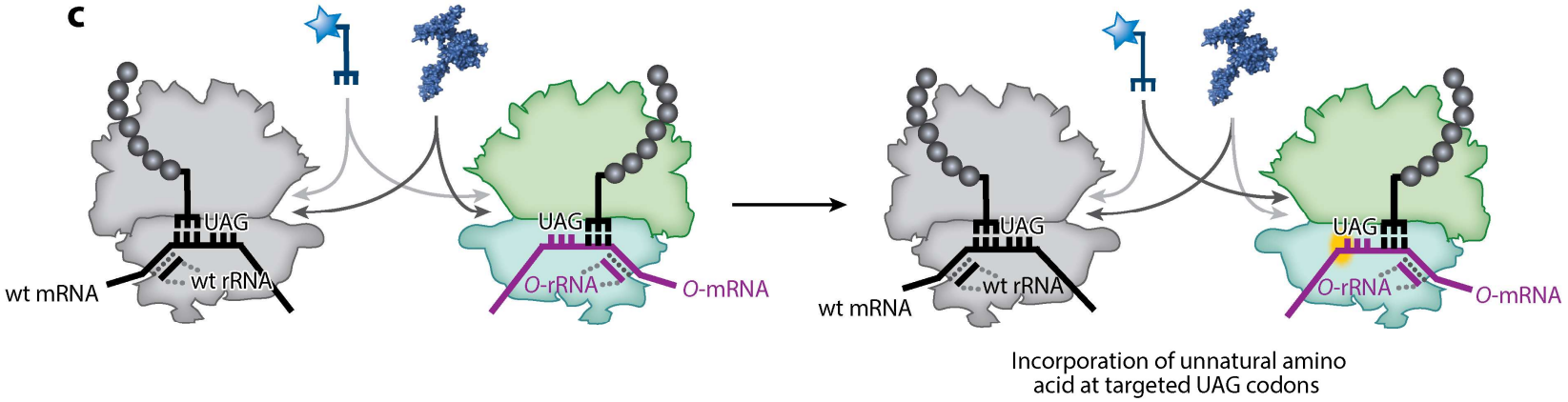
The design of ribosome and mRNA libraries for the selection of orthogonal pairs.
(a) The classic SD*ASD interaction (top) and the nucleotides randomized in mRNAlib and rRNAlib (bottom). (b) The SD*ASD interaction helix in the ribosome.

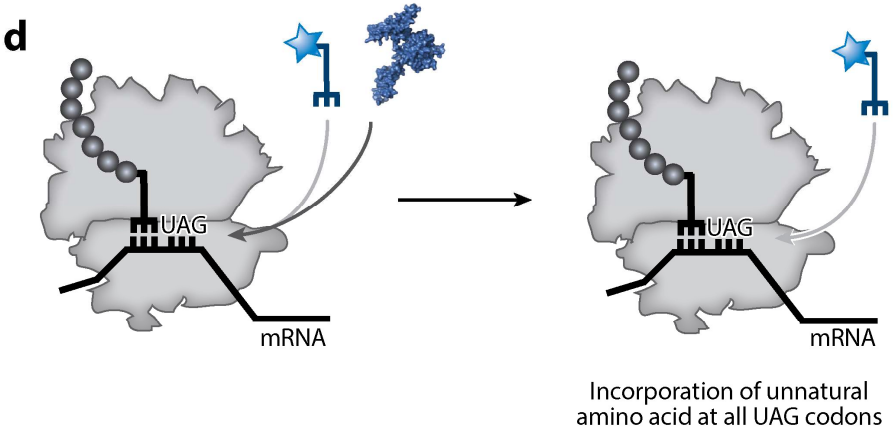




O. Rackham, J. W. Chin
Nature Chem. Biol, **2005**, *1*, 159-166

Strategies to enhance unnatural amino acid incorporation in response to the amber stop codon in Escherichia coli.



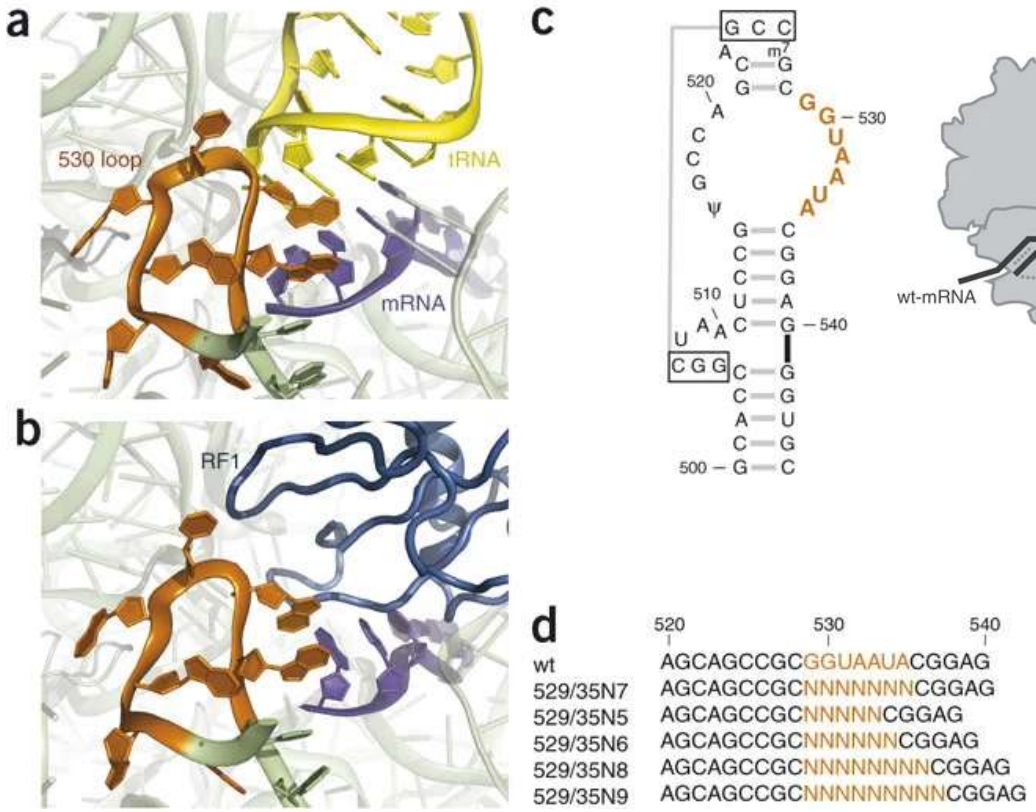


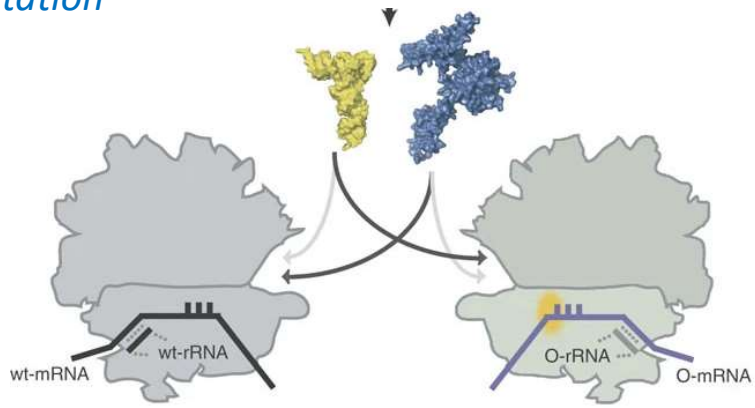
Chin JW. 2014.
Annu. Rev. Biochem. 83:379–408

(c) The orthogonal ribosome has been evolved to efficiently decode amber-suppressor tRNAs, differentiating the decoding of amber codons on the orthogonal and cellular messages and enhancing unnatural amino acids on orthogonal messages without enhancing the incorporation of unnatural amino acids at genomically encoded stop codons.

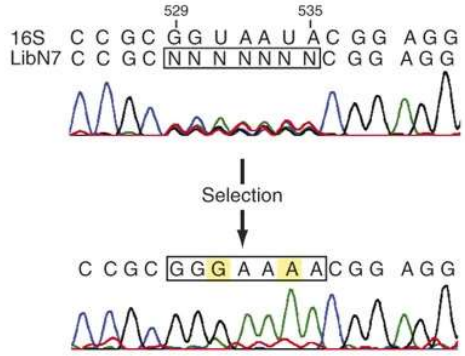
(d) RF1 knockouts and knockdowns for unnatural amino acid incorporation in E. coli. The strategies increase the incorporation of unnatural amino acids in response to the desired stop codon and any genomically encoded stop codons.

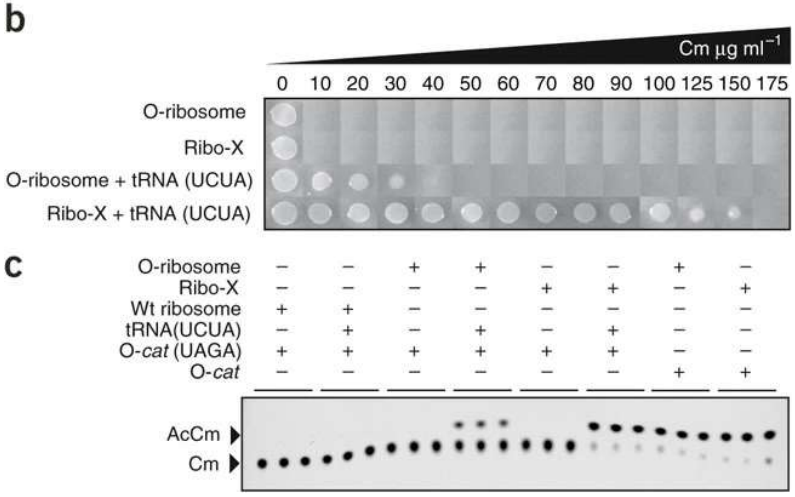
Ribosome evolution to overcome the RF-1 limitation





Selection and phenotypic characterization of ribo-X. The selection of ribosomes that decode UAGA and UAG codons, using cognate tRNAs derived from $tRNA^{ser2}$.





Design of ribosome-decoding libraries. (a) Structure of a tRNA anticodon stem loop (yellow) bound to mRNA (purple) in the A-site of the ribosome (green). The 530 loop is shown in orange. (b) Structural model of RF-1 (blue) bound in the A-site of the ribosome. (c) Secondary structure of the 530 loop. The region targeted for mutation is colored orange. (d) The sequence of ribosome-decoding libraries.

K. Wang, H. Neumann, S.Y. Peak-Chew, J. W. Chin *Nature Biotech.*, **2007**, *25*, 770-777

Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome

The *Methanococcus jannaschii* TyrRS–$tRNA_{CUA}$ and the *Methanosarcina barkeri* MbPylRS–$tRNA_{CUA}$ orthogonal pairs have been evolved to incorporate a range of unnatural amino acids in response to the amber codon in *Escherichia coli*.

The general limitation: low efficiency incorporation of a single type of unnatural amino acid at a time, because every triplet codon in the universal genetic code is used in encoding the synthesis of the proteome.

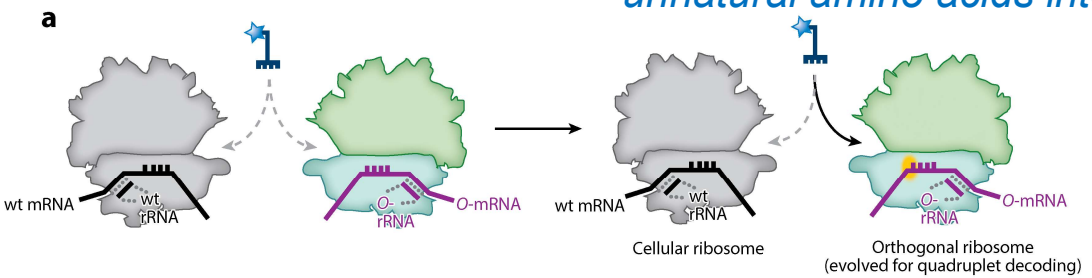
An orthogonal ribosome (ribo-Q1) efficiently decodes a series of quadruplet codons and the amber codon, providing several blank codons on an orthogonal messenger RNA, which it specifically translates. By creating mutually orthogonal aminoacyl-tRNA synthetase–tRNA pairs and combining them with ribo-Q1, incorporation of distinct unnatural amino acids in response to two of the new blank codons on the orthogonal mRNA has been achieved.

It will be possible to encode more than 200 unnatural amino acid combinations using this approach.

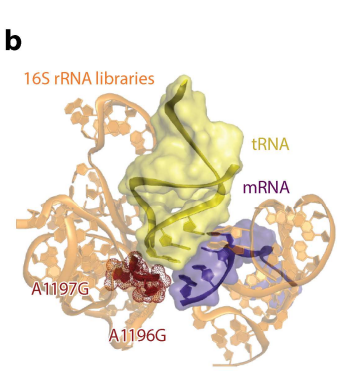
The ribo-Q1 independently decodes a series of quadruplet codons

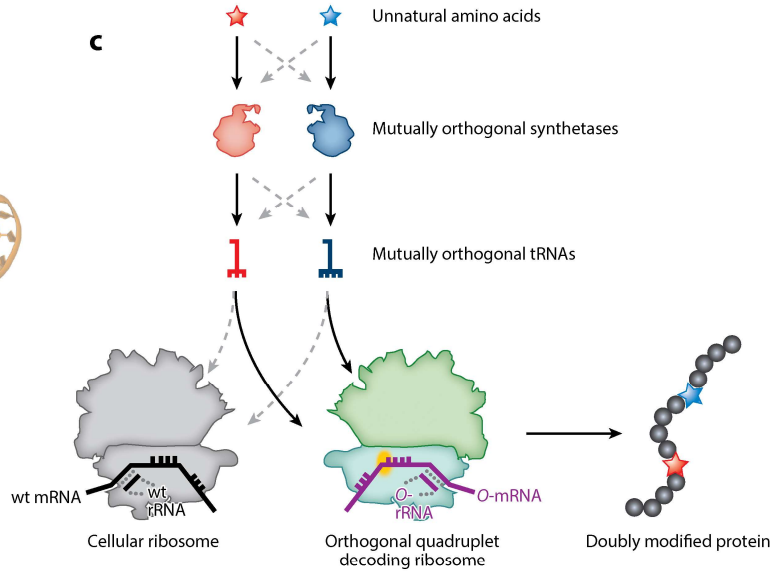
H. Neumann, K. Wang, L. Davis, M. Garcia-Alai, J. W. Chin *Nature*, **2010**, *464*, 441-444

Evolution of an orthogonal quadruplet decoding ribosome enables the incorporation of multiple distinct unnatural amino acids into a single polypeptide.



(*a*) The orthogonal ribosome has been evolved in the laboratory to efficiently decode quadruplet codons.



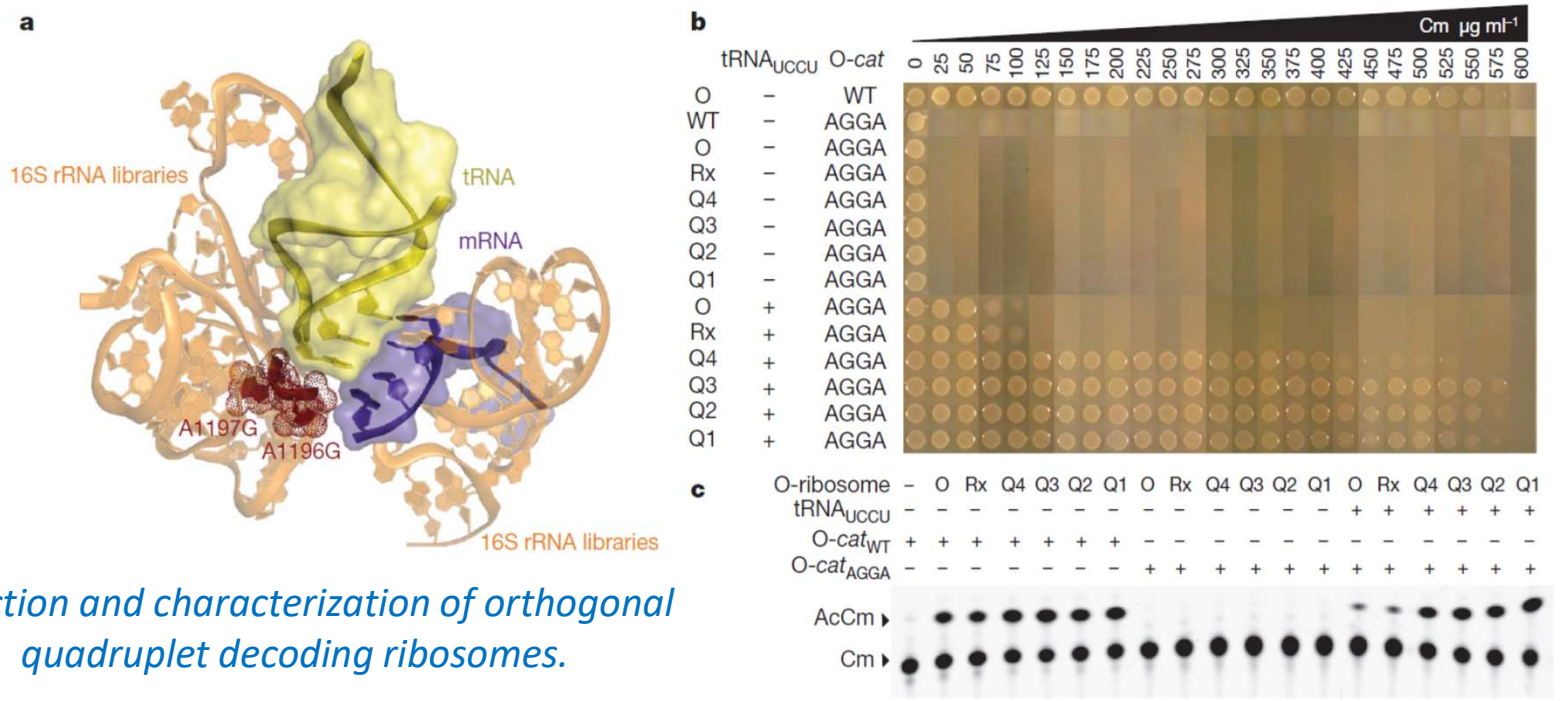


(*b*) Mutations in the A site of 16S ribosomal RNA (rRNA) facilitate quadruplet decoding on the orthogonal ribosome.

