Control over Translation: The Third Genetic code

E. Westhof
Institut de biologie moléculaire et cellulaire du CNRS
Université de Strasbourg, France
2018

The genetic code is **redundant quasi-universal.**

But the representation of the codon table is arbitrary

From
« FRANCIS CRICK, Hunter of Life's Secrets »
Robert Olby
CSHLP, 2009.

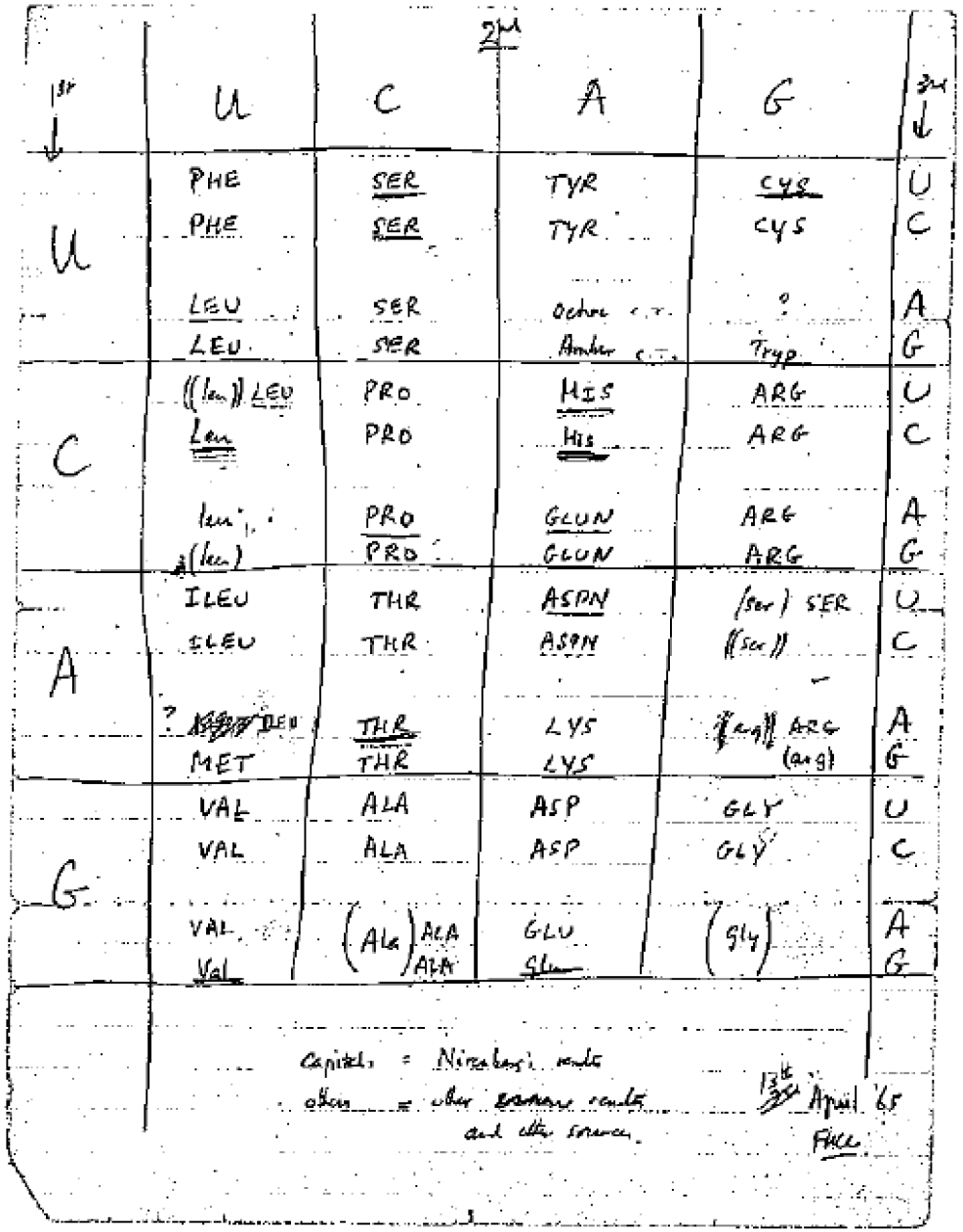
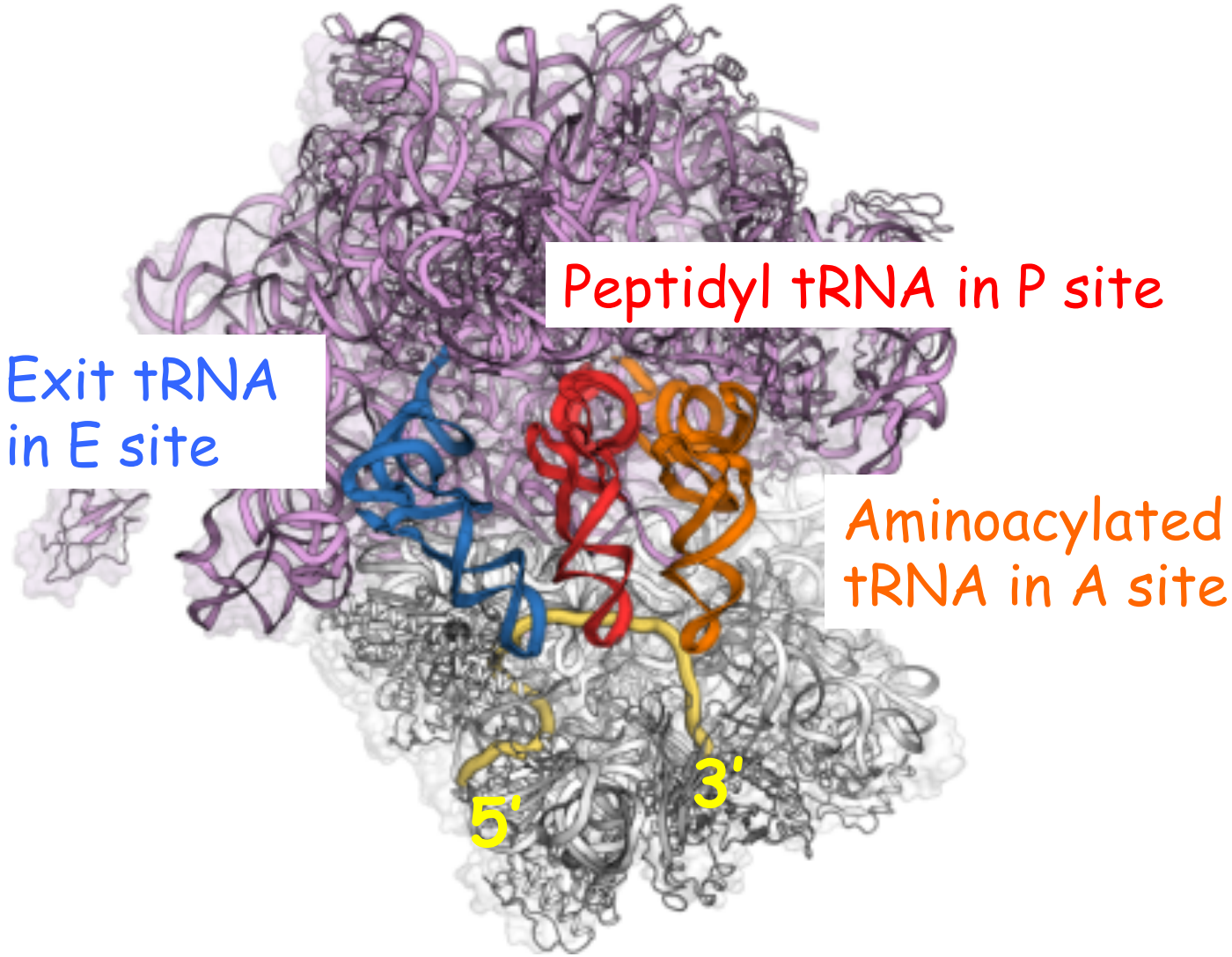