Chin JW. 2014.
Annu. Rev. Biochem. 83:379–408

(*c*) Genetically encoding multiple unnatural amino acids via orthogonal translation. Mutually orthogonal synthetase/tRNA (transfer RNA) pairs have been used to direct the incorporation of distinct unnatural amino acids into a single polypeptide. The extended anticodon or amber-suppressor tRNAs are selectively decoded on the evolved orthogonal ribosome, creating a parallel translation pathway in the cell.

Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome



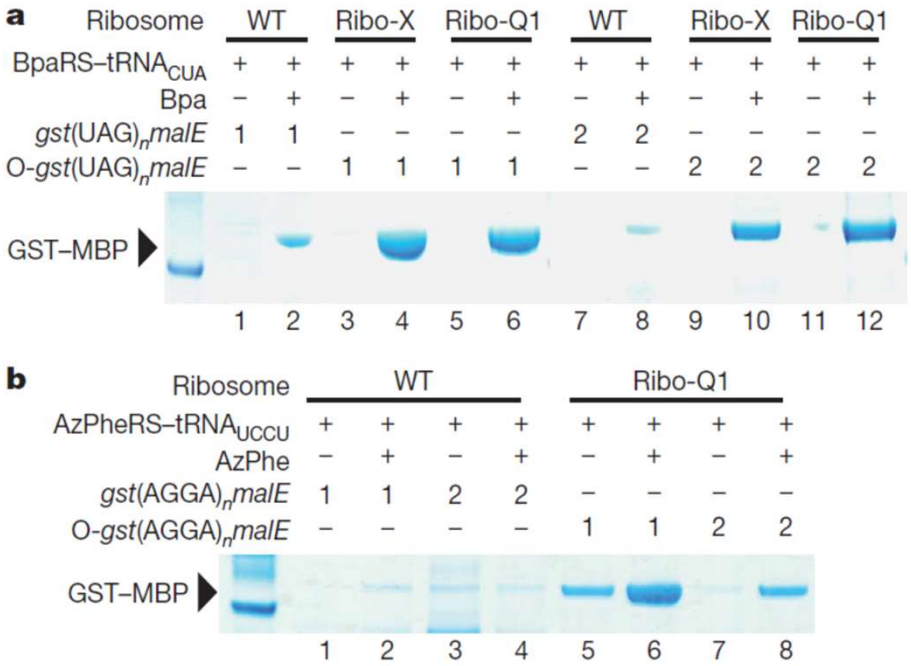
Selection and characterization of orthogonal quadruplet decoding ribosomes.

a) Mutations in quadruplet decoding ribosomes form a structural cluster close to the space potentially occupied by an extended anticodon tRNA. Selected nucleotides are shown in red. **b)** Ribo-Qs substantially enhance the decoding of quadruplet codons. The $tRNA^{Ser}_{UCCU}$-dependent enhancement in decoding AGGA codons in the O-*Cat* (AGGA 103, AGGA 146) gene was measured by survival on increasing concentrations of chloramphenicol (Cm). WT, wild type. **c)** as in **b** , but measuring CAT enzymatic activity directly by thin-layer chromatography.AcCm, acetylated chloramphenicol; O, O-ribosome; Q1–Q4, ribo-Q1–Q4;Rx, ribo-X

H. Neumann, K. Wang, L. Davis, M. Garcia-Alai, J. W. Chin *Nature*, **2010**, *464*, 441-444

Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome

Enhanced incorporation of unnatural amino acids in response to amber and quadruplet codons with ribo-Q.



a) Ribo-Q1 incorporates Bpa as efficiently as ribo-X.

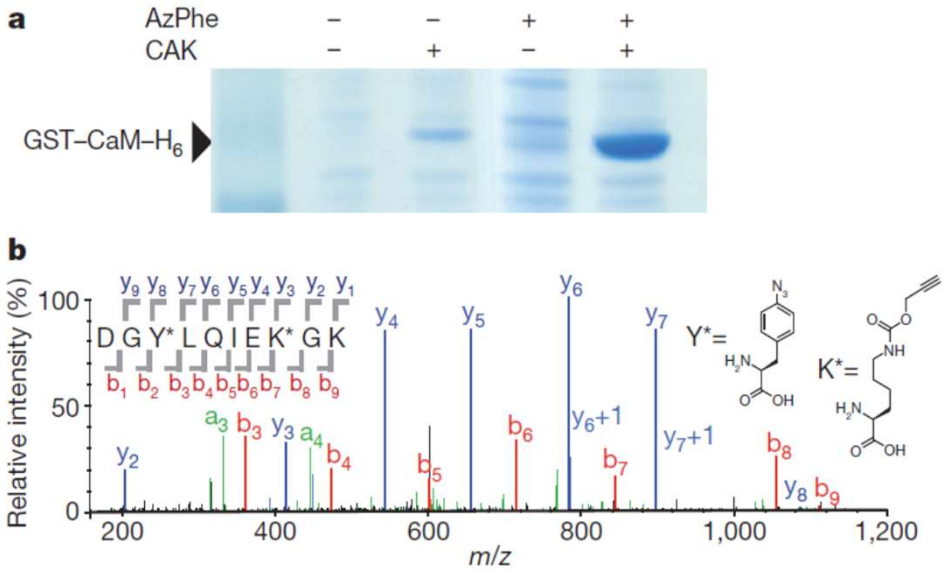
b) Ribo-Q1 enhances the efficiency of AzPhe incorporation in response to the AGGA quadruplet codon using AzPheRS*–tRNA$_{UCCU}$.

$(UAG)_n$ or $(AGGA)_n$ describes the number of amber or AGGA codons (n) between gst and malE.

H. Neumann, K. Wang, L. Davis, M. Garcia-Alai, J. W. Chin *Nature*, **2010**, *464*, 441-444

Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome

Encoding an azide and an alkyne in a single protein by orthogonal translation

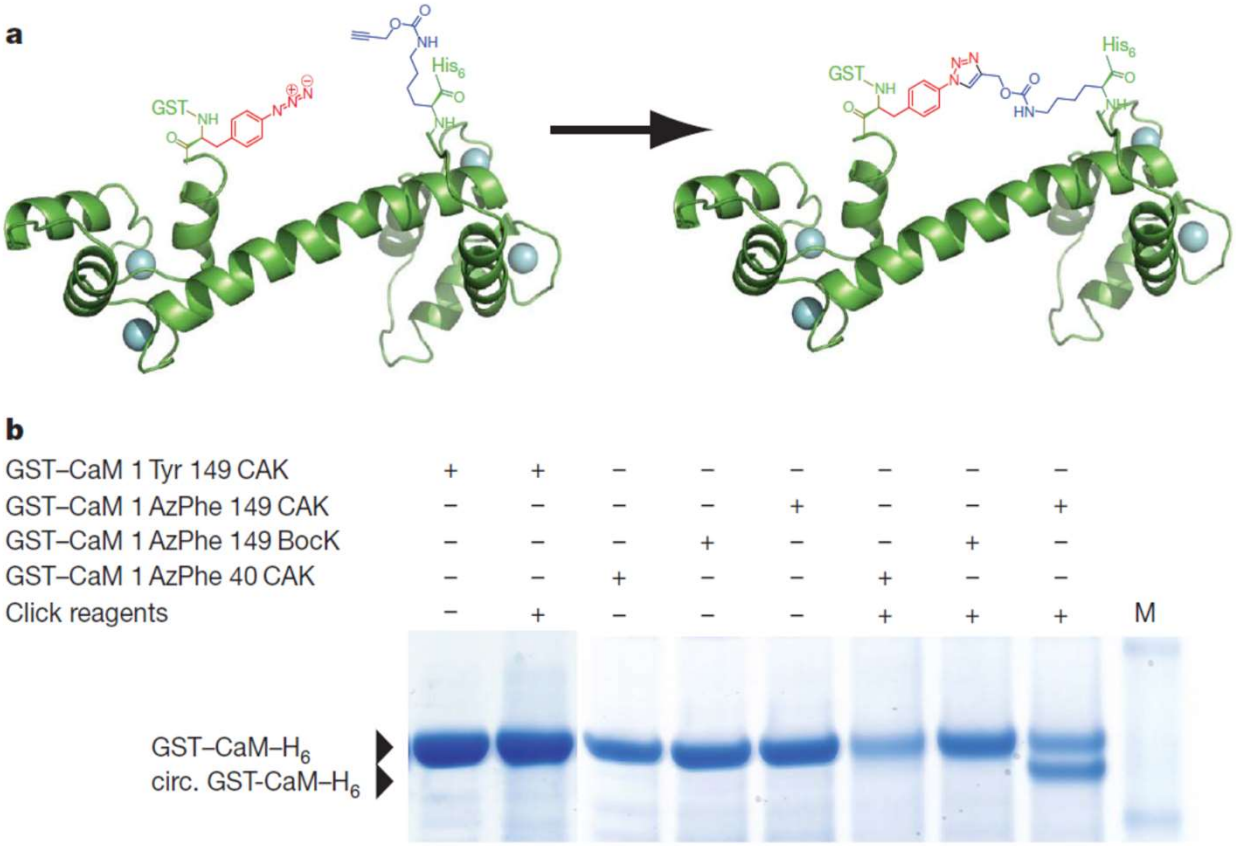


a) Expression of GST–CaM–His_6 (a GST–calmodulin–His_6 fusion) containing two unnatural amino acids. An orthogonal gene producing a GST–CaM–His_6 fusion that contains an AGGA codon at position 1 and an amber codon at position 40 of calmodulin was translated by ribo-Q1 in the presence of AzPheRS*–$tRNA_{UCCU}$ and MbPylRS–$tRNA_{CUA}$.
b) LC–MS/MS analysis of the incorporation of two distinct unnatural amino acids into the linker region of GST–MBP.
Y* - AzPhe; K* - CAK.

H. Neumann, K. Wang, L. Davis, M. Garcia-Alai, J. W. Chin *Nature*, **2010**, *464*, 441-444

Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome

Genetically directed cyclization of calmodulin by a Cu(I)-catalysed Huisgen's [2+3]-cycloaddition



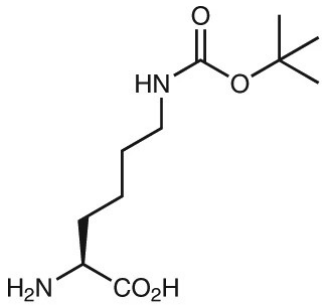
a) Structure of calmodulin indicating the sites of incorporation of AzPhe and CAK and their triazole product.

b) GST–CaM–His_6 1 AzPhe 149 CAK specifically cyclizes with Cu(I)-catalyst.

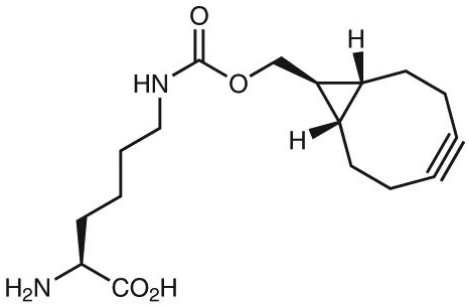
BocK - N^{ε}-*tert*-butyl-oxycarbonyl-*L*-lysine;
circ. - circularized protein.

H. Neumann, K. Wang, L. Davis, M. Garcia-Alai, J. W. Chin *Nature*, **2010**, *464*, 441-444

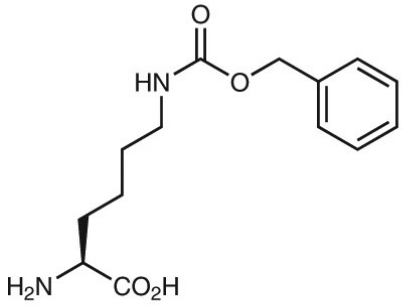
Mutually orthogonal pyrrolysyl-tRNA synthetase(PylRS)/tRNA pairs



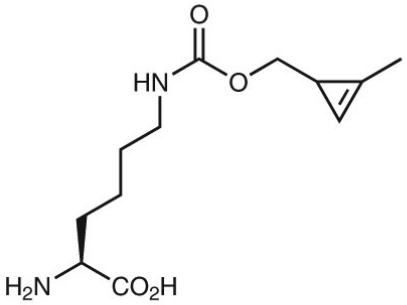
1
BocK



2
BCNK



3
CbzK



4
CypK

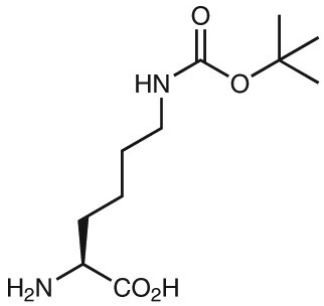
The *Methanosarcina mazei* (*Mm*) pyrrolysyl-tRNA synthetase (PylRS encoded by *PylS*)/*Mm*^{Pyl}tRNA$_{CUA}$ (encoded by *MmPylT*) pair, along with the homologous pair from *Methanosarcina barkeri* (*Mb*), has been extensively developed for the co-translational incorporation of ncAAs into proteins via genetic code expansion.

*Mm*PylRS and *Mb*PylRS do not recognize the anticodon of their cognate ^{Pyl}tRNA$_{CUA}$s, a feature that facilitates the decoding of diverse codons by these pairs, through mutation of their anticodons.

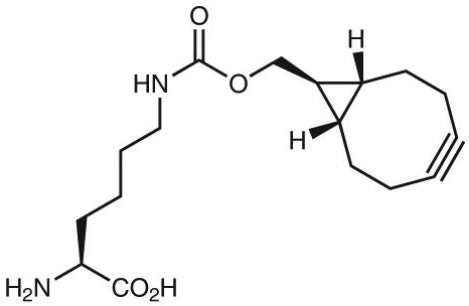
The *Mm* or *Mb* PylRS/^{Pyl}tRNA pairs have been used in combination with derivatives of the *Methanocaldococcus janaschii (Mj)* TyrRS/^{Tyr}tRNA$_{CUA}$ pair, to direct the co-translational incorporation of several pairs of ncAAs into proteins in *Escherichia coli*. However, *Mj*TyrRS recognizes the anticodon of its cognate tRNA, which restricts the codons this pair can be easily altered to efficiently decode, and this pair has primarily been used to incorporate aromatic ncAAs related to phenylalanine.

J.C.W. Willis, J. W. Chin *Nature Chem.*, **2018**, *10*, 831-837

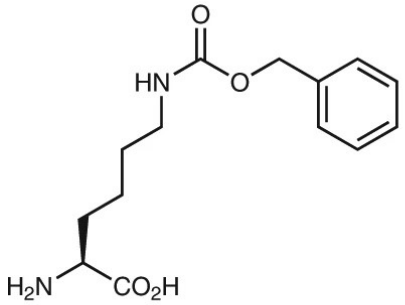
Mutually orthogonal pyrrolysyl-tRNA synthetase(PylRS)/tRNA pairs



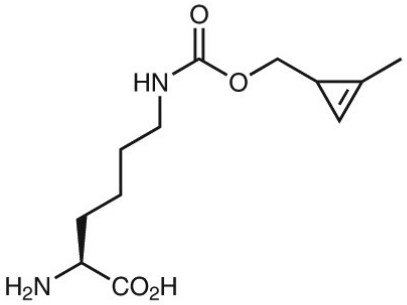
1
BocK



2
BCNK



3
CbzK



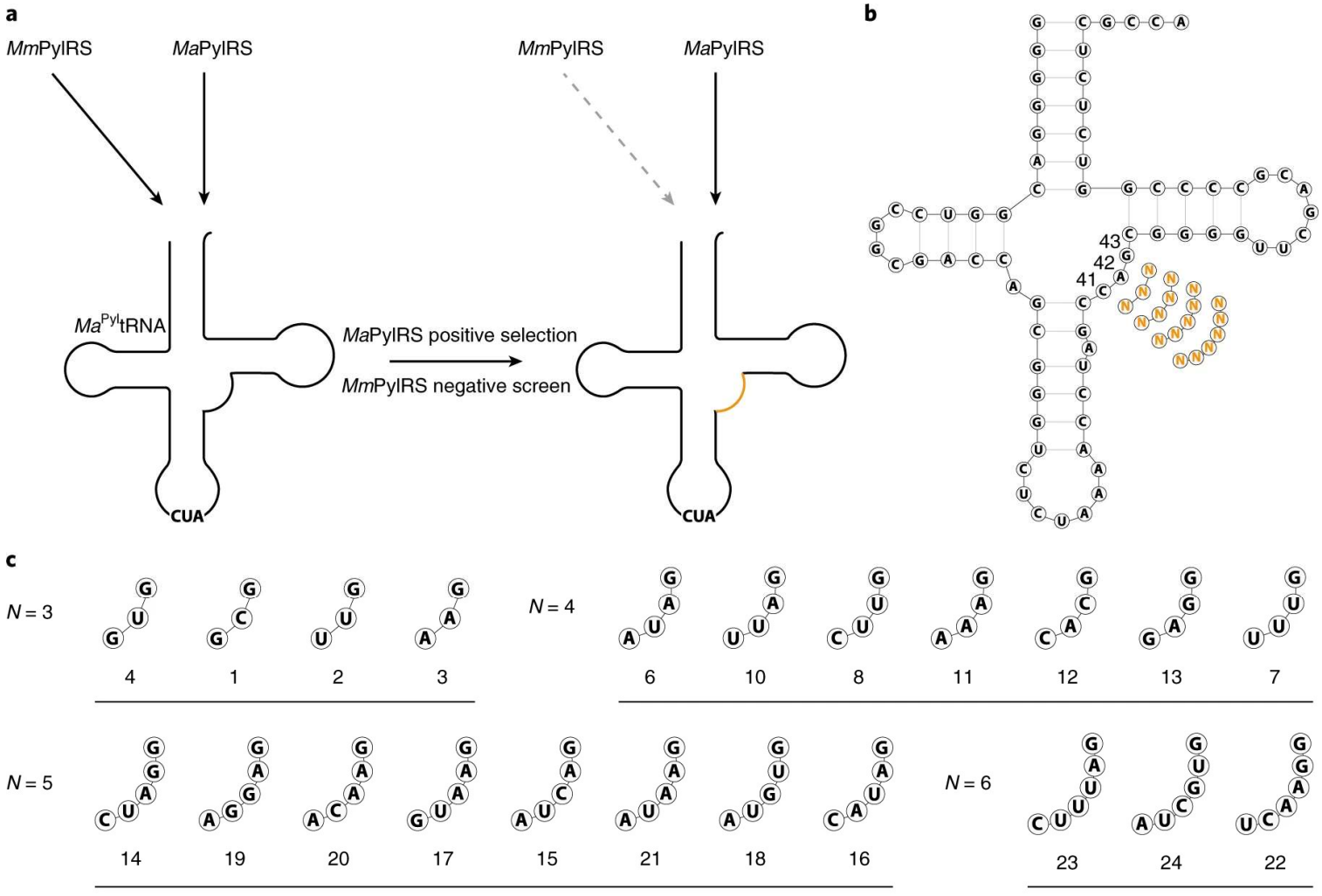
4
CypK

*Mm*PylRS and *Mb*PylRS are composed of an N-terminal domain that binds to the T-arm and variable loop of their cognate tRNAs, linked to a C-terminal catalytic domain (PylRSΔ184 in *M. mazei*). Although the catalytic domain can (inefficiently) aminoacylate its cognate tRNA in vitro, the full-length protein is absolutely required for measureable amber suppression activity in cells. Another group of PylRS enzymes, commonly exemplified by *Desulfitobacterium hafniense* (*Dh*), have separate genes encoding the N-terminal domain (*Dh*PylSn) and the C-terminal catalytic domain (*Dh*PylSc) of PylRS as distinct polypeptides. These polypeptides are believed to assemble to create a functional synthetase in vivo. In *E. coli*, the C-terminal protein, *Dh*PylSc, is reported to have less than 1% of the activity of *Mb*PylRS.

New PylRS/^{Pyl}tRNA pairs were created, that are mutually orthogonal to the *Mm*PylRS/*Mm*^{Pyl}tRNA pair. The two PylRS/^{Pyl}tRNA-derived pairs can function in the same cell, decode distinct codons, and incorporate distinct ncAAs.

J.C.W. Willis, J. W. Chin *Nature Chem.*, **2018**, *10*, 831-837

Mutually orthogonal pyrrolysyl-tRNA synthetase(PylRS)/tRNA pairs



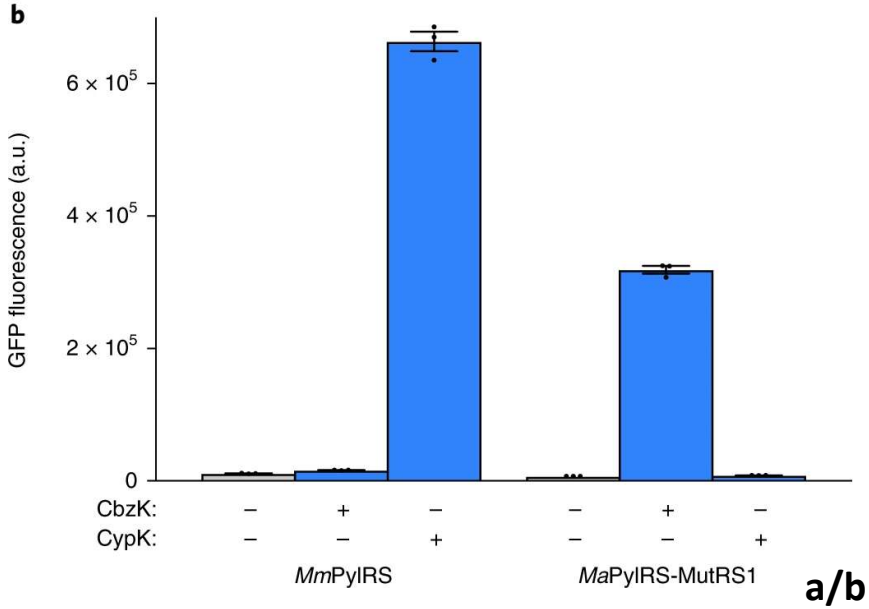
a, Evolution of *Ma*^{Pyl}tRNA to abolish its function with *Mm*PylRS while preserving its function with *Ma*PylRS.

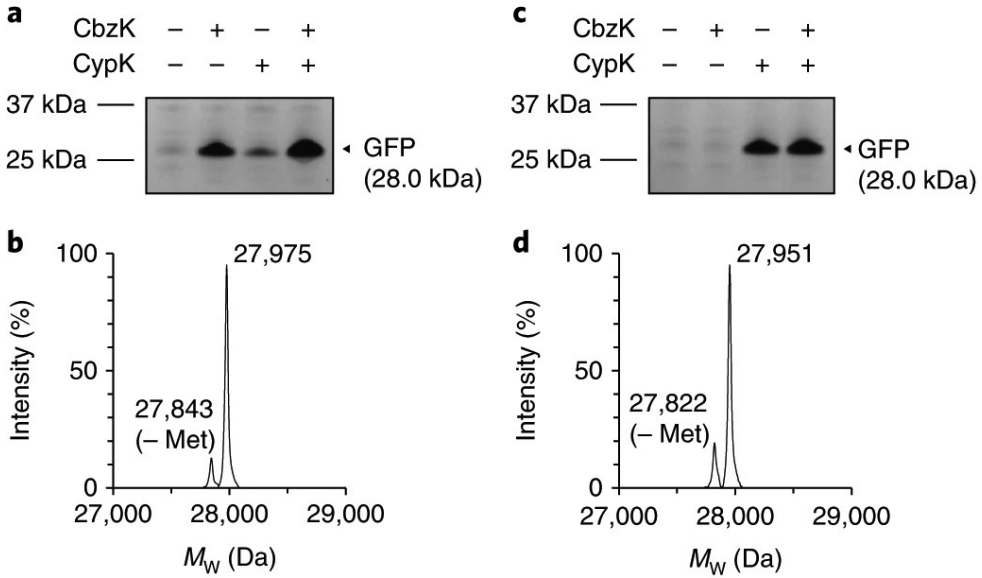
b, Libraries of *Ma*^{Pyl}tRNA created by randomizing or expanding the length of the variable loop to four, five or six randomized nucleotides.

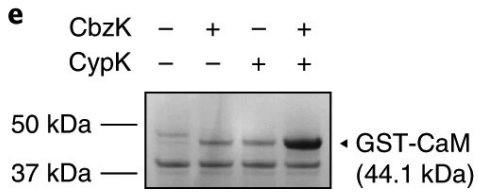
c, Variable loop sequences for the *Ma*^{Pyl}tRNA hits identified.

J.C.W. Willis, J. W. Chin *Nature Chem.*, **2018**, *10*, 831-837

Mutually orthogonal pyrrolysyl-tRNA synthetase(PylRS)/tRNA pairs







In vivo amber suppression activity assay using *E. coli* DH10B bearing pBAD *GFP(150TAG)His*$_6$ and the corresponding pKW PylRS/^{Pyl}tRNA$_{CUA}$ plasmid in the presence and absence of CbzK (**3**) and CypK (**4**) demonstrates the selective incorporation of CypK by *Mm*PylRS and the selective incorporation of CbzK by *Ma*PylRS-MutRS1

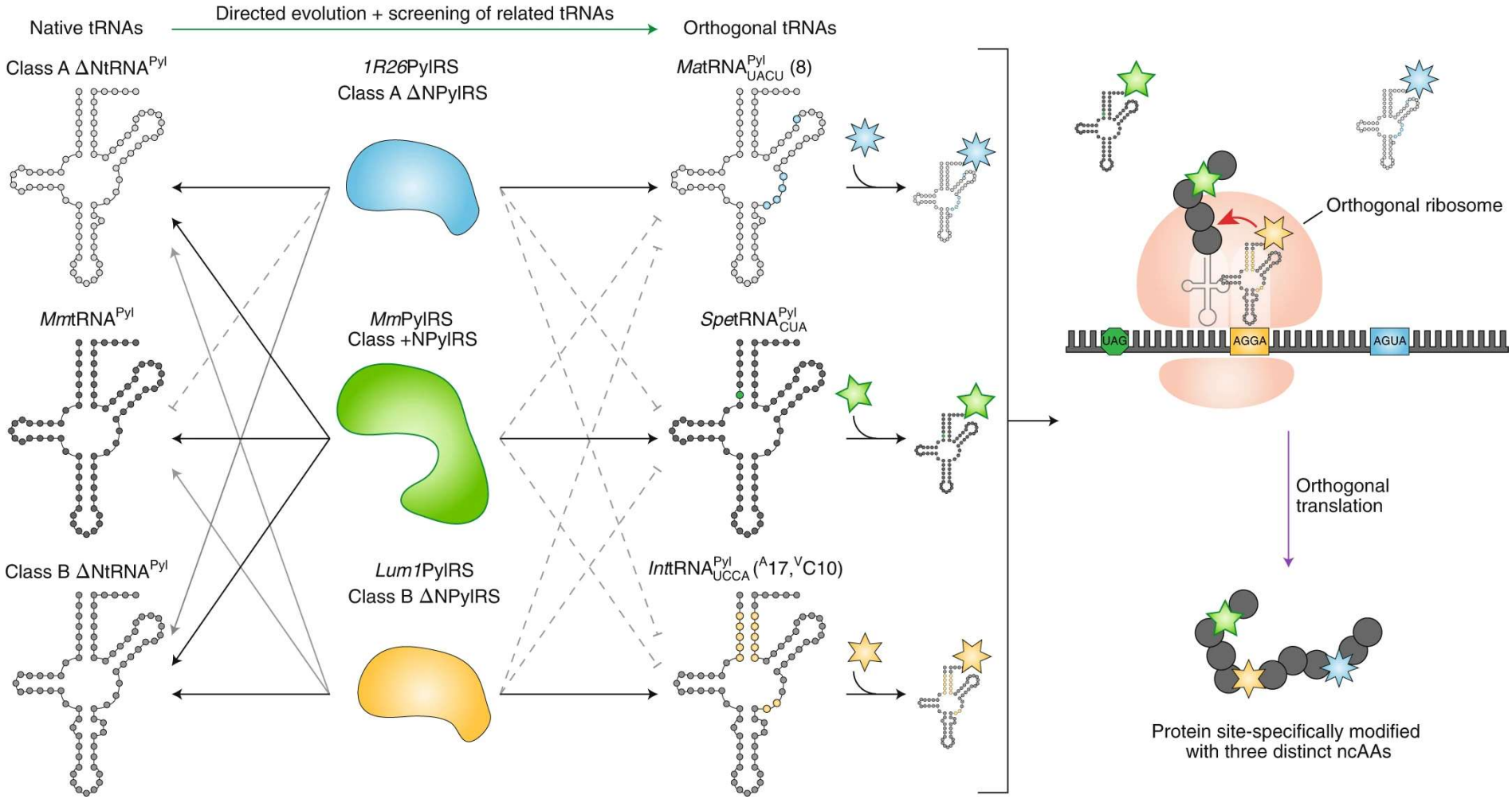
a/b GFPHis$_6$ purified from *E. coli* containing *GFP(150TAG)His*$_6$, *Mm*PylRS/*Mm*^{Pyl}tRNA$_{UCCU}$ and *Ma*PylRS-MutRS1/*Ma*^{Pyl}tRNA(6)$_{CUA}$

c/d GFPHis$_6$ purified from *E. coli* containing *GFP(150TAG)His*$_6$, *Mm*PylRS/*Mm*^{Pyl}tRNA$_{CUA}$ and *Ma*PylRS-MutRS1/*Ma*^{Pyl}tRNA(6)$_{UACU}$

e Glutathione-Stransferase-calmodulin (GST-CaM) purifications from *E. coli* containing ribo-Q1, *o*-GST-CaM(1TAG, 40AGGA), *Mm*PylRS/*Mm*^{Pyl}tRNA$_{UCCU}$ and *Ma*PylRS-MutRS1/*Ma*^{Pyl}tRNA(6)$_{CUA}$ grown in the presence and absence of the indicated ncAAs.

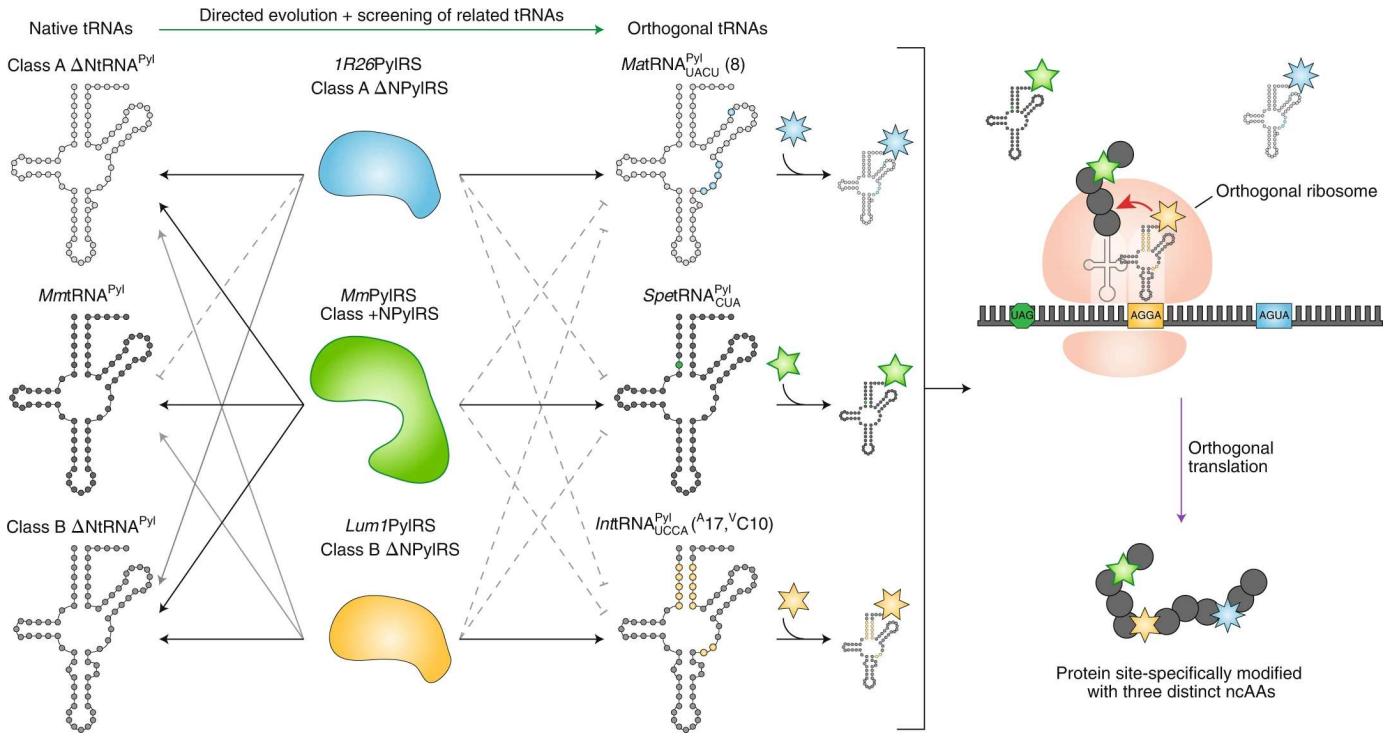
J.C.W. Willis, J. W. Chin *Nature Chem.*, **2018**, *10*, 831-837