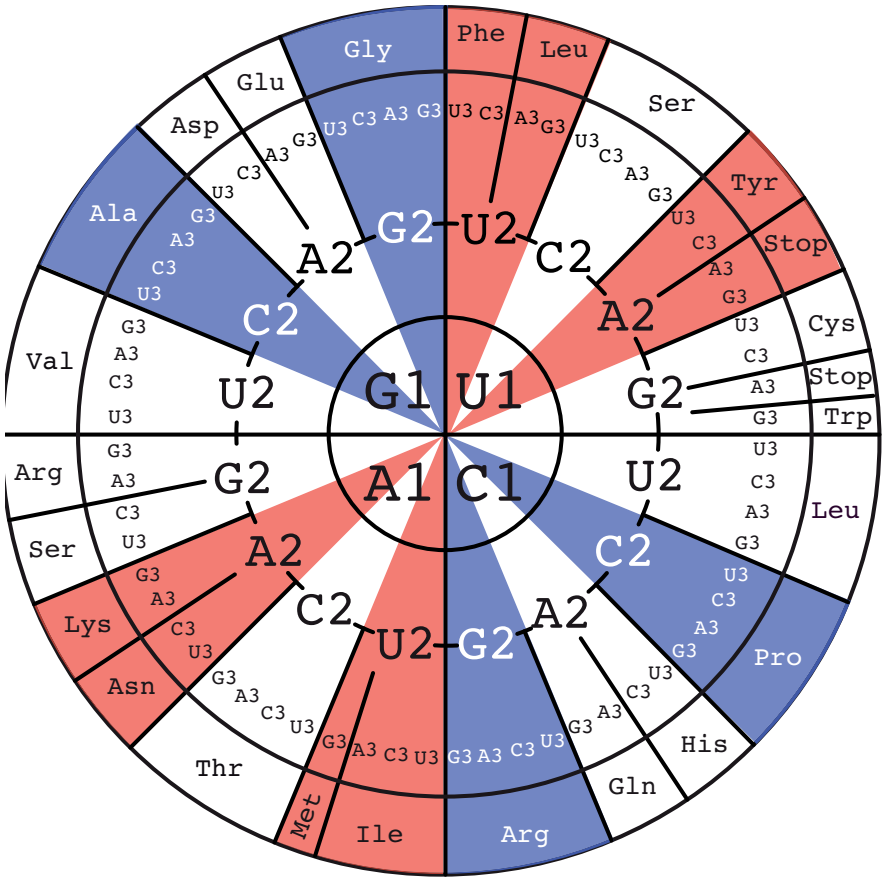
Figure 15.2 Crick's rough sketch of his checkerboard showing the stage reached in solving the genetic code in April 1965.