Triply orthogonal PylRS/tRNA – three distinct noncanonical aminoacids



D. Dunkelmann, J.C.W. Willis, A.T. Beattie, J. W. Chin *Nature Chem.*, **2020**, *12*, 535-544

Triply orthogonal PylRS/tRNA – three distinct noncanonical aminoacids

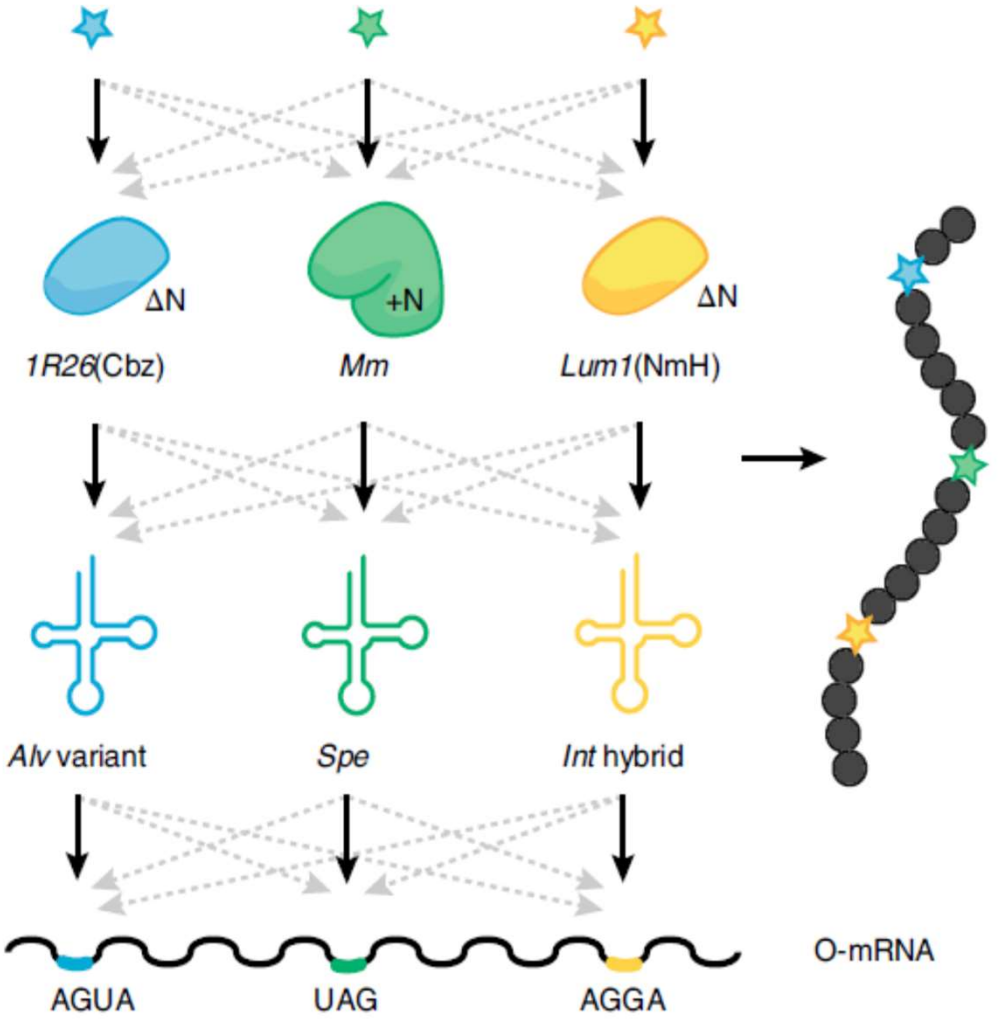


Combining three orthogonal tRNA^{Pyl}/PylRS based systems with orthogonal quadruplet decoding ribosomes enables the selective incorporation of three distinct ncAAs into one faithfully terminated protein. PylRS enzymes are classified into three separate classes: one containing an N-terminal domain that binds the variable loop of tRNA^{Pyl} (PylRS shown in green), and two functionally distinct ΔNPylRS classes (shown in blue and orange). By exploring natural diversity and combining this with directed evolution of the tRNA^{Pyl} variants, these three pairs were made orthogonal to each other (black arrows indicate amino acylation activity, grey arrows indicate residual amino acylation, blunt arrows indicate orthogonality).

When combined with specific active site mutations and orthogonal ribosomes these pairs enable the incorporation of three different ncAAs (indicated by blue, green and orange stars, respectively) into a protein by decoding an amber stop codon, as well as two different quadruplet codons

D. Dunkelmann, J.C.W. Willis, A.T. Beattie, J. W. Chin *Nature Chem.*, **2020**, *12*, 535-544

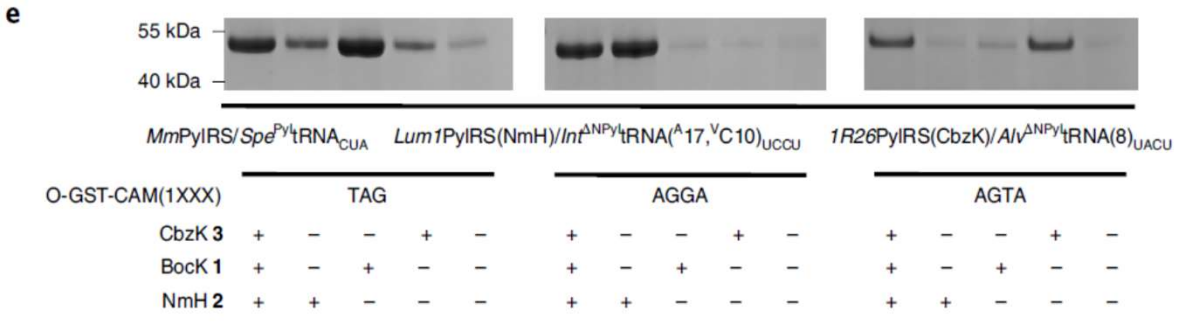
Triply orthogonal PylRS/tRNA – three distinct noncanonical aminoacids

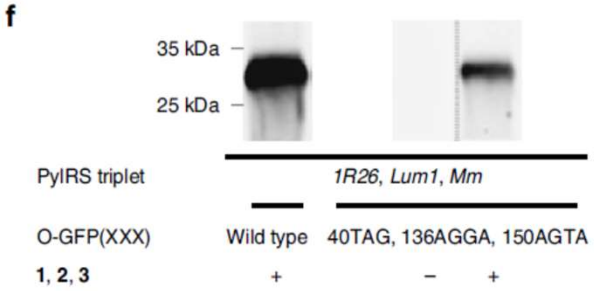


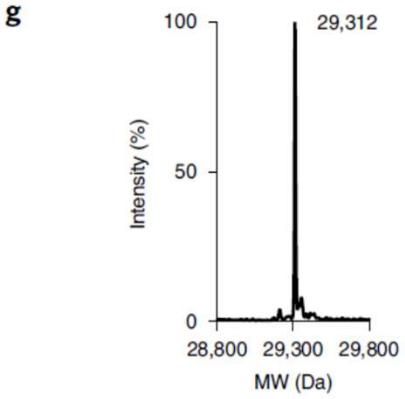
Schematic showing the nine specific functions (three specific ncAA recognitions, three specific aaRS/tRNA recognitions and three specific codon–anticodon interactions) that have been created to generate active pairs (black arrows). The schematic also shows the 18 potential interactions that have been eliminated (dashed grey arrows) to ensure orthogonality between the three pairs. The large number of potential interactions (between the endogenous amino acids, synthetases and tRNAs and the engineered triply orthogonal pairs) that have been eliminated to ensure orthogonality are omitted for clarity.

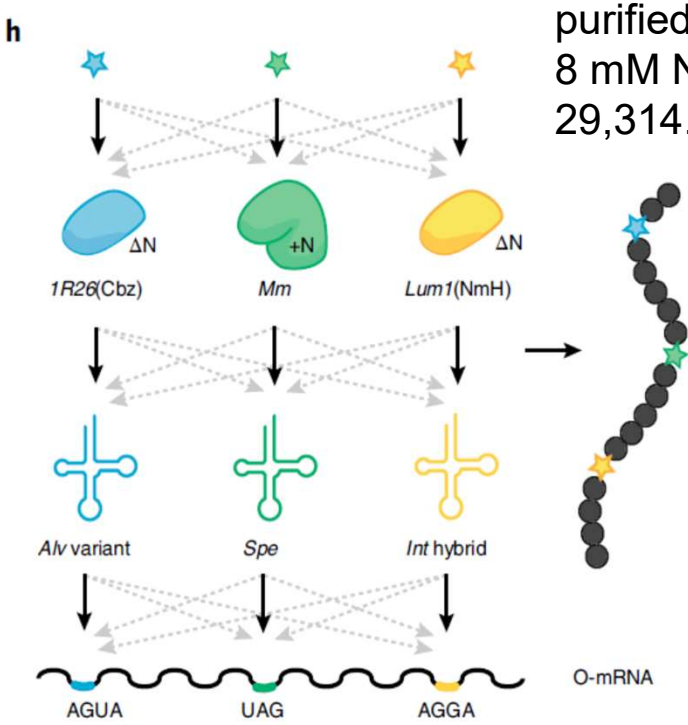
D. Dunkelmann, J.C.W. Willis, A.T. Beattie, J. W. Chin *Nature Chem.*, **2020**, *12*, 535-544

Triply orthogonal PylRS/tRNA – three distinct noncanonical aminoacids





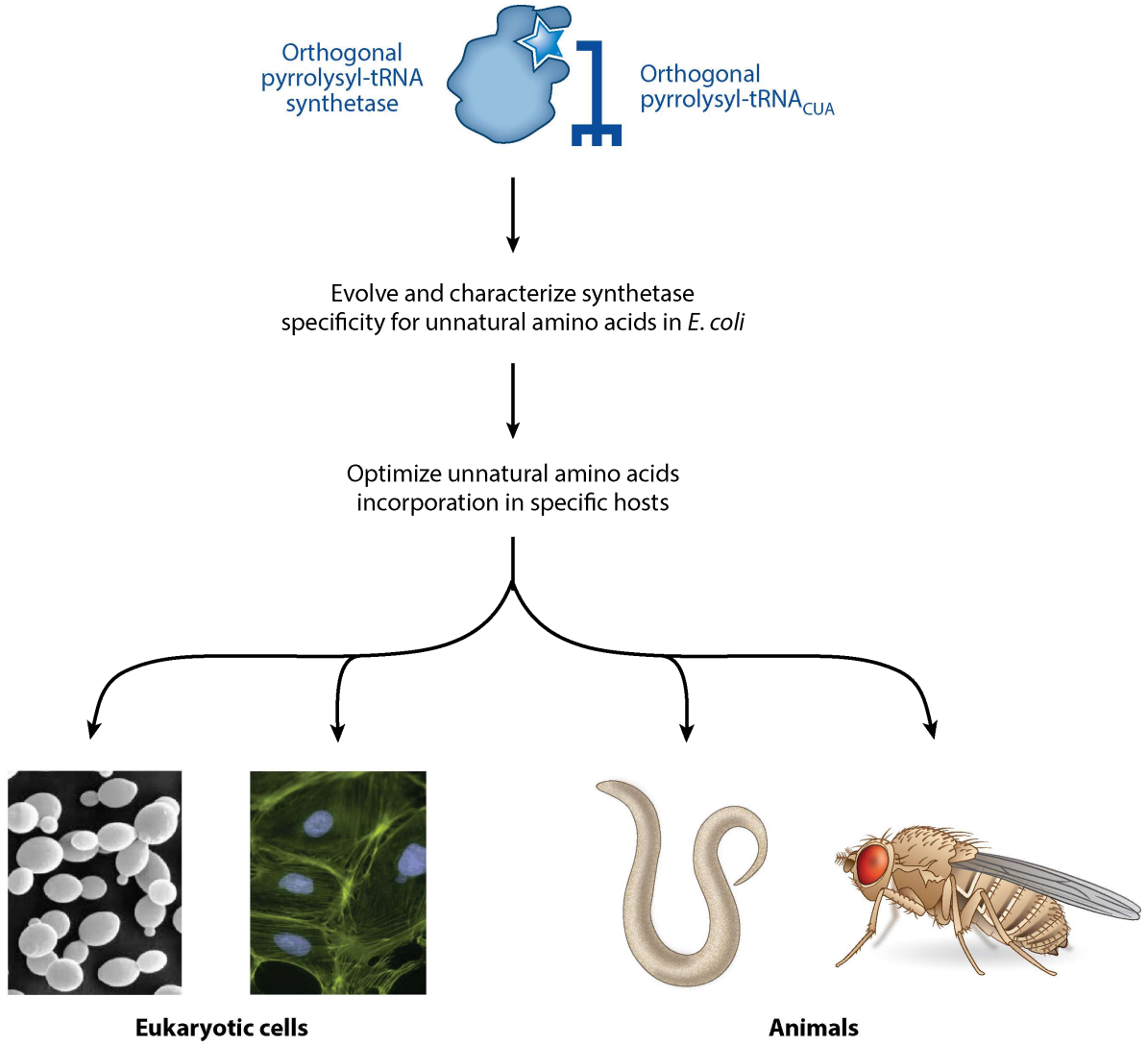




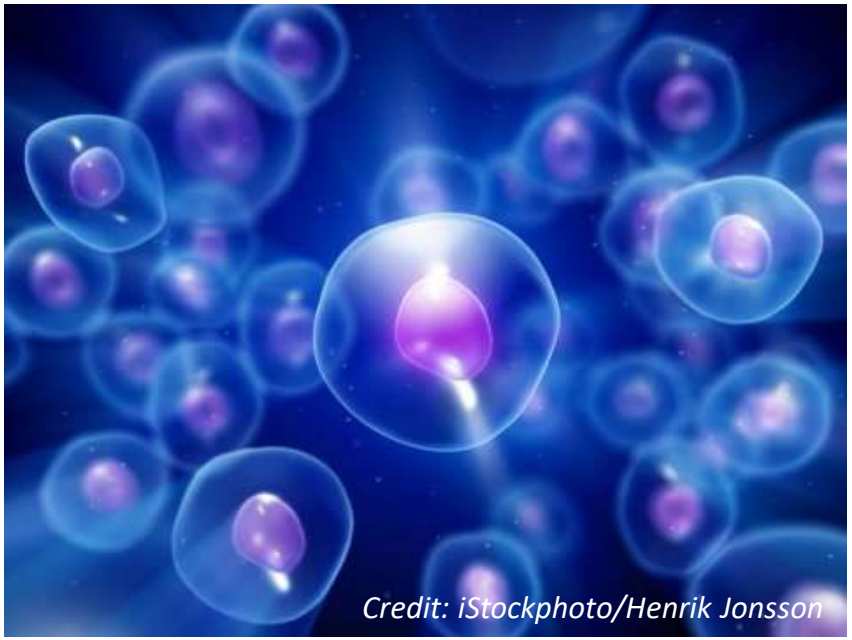
e, Divergence of the triply orthogonal pair derivatives that recognize distinct ncAAs to decode distinct codons.
f, Incorporation of three ncAAs into a single polypeptide.
g, Electrospray TOF-MS of purified $_{Strep}$GFPHis$_6$ purified from cells in the presence of 8 mM BocK, 8 mM NmH and 2 mM CbzK. mass predicted: 29,314.5 Da; mass found: 29,312 Da.

D. Dunkelmann, J.C.W. Willis, A.T. Beattie, J. W. Chin *Nature Chem.*, **2020**, *12*, 535-544

The pyrrolysyl-tRNA (transfer RNA) synthetase/tRNA$_{CUA}$ pair can be used to site-specifically encode the incorporation of unnatural amino acids into proteins in cells and animals.

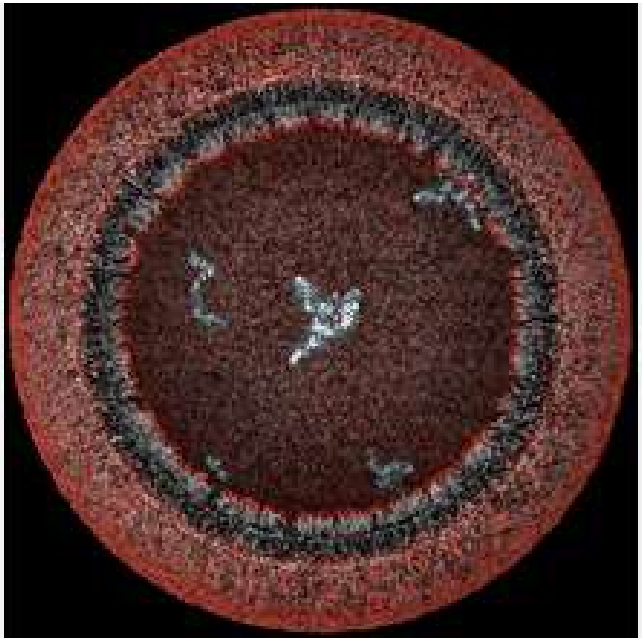


Encapsulation – essential for life



Membrane compartments

Assembly of amphiphilic monomers into protocellular compartments



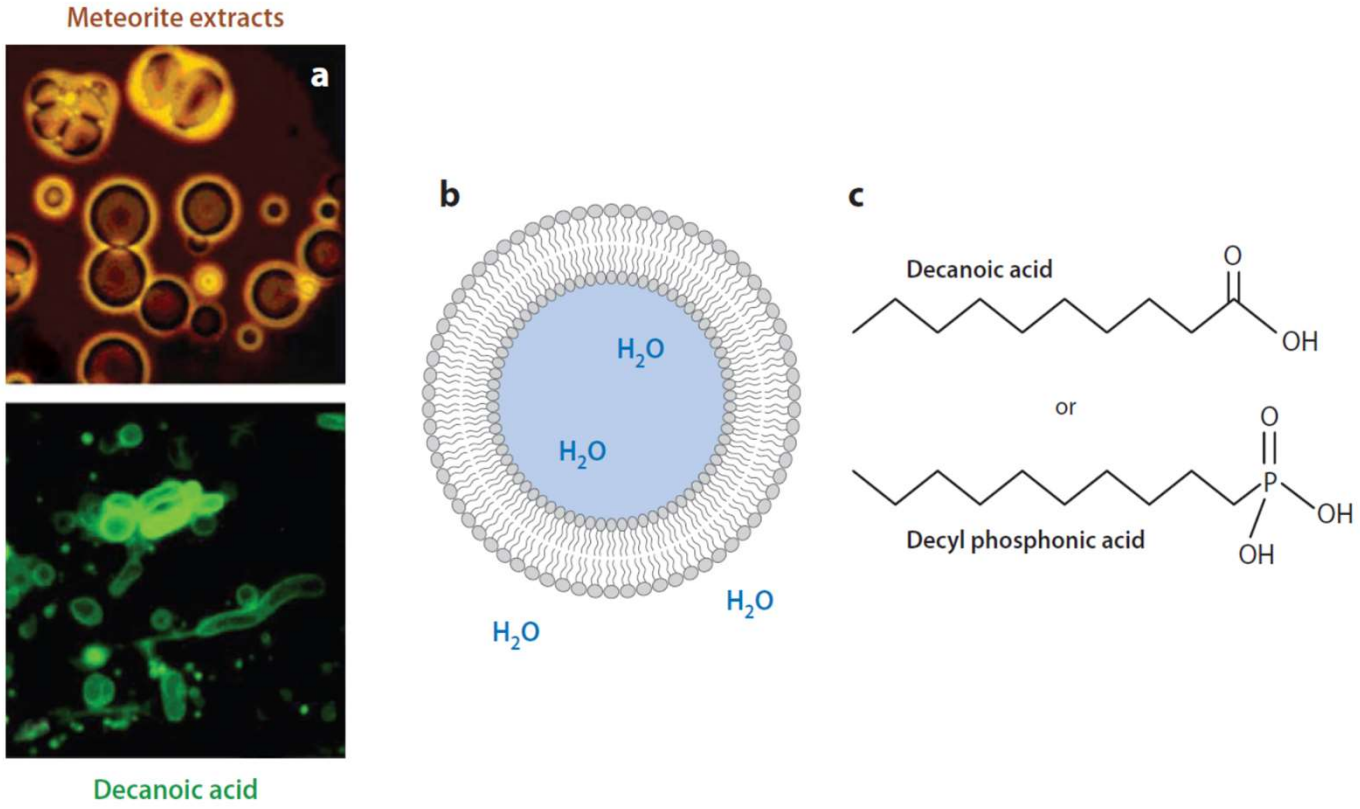
Credit: *Janet Iwasa*

A three-dimensional view of a model protocell (a primitive cell) approximately 100 nanometers in diameter.

The protocell's fatty acid membrane allows nutrients and DNA building blocks to enter the cell and participate in non-enzymatic copying of the cell's DNA. The newly formed strands of DNA remain in the protocell

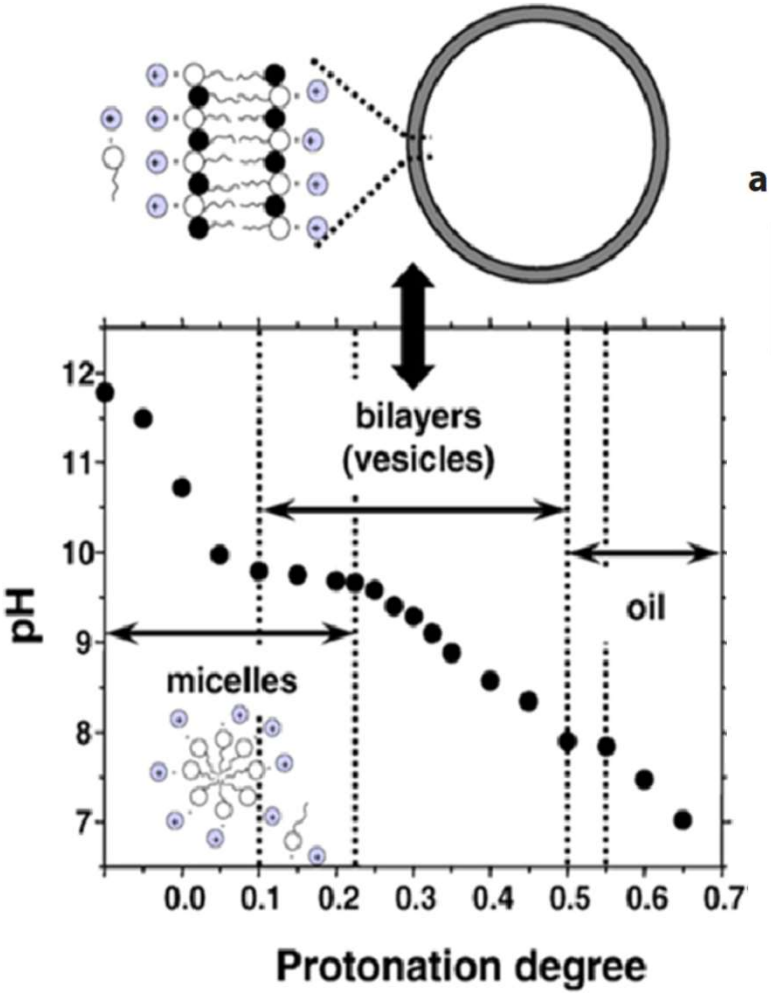
Encapsulation – essential for life

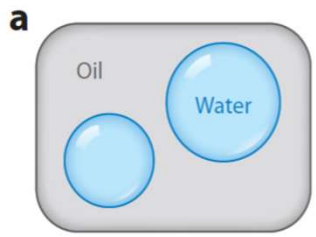
Fatty acids have been found in meteorites – plausible prebiotic synthesis pathways existed in the early Solar System

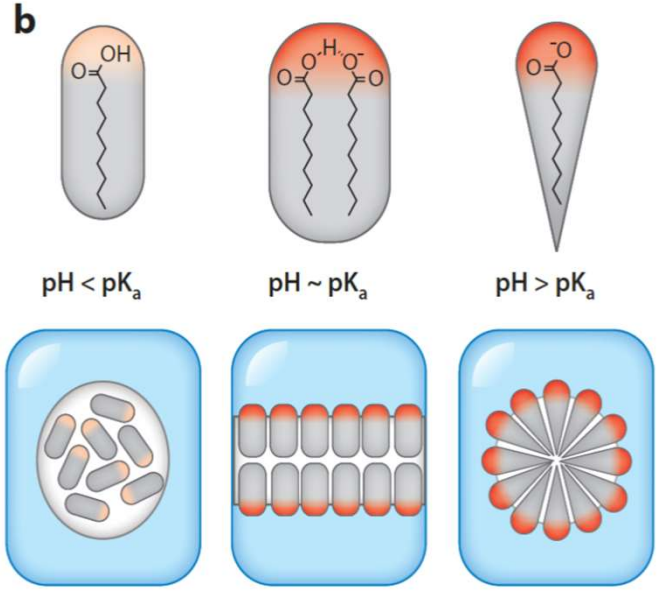


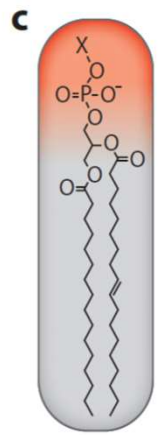
Extracts of meteorites containing these compounds spontaneously form vescicles when hydrated

pH-dependent phase behavior of fatty acids in water



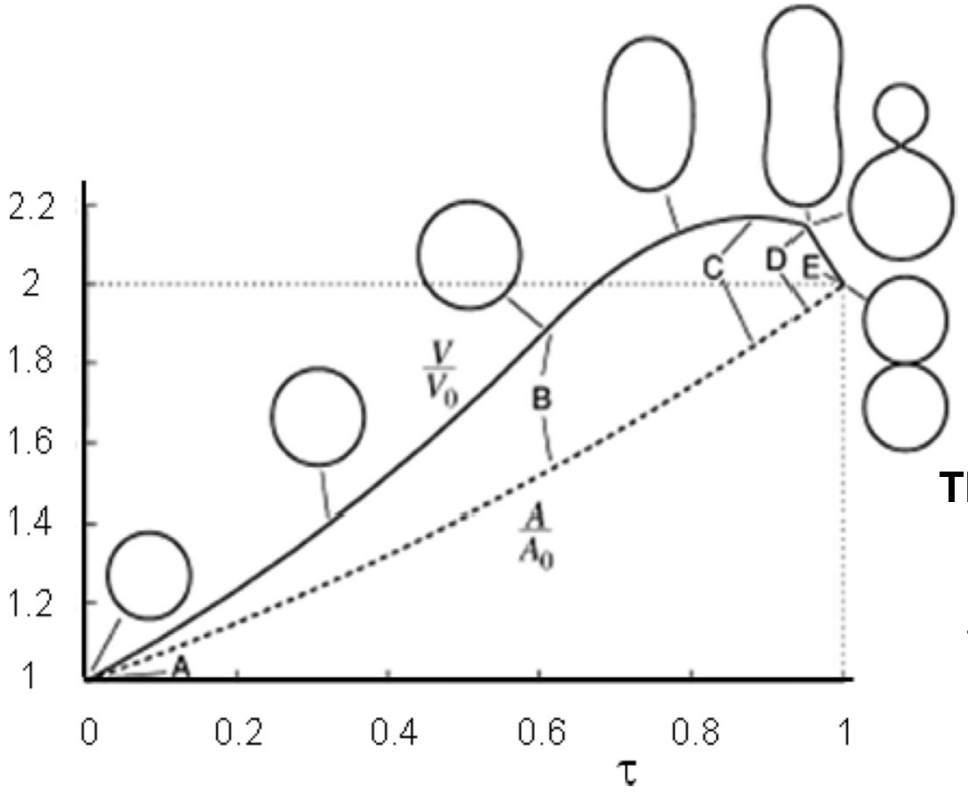


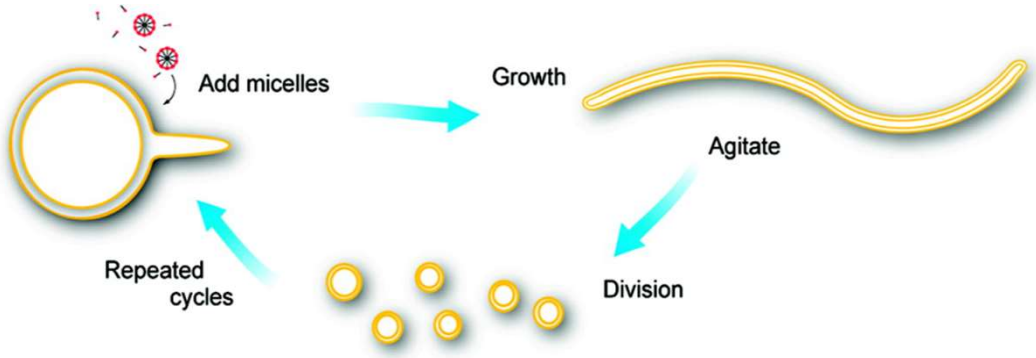




80 mM oleic acid/ sodium oleate in water

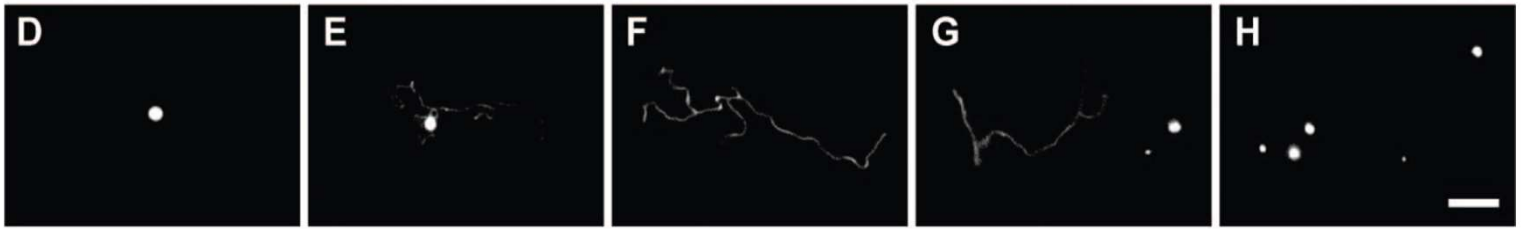
Growth and division of vesicles





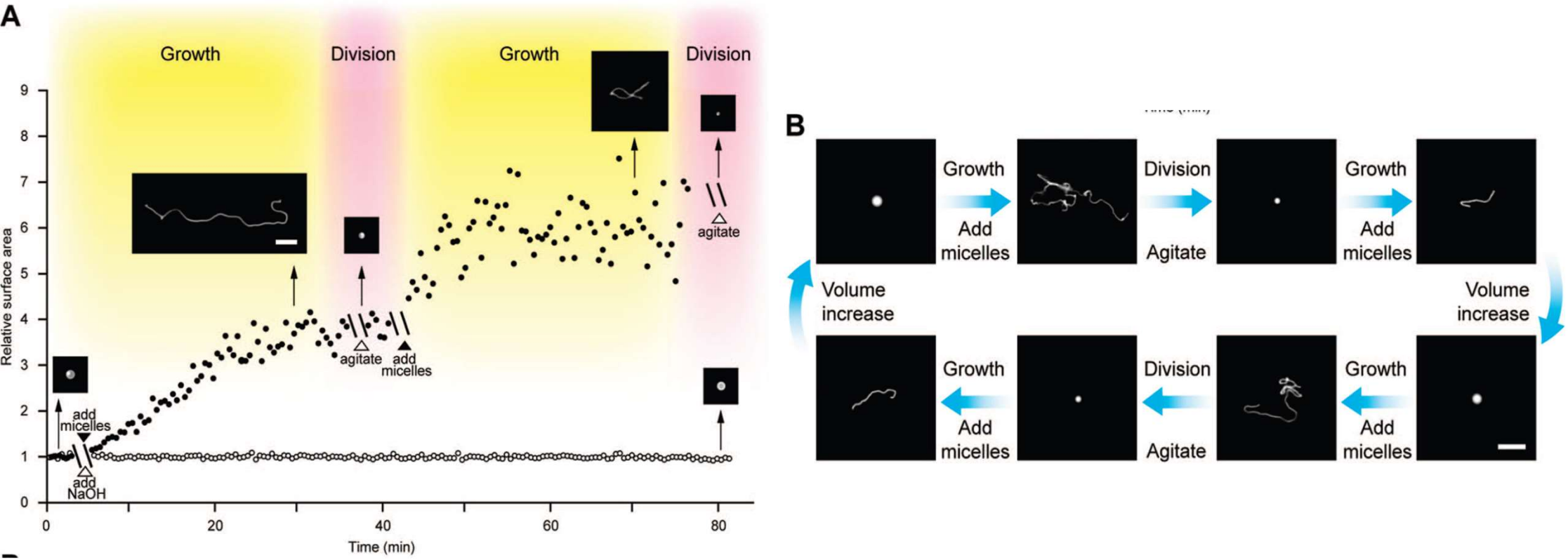
The growth of large multilamellar fatty acid vesicles fed with fatty acid micelles:

when solute permeation across the membranes is slow, the transient imbalance between surface area and volume growth causes formation of long thread-like vesicles. Modest shear forces are then sufficient to divide them into multiple daughter vesicles without loss of internal contents.