Only A-U pairs at 1st & 2nd

Only G=C pairs at 1st & 2nd

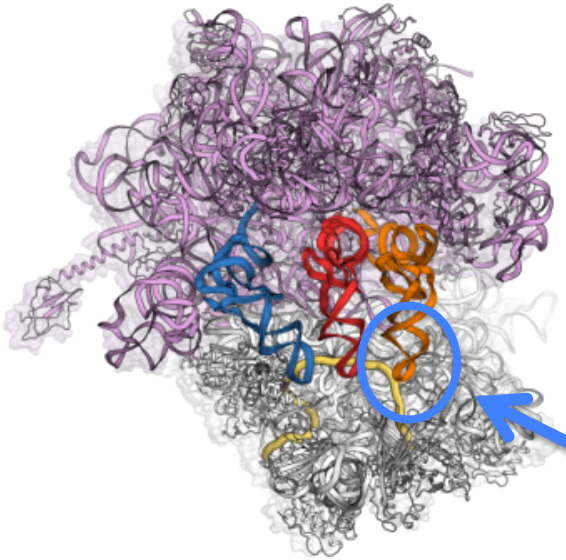
U		C		A		G	
UUU UUC	Phe	UCU UCC UCA UCG	Ser	UAU UAC	Tyr	UGU UGC	Cys
UUA UUG	Leu			UAA UAG	*Stop*	UGA	*Stop*
						UGG	Trp
CUU CUC CUA CUG	Leu	CCU CCC CCA CCG	Pro	CAU CAC	His	CGU CGC CGA CGG	Arg
				CAA CAG	Gln		
AUU AUC AUA	Ile	ACU ACC ACA ACG	Thr	AAU AAC	Asn	AGU AGC	Ser
AUG	Met			AAA AAG	Lys	AGA AGG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

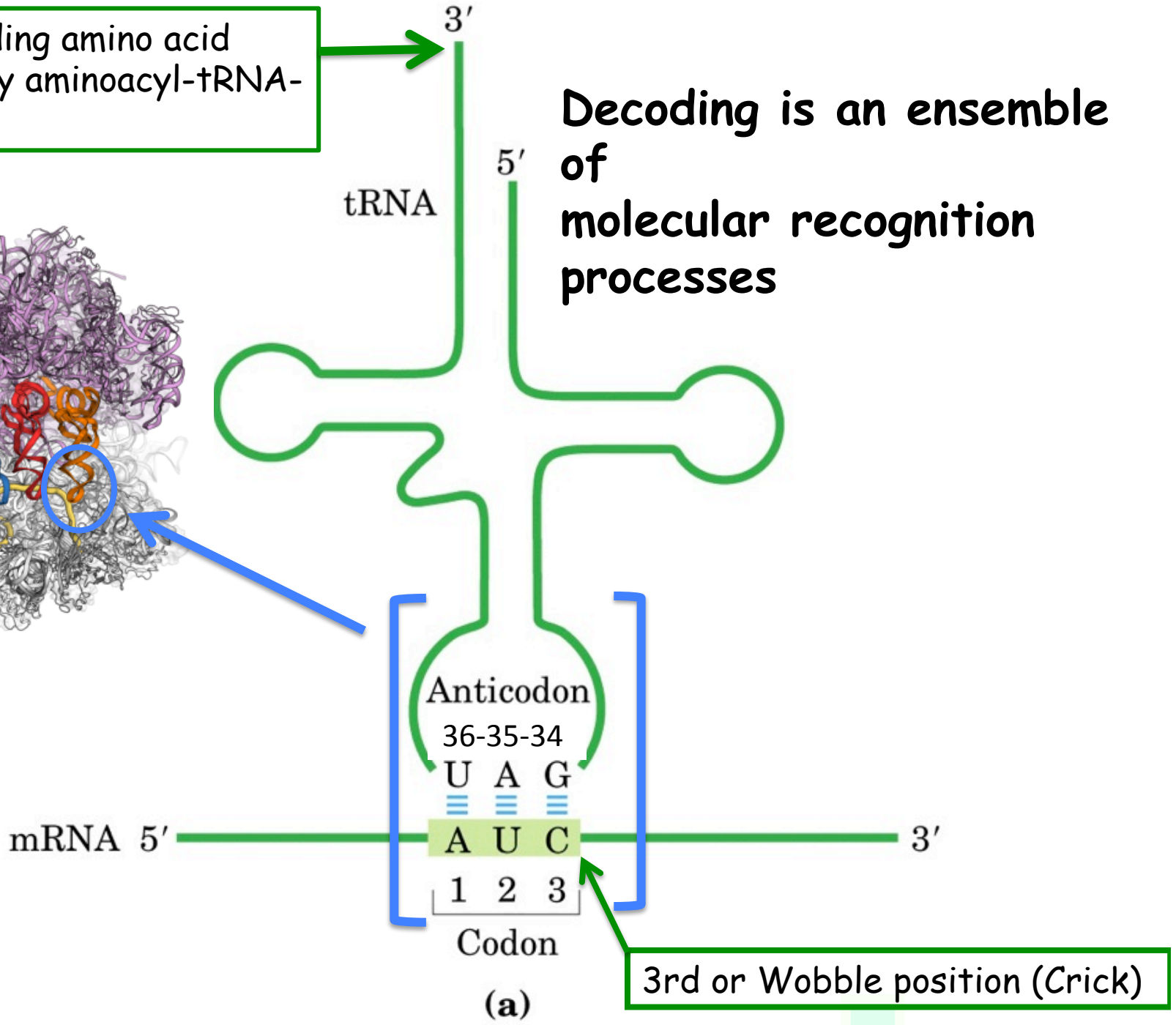


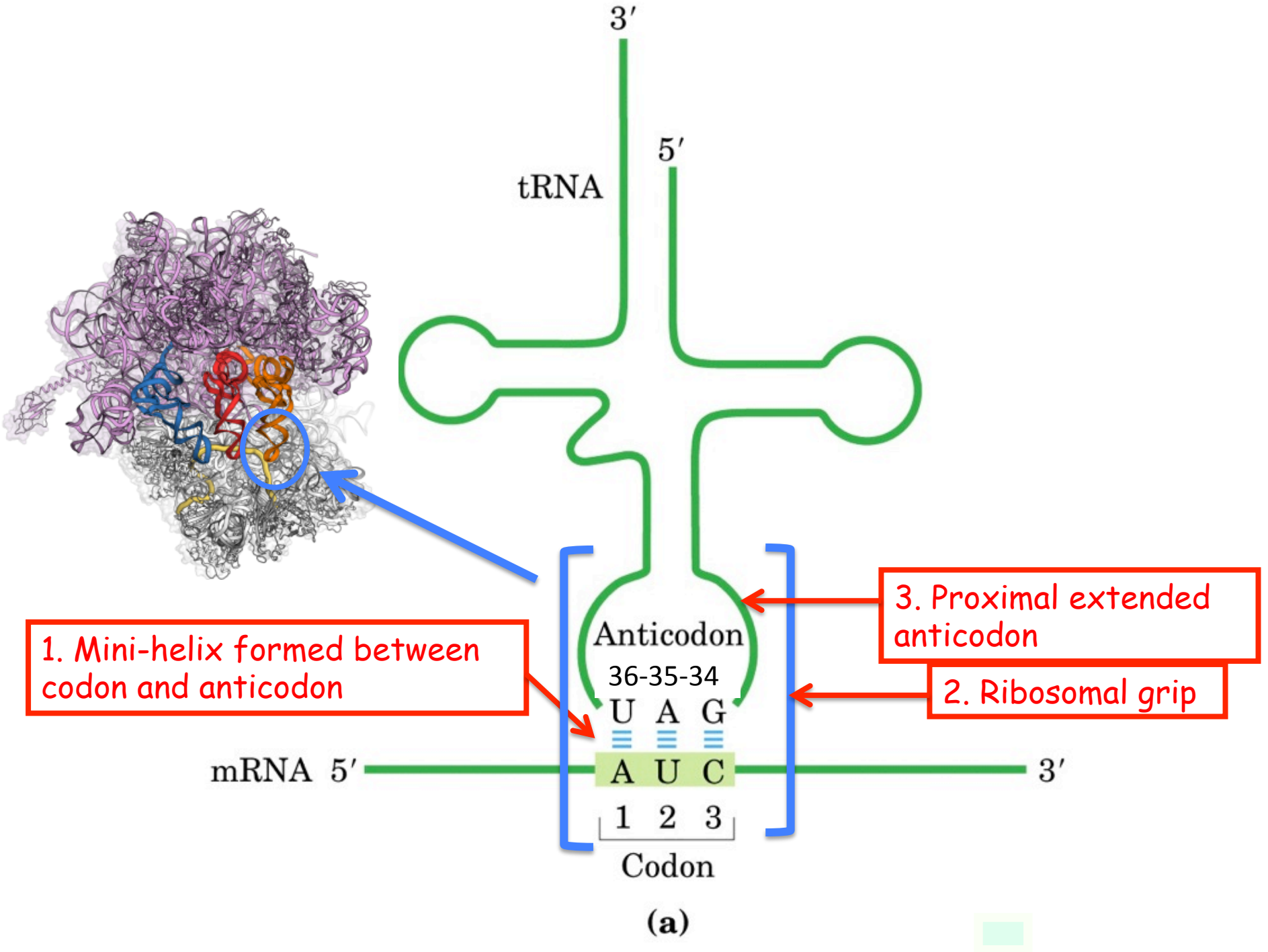

Gly
Phe
Leu
Glu
Ser
Asp
Tyr
Ala
Stop
Cys
Val
Stop
Trp
Arg
Leu
Ser
Pro
Lys
His
Asn
Gln
Thr
Met
Ile
Arg
G1
U1
A1
C1
G2
U2
A2
C2
U3
C3
A3
G3