Ting F. Zhu, and Jack W. Szostak *J. Am. Chem. Soc.*, **2009**, 131 (15), 5705-5713

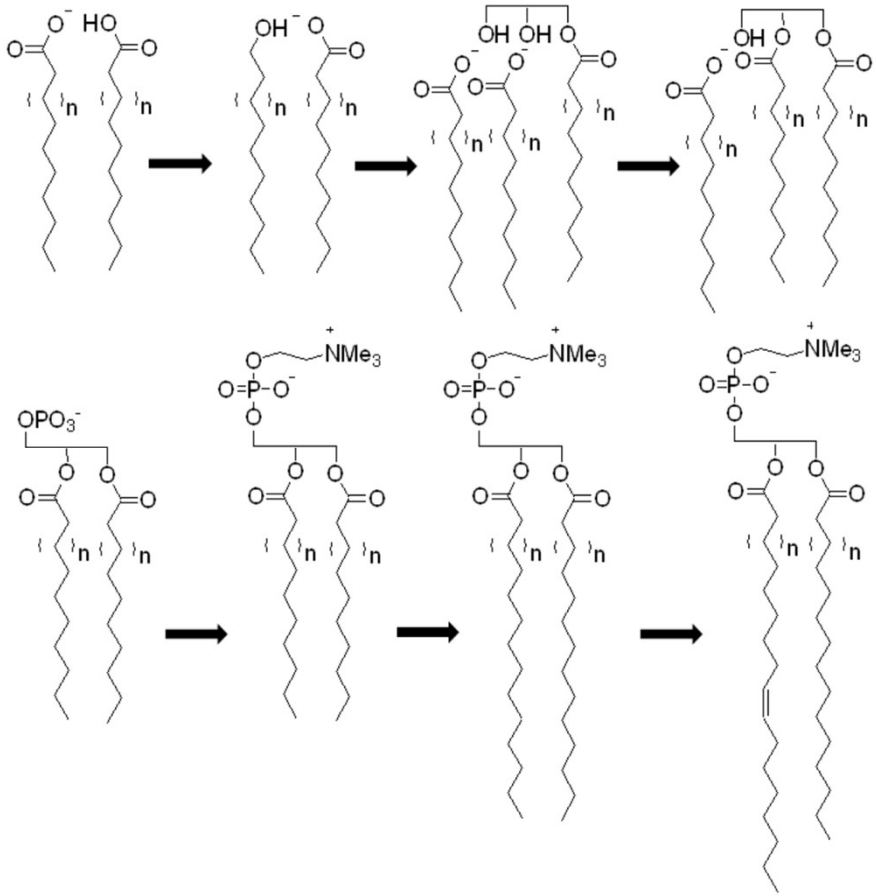
Coupled growth and division of model protocell membranes



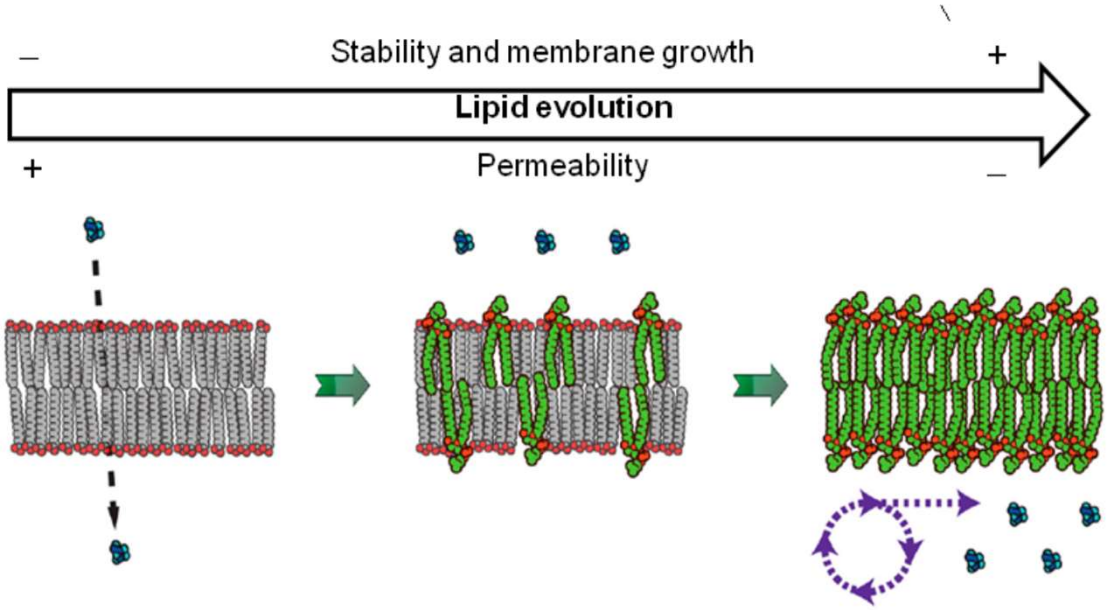
Cycles of vesicle growth and division. (A) Relative surface area after two cycles of addition of 5 equiv of oleate micelles (solid circles) or 5 equiv of NaOH (open circles) to oleate vesicles, each followed by agitation. Inset micrographs show vesicle shapes at indicated times. Scale bar, 10 μm. (B) Vesicle shapes during cycles of growth and division in a model prebiotic buffer (0.2 M Na-glycine, pH 8.5, ~1 mM initial oleic acid, vesicles contain 10 mM HPTS for fluorescence imaging). Scale bar, 20 μm.

Ting F. Zhu, and Jack W. Szostak *J. Am. Chem. Soc.*, **2009**, 131 (15), 5705-5713

Scheme of the membrane evolution

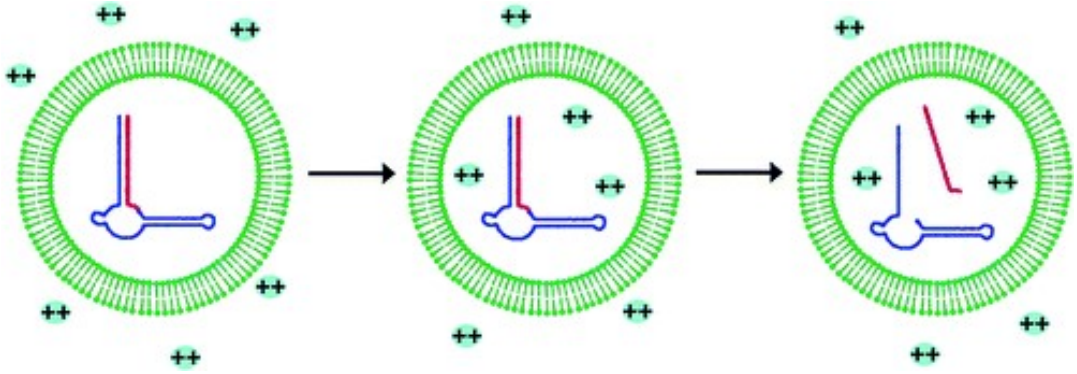


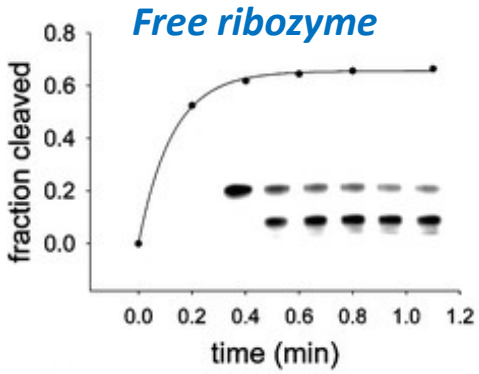
Chemical evolution of membrane components

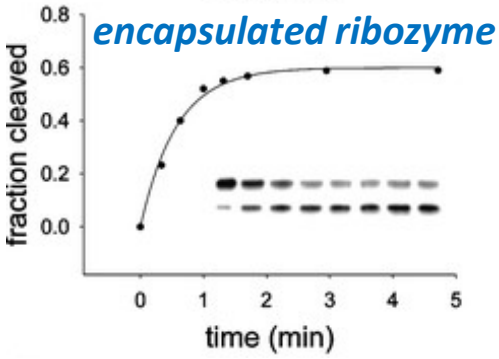


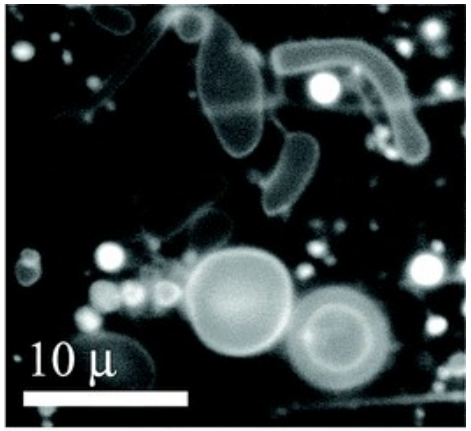
More complex components lead to slower amphiphile desorption and thus faster growth of the protocell. Decreasing permeability is a selective pressure for the emergence of internalized metabolic and transport machinery in the system

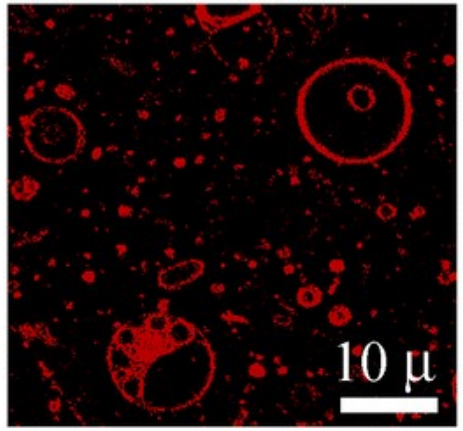
RNA Catalysis in Model Protocell Vesicles











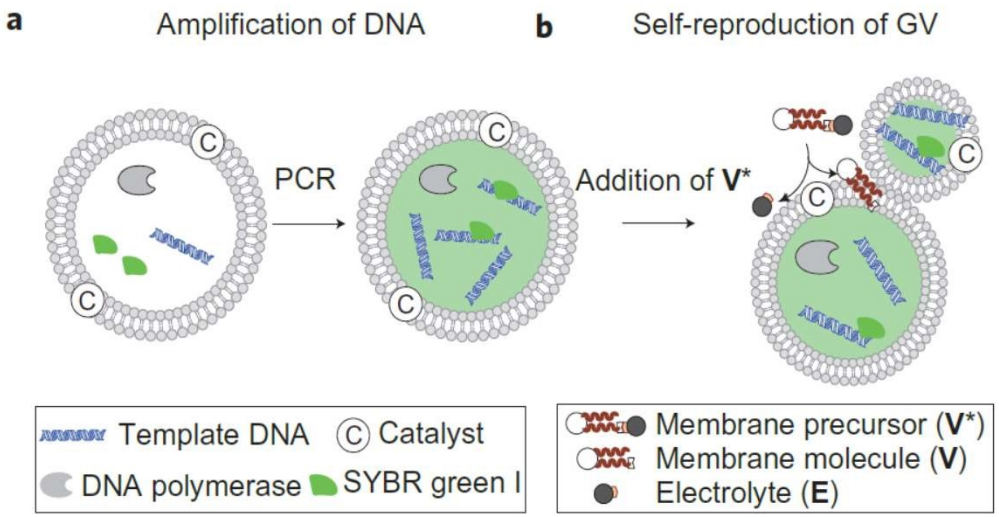
A mixture of myristoleic acid and its glycerol monoester forms vesicles that were Mg^{2+}-tolerant. Mg^{2+} cations can permeate the membrane and equilibrate within a few minutes.

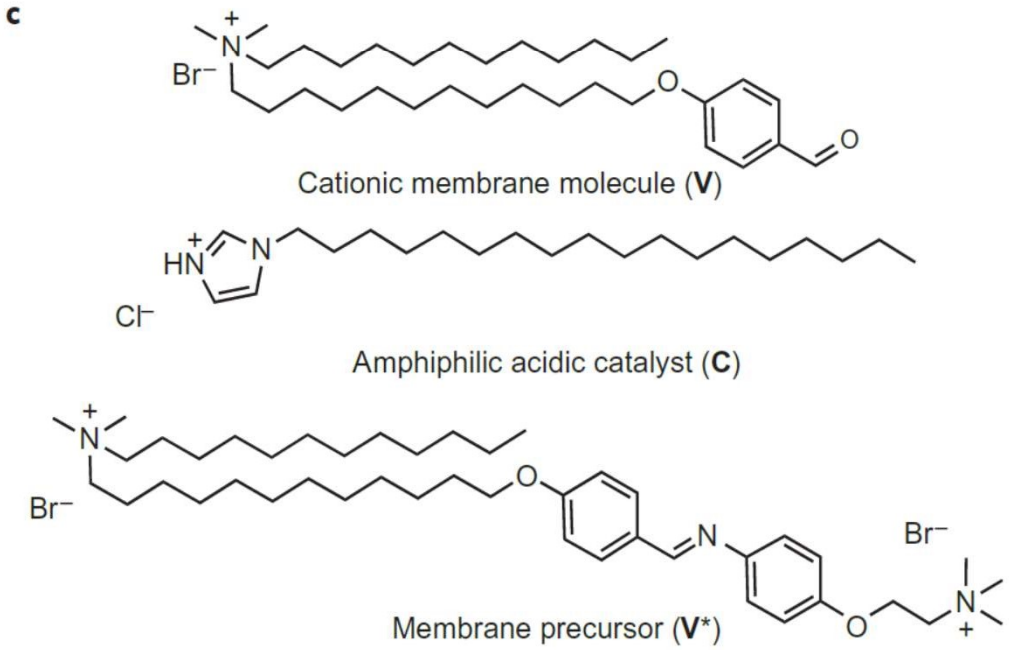
In vesicles encapsulating a hammerhead ribozyme, the addition of external Mg^{2+} led to the activation and self-cleavage of the ribozyme molecules. These vesicles can grow upon addition of micelles. It demonstrates that membranes made from simple amphiphiles can form vesicles that are stable enough to retain encapsulated RNAs in the presence of divalent cations.

Fluorescence microscopy of 2:1:0.3 MA:GMM:dodecane vesicles containing hammerhead ribozyme in the presence of 3 mM $MgCl_2$,

I. A. Chen, K. Salehi-Ashtiani, and J. W. Szostak *J. Am. Chem. Soc.*, **2009**, 127, 13213-13219

Self-reproduction of giant vesicles combined with the amplification of DNA





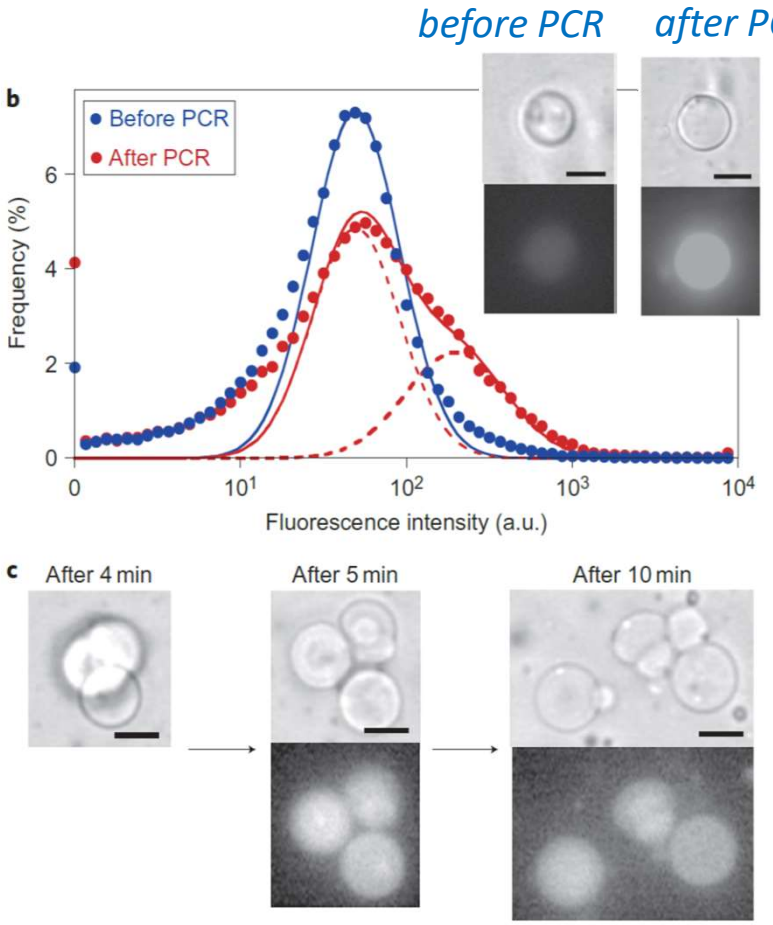
a, ***Amplification of DNA within a GV***. An aqueous dispersion of GVs containing PCR reagents was prepared using a film-swelling method with a buffered solution containing template DNA, primers, fluorescent tag SYBR Green I, deoxynucleoside triphosphates, DNA polymerase and Mg^{2+}.

b, ***Vesicular self-reproduction induced by adding membrane precursor V*.*** Addition of V* produces membrane molecules and electrolytes through hydrolysis assisted by an amphiphilic catalyst. Adhesion of the amplified DNA to the inner leaflet accelerates vesicular growth and division.

c, Chemical structures of membrane molecule V, amphiphile catalyst C and membrane precursor V*.

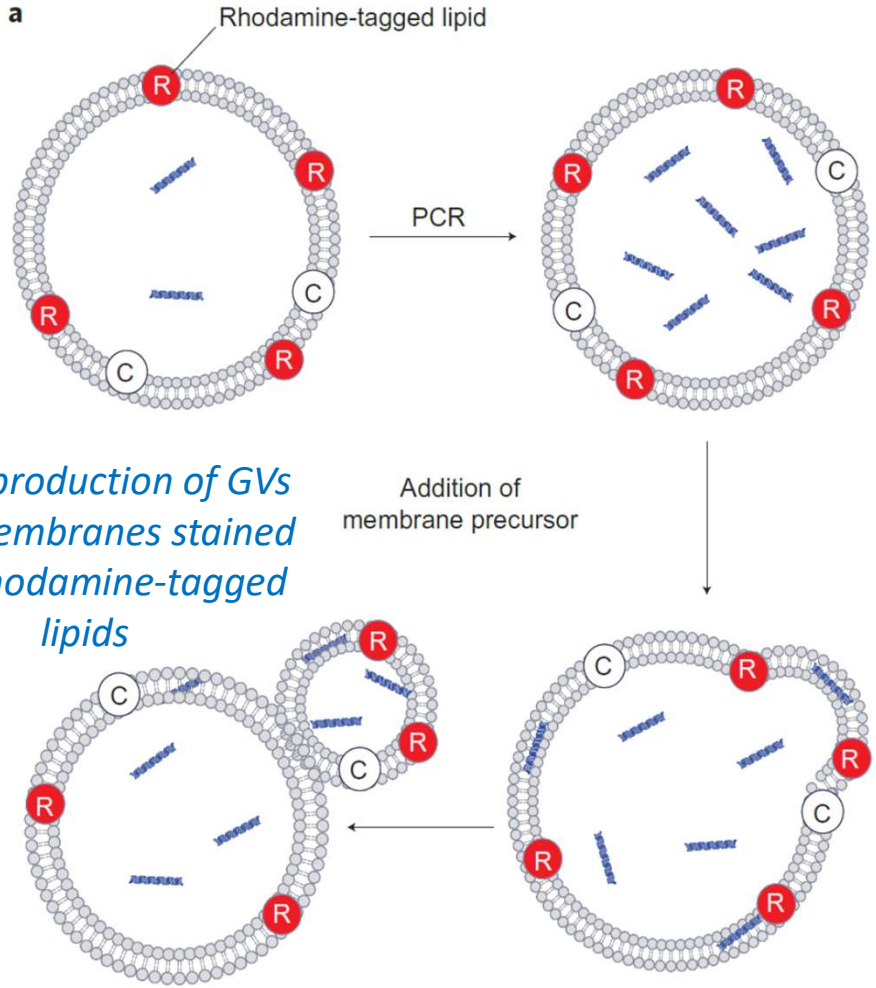
K. Kurihara *et al., Nat. Chem.*, **2011**, 3, 775-781

Self-reproduction of giant vesicles combined with the amplification of DNA



Real-time observation of morphological changes of DNA-amplified GVs after addition of V* .