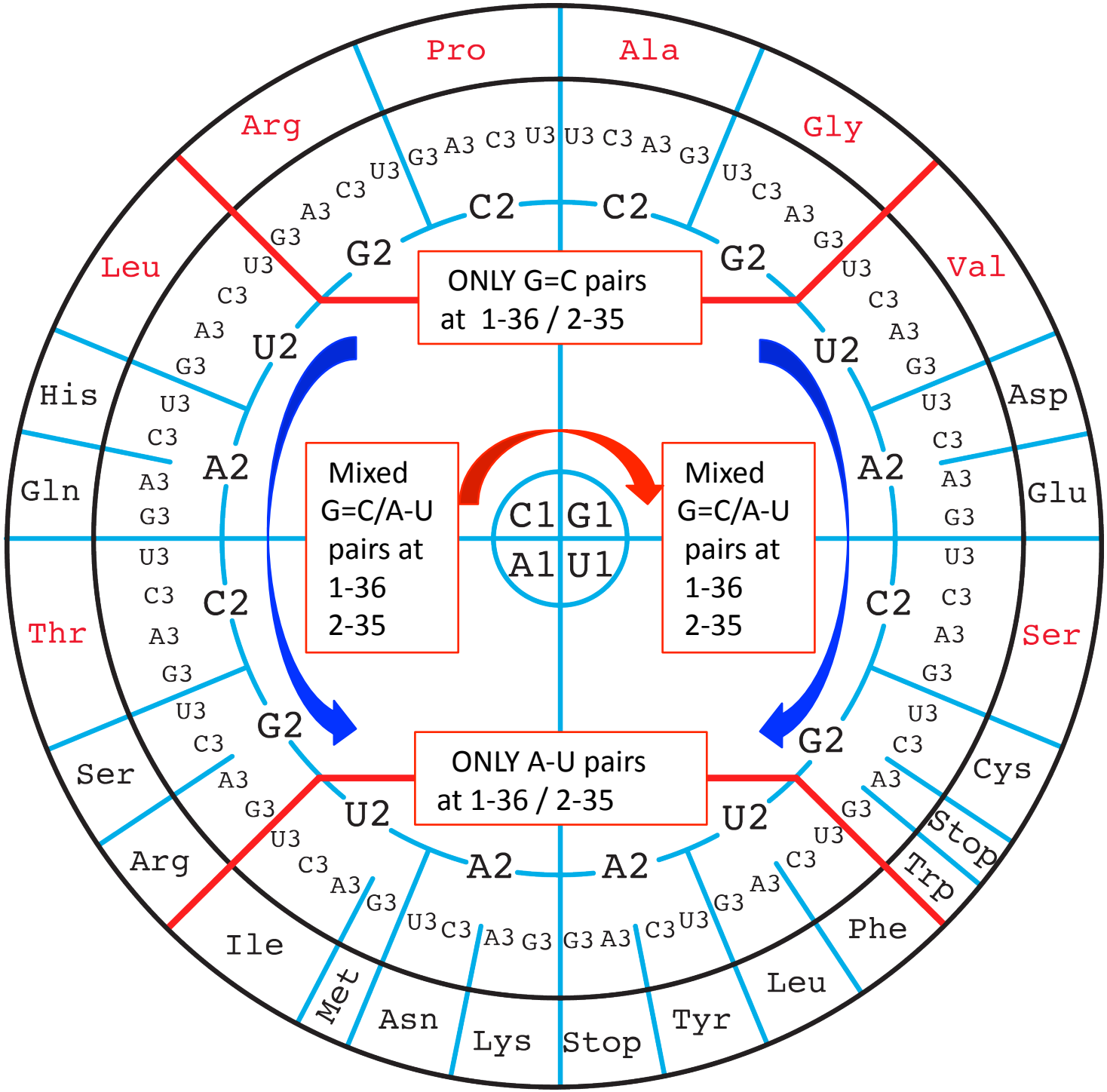
Corresponding amino acid attached by aminoacyl-tRNA-synthetase