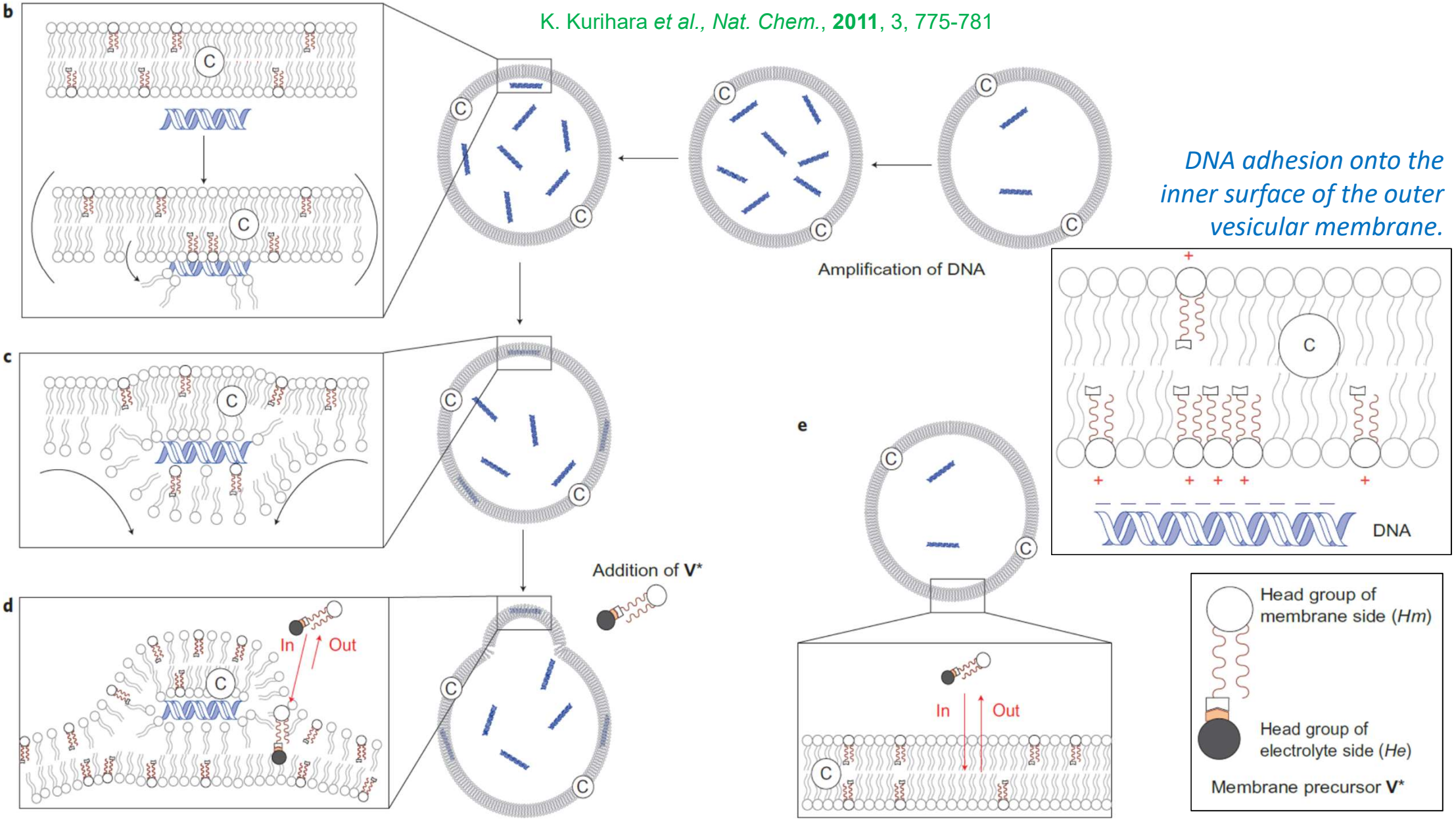
Original GVs began to grow and divide 4 min after adding V. Complete division into four GVs occurred at 5.5 min, and separation occurred at 7 min. Scale bars, 10 μm.*



Self-reproduction of GVs with membranes stained with rhodamine-tagged lipids

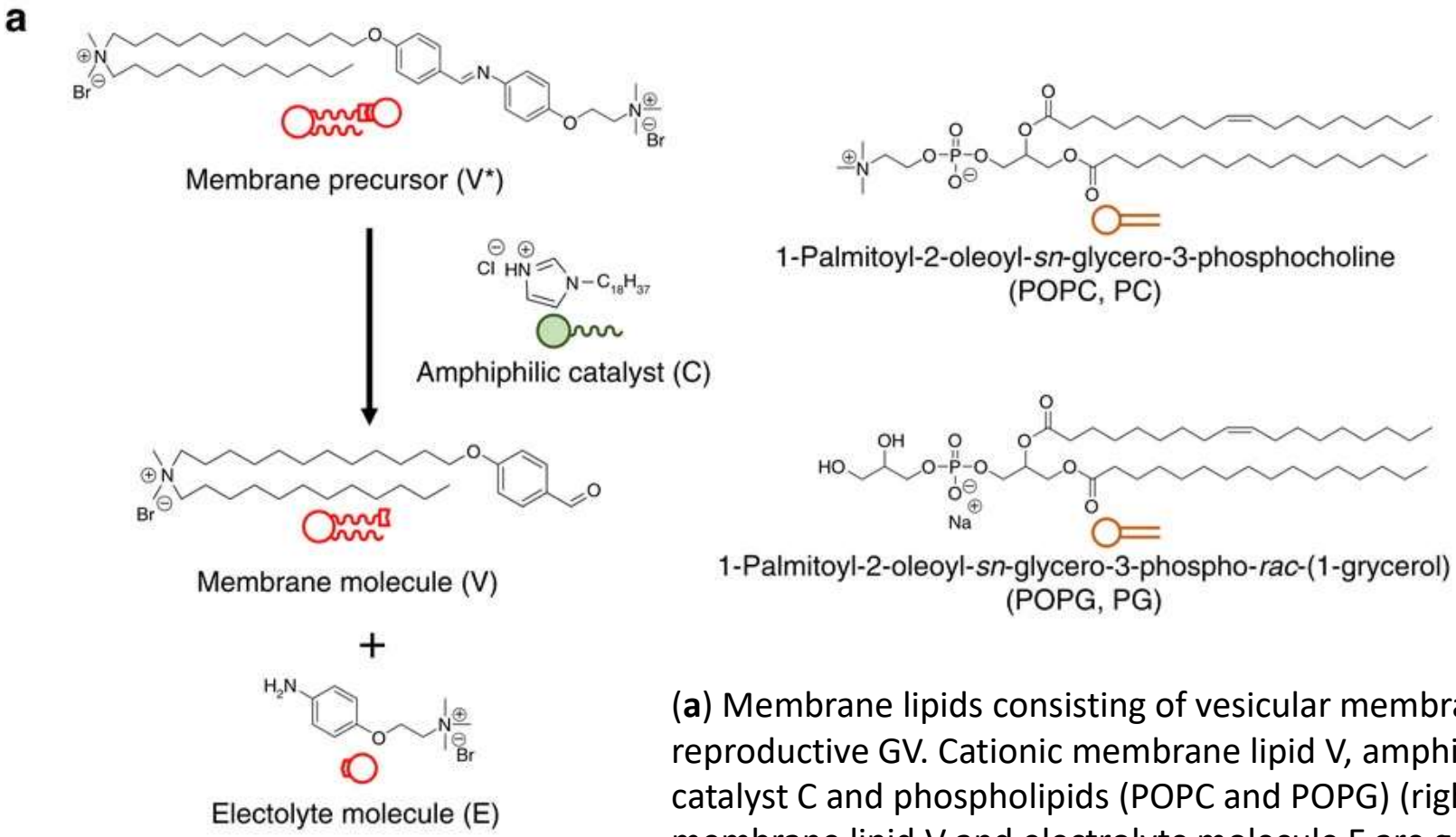
K. Kurihara *et al., Nat. Chem.*, **2011**, 3, 775-781

K. Kurihara *et al., Nat. Chem.*, **2011**, 3, 775-781



DNA adhesion onto the inner surface of the outer vesicular membrane.

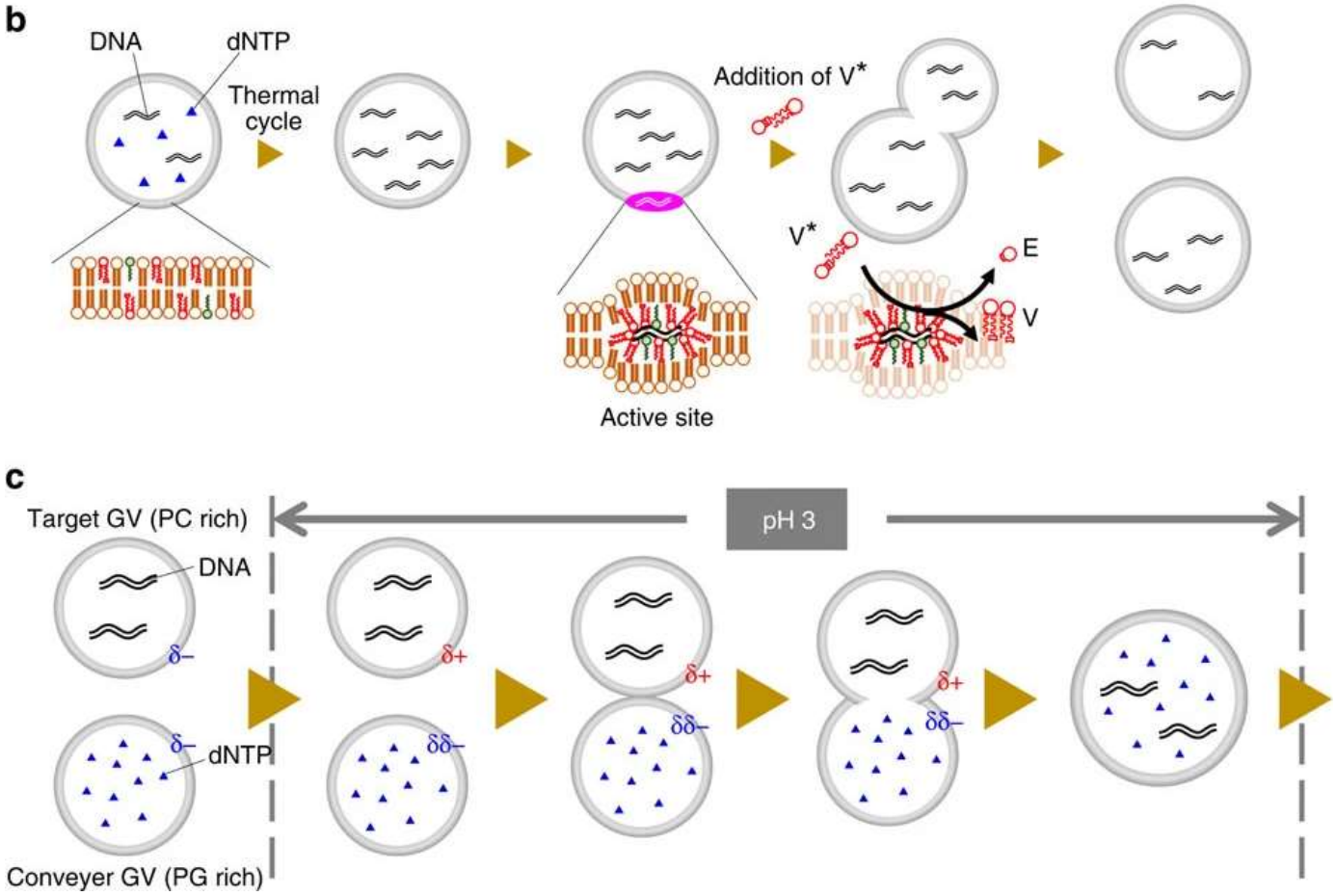
A protocell model with a primitive cell cycle



(**a**) Membrane lipids consisting of vesicular membrane of self-reproductive GV. Cationic membrane lipid V, amphiphilic catalyst C and phospholipids (POPC and POPG) (right). The membrane lipid V and electrolyte molecule E are generated through the hydrolysis of the membrane lipid precursor V*.

K. Kurihara *et al., Nat. Commun.*, **2016**, 6:8352 | DOI: 10.1038/ncomms9352

A protocell model with a primitive cell cycle

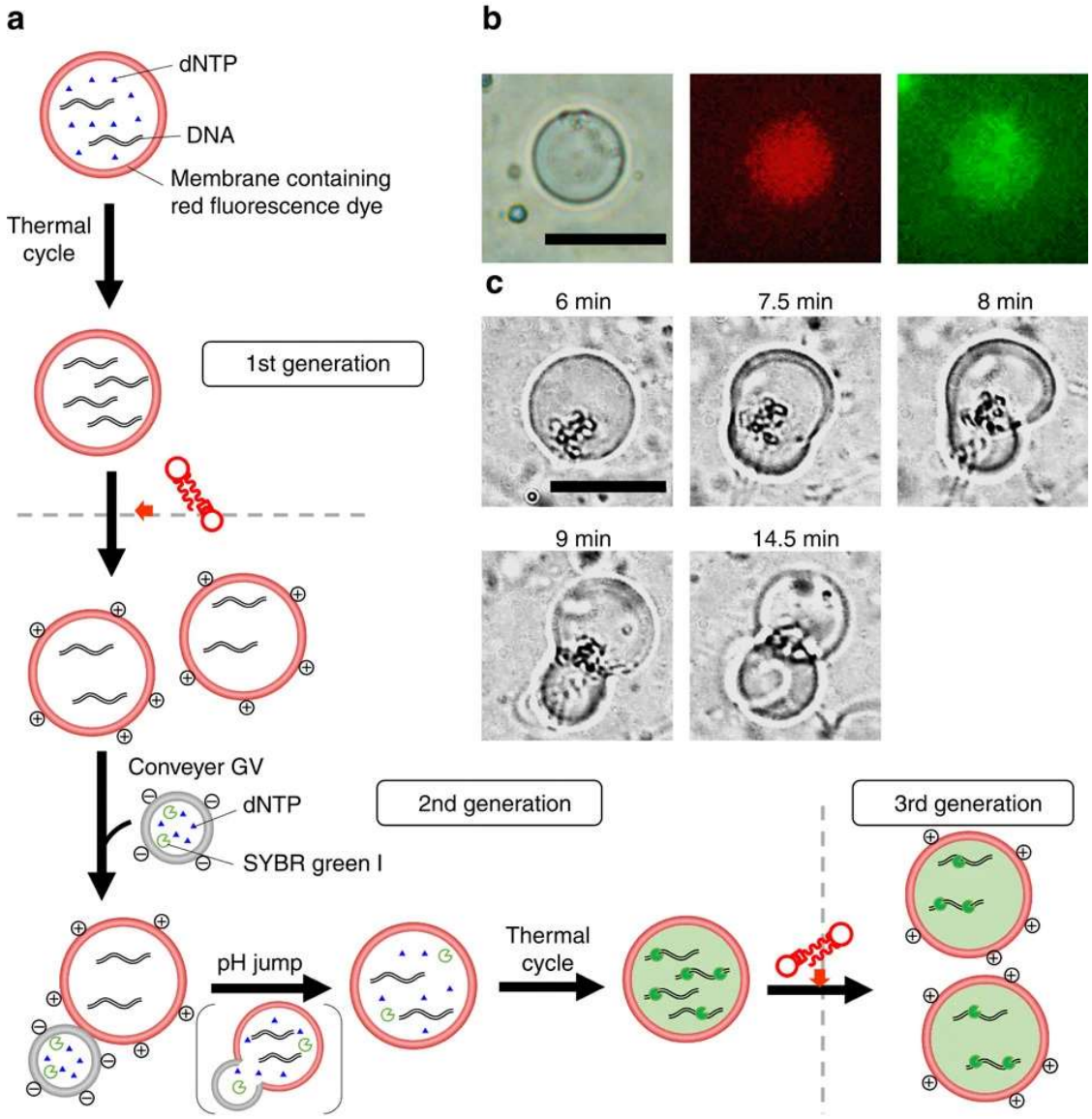


(**b**) The production of cationic membrane lipid V from its precursor V*. The cationic membrane V is produced together with the electrolyte E at an active site comprised of amplified DNA and amphiphilic catalyst C in the giant vesicular membrane.

(**c**) pH lowering induced adhesion and fusion between the target GV and the conveyer GV. The surface charge of the target GV changes to cationic due to the protonation of the POPC as well as the increase of the cationic membrane lipid V from its precursor, and the target GV adheres to the conveyer GV with a negative surface charge at pH=3. These two types of GVs fuse, and the transport of dNTP from the conveyer GV to the target GV proceeds.

K. Kurihara *et al., Nat. Commun.*, **2016**, 6:8352 | DOI: 10.1038/ncomms9352

A protocell model with a primitive cell cycle



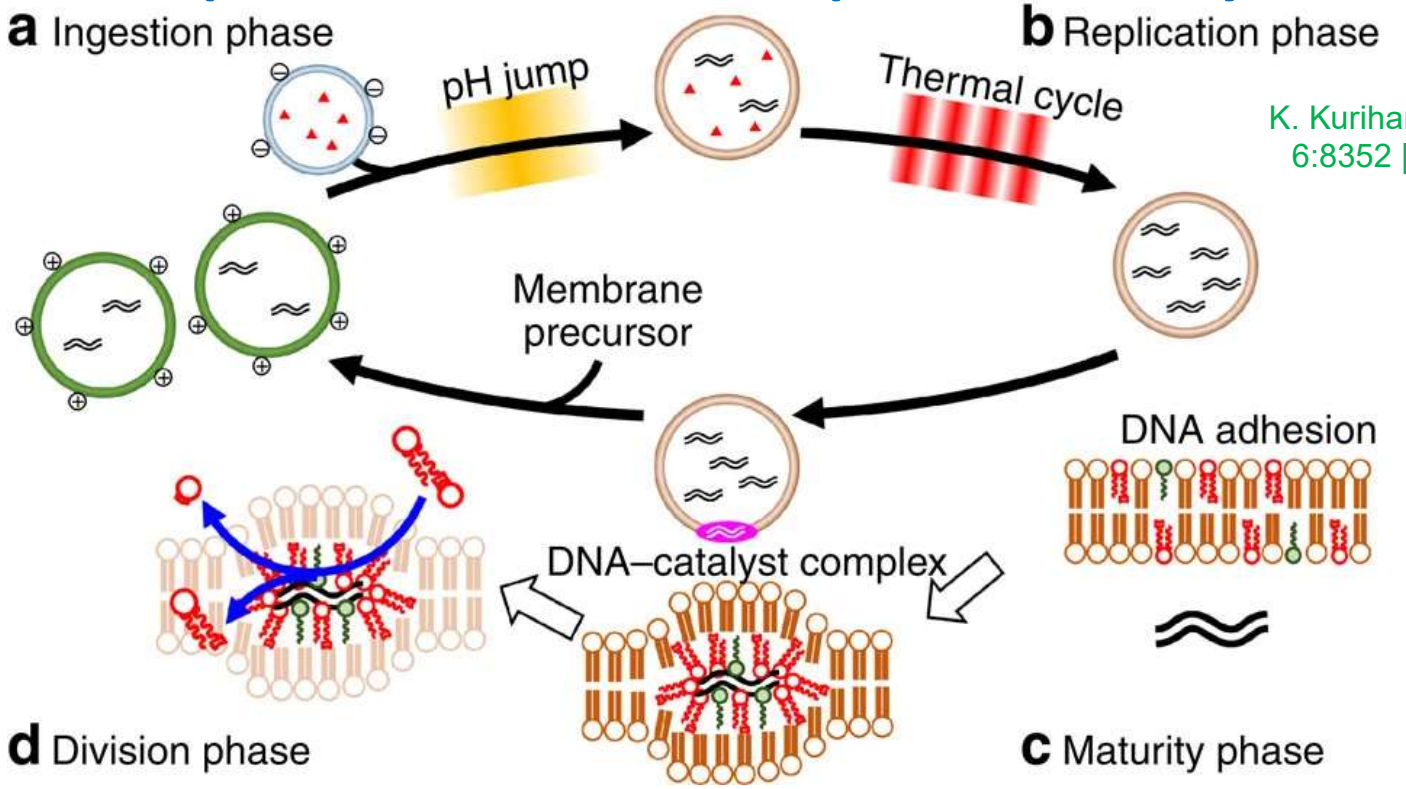
(**a**) Self-proliferation of GV-based model protocell from 1st generation to 3rd generation. DNA amplification in mother GV was followed by the first division to give rise to daughter GVs. Ingestion of dNTP in conveyer GV by daughter GVs and DNA amplification in daughter GV led the second division to give granddaughter GVs (bottom).

(**b**) Differential interface contrast microscope image of DNA-amplified daughter GV (left). Fluorescence microscope images of the red fluorescence emitted from the vesicular membrane (center) and the green fluorescence from inside the daughter GV (right). Scale bar, 10 µm.

(**c**) Division of the daughter GV to afford granddaughter GVs by the addition of precursor V* of the membrane lipid. Scale bar, 20 µm.

K. Kurihara *et al., Nat. Commun.*, **2016**, 6:8352 | DOI: 10.1038/ncomms9352

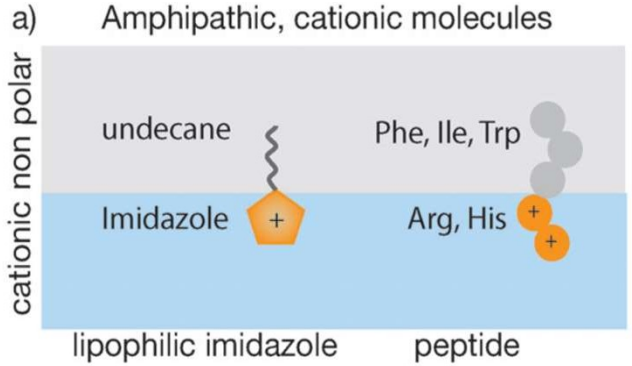
A protocell model with a primitive cell cycle

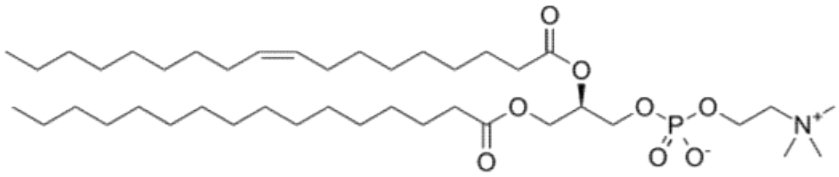


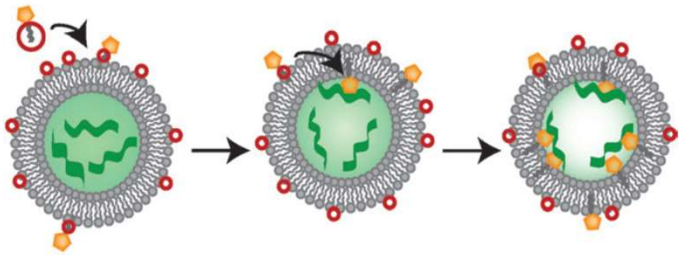
K. Kurihara *et al., Nat. Commun.*, **2016**, 6:8352 | DOI: 10.1038/ncomms9352

(**a**) In the ingestion phase, the GV of the next generation ingests substrates through vesicular fusion with conveyer GV containing dNTP, triggered by a pH jump. (**b**) In the replication phase, the replication of DNA in the next-generation GV proceeds using ingested dNTP. (**c**) In the maturity phase, the catalytic ability of the vesicular membrane matures in a sense that a complex between amplified DNA, amphiphilic catalyst C and cationic lipids V intrudes into the vesicular membrane, forming an active site for converting membrane precursor V* to lipid membrane V. (**d**) In the division phase, the self-proliferative GV grows and exhibits a budding deformation and an equivolume division when the precursor V* of the membrane lipid is added to the exterior of GVs.

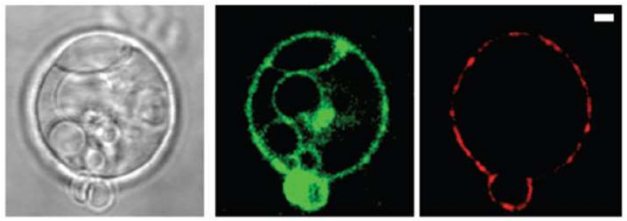
Noncovalent nucleotide association with membranes

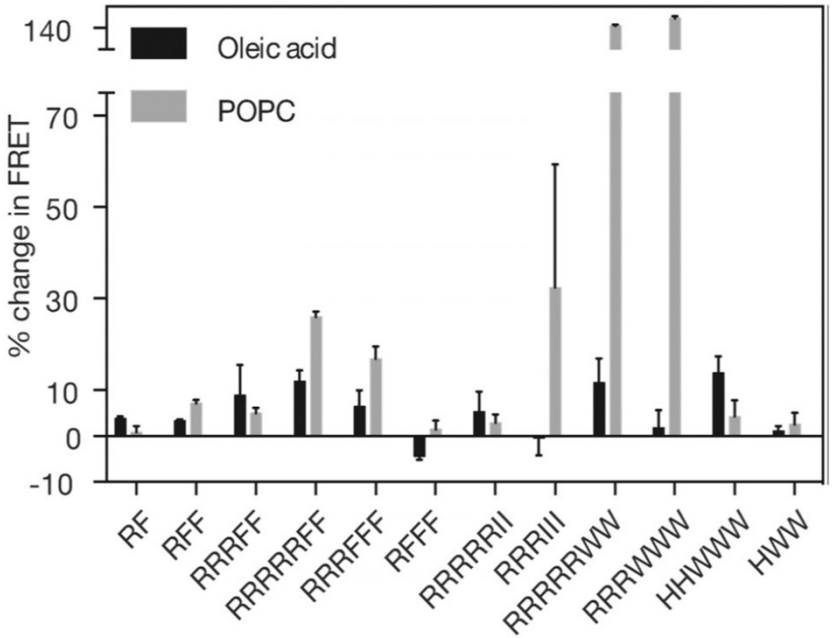






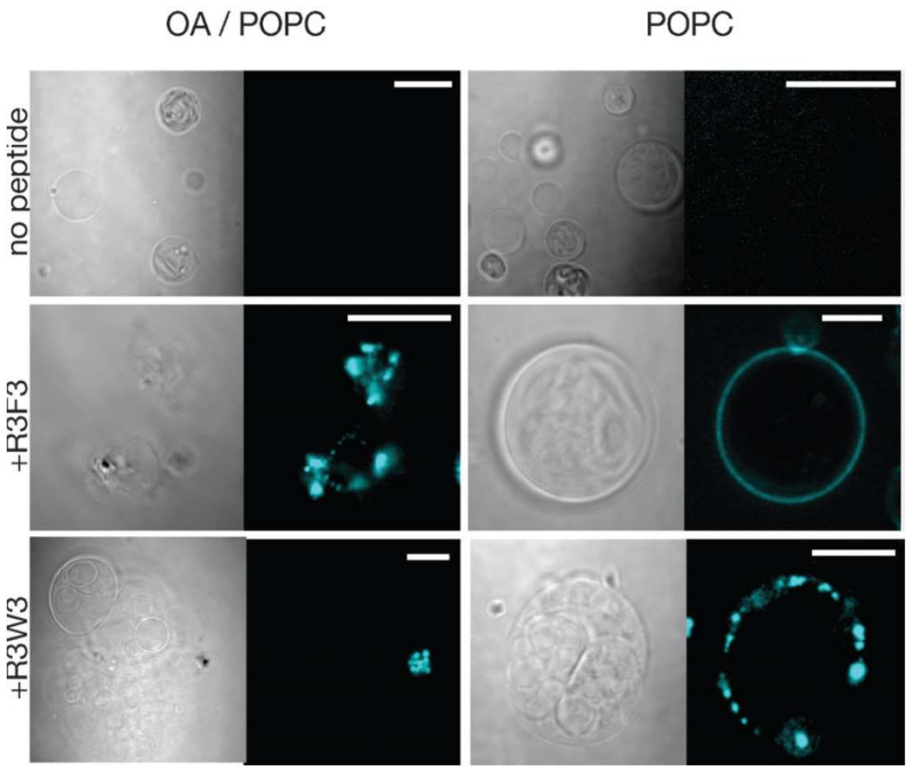
+POPC SUVs with undecylimidazole

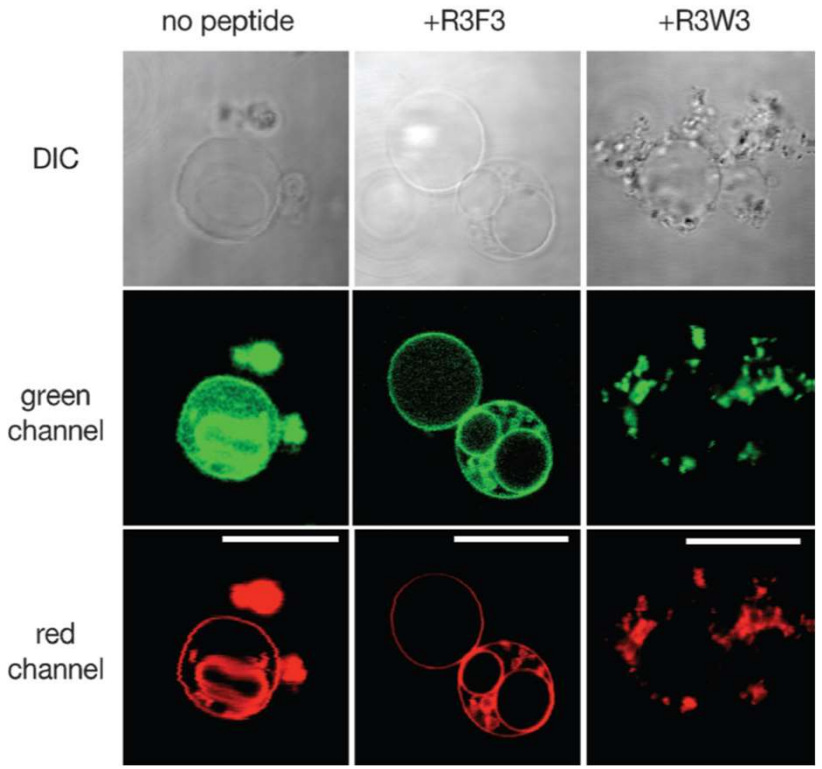




Neha P. Kamat, Sylvia Tobe, Ian T. Hill, and Jack W. Szostak *Angew. Chem. Int. Ed.* **2015**, *54*, 11735 –11739

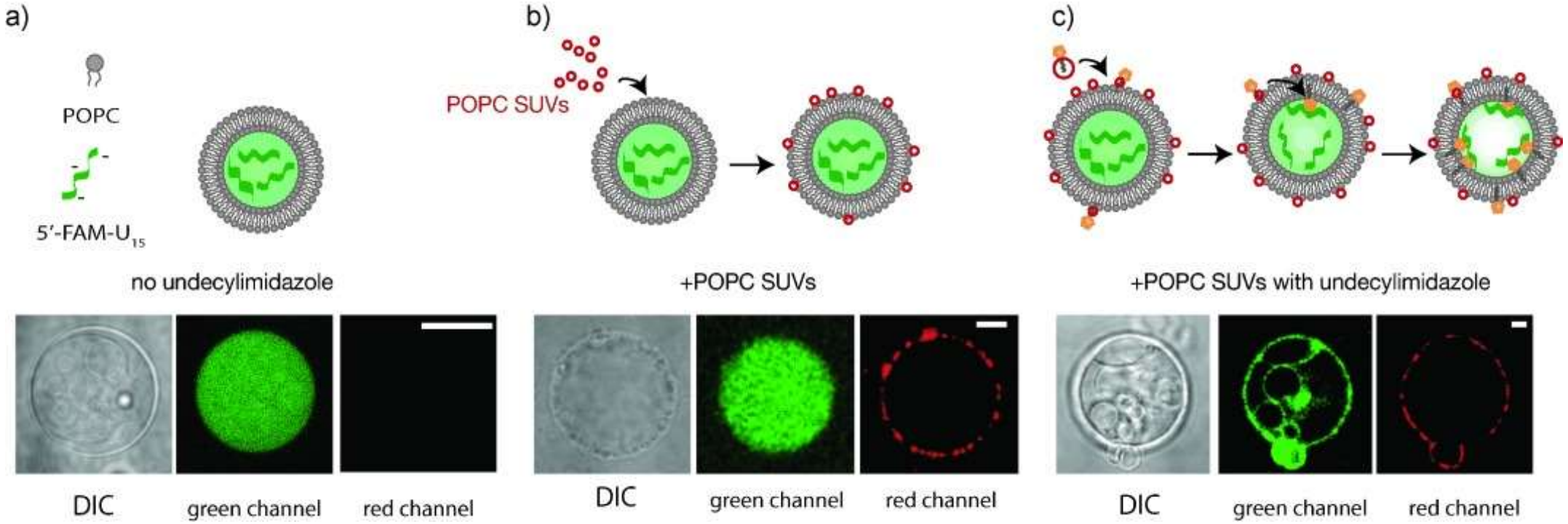
Noncovalent nucleotide association with membranes





Neha P. Kamat, Sylvia Tobe, Ian T. Hill, and Jack W. Szostak *Angew. Chem. Int. Ed.* **2015**, *54*, 11735 –11739

Noncovalent nucleotide association with membranes



Neha P. Kamat, Sylvia Tobe, Ian T. Hill, and Jack W. Szostak *Angew. Chem. Int. Ed.* **2015**, *54*, 11735 –11739