Decoding is an ensemble of molecular recognition processes

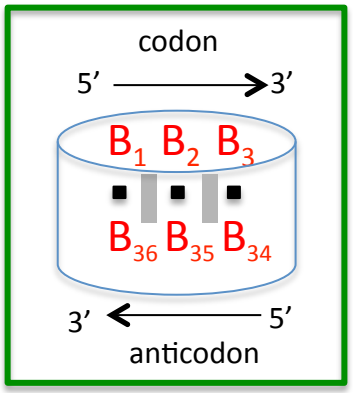


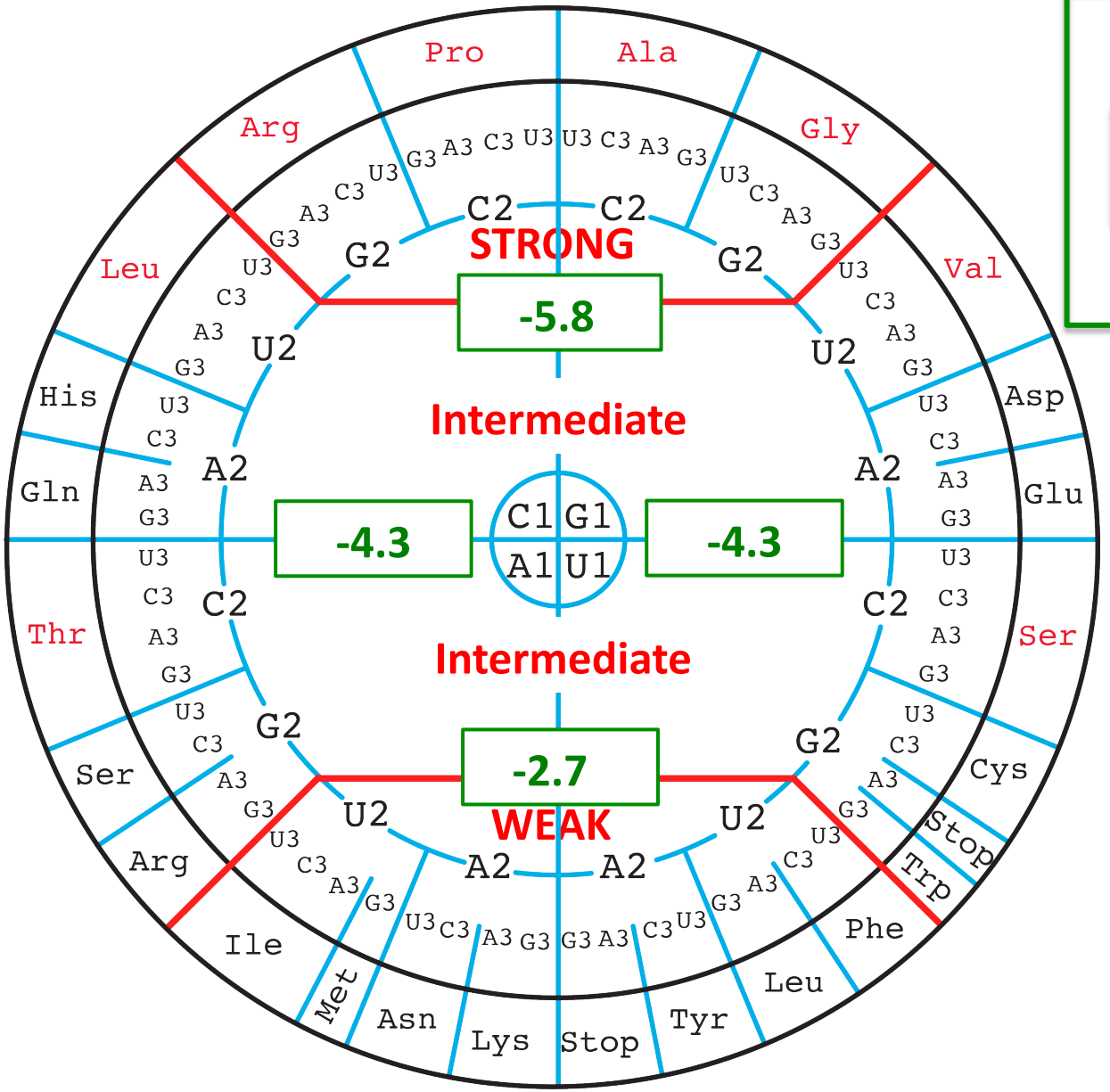


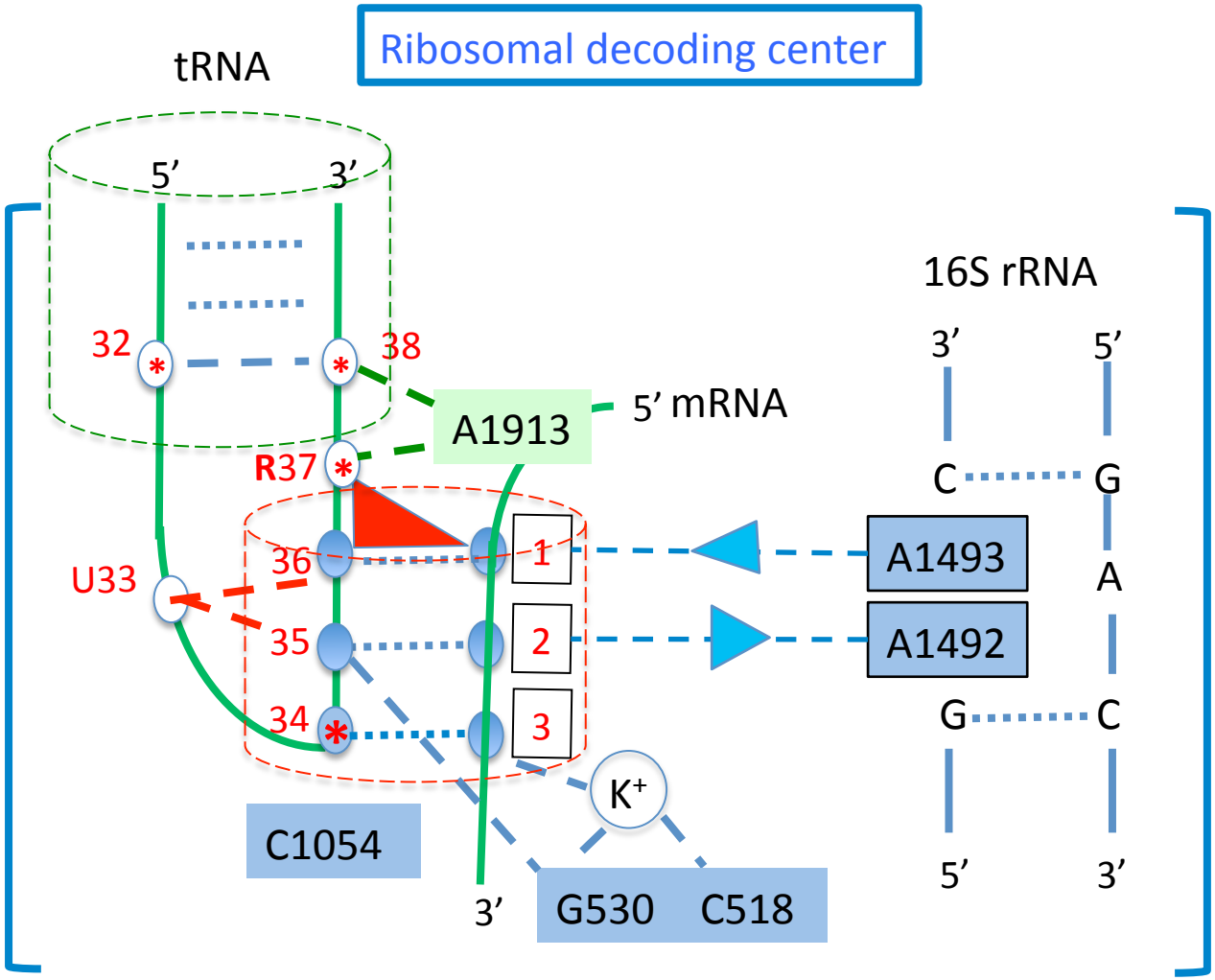

3′
5′
tRNA
1. Mini-helix formed between codon and anticodon
3. Proximal extended anticodon
2. Ribosomal grip
Anticodon
36-35-34
U A G
mRNA 5′
A U C
3′
1 2 3
Codon
(a)


Pro
Ala
Arg
Gly
Leu
Val
His
Asp
Gln
Glu
Thr
Ser
Ser
Cys
Arg
Stop
Trp
Ile
Phe
Met
Asn
Lys
Stop
Tyr
Leu
ONLY G=C pairs
at 1-36 / 2-35
ONLY A-U pairs
at 1-36 / 2-35
Mixed
G=C/A-U
pairs at
1-36
2-35
Mixed
G=C/A-U
pairs at
1-36
2-35
C1 G1
A1 U1
U2
C2
A2
G2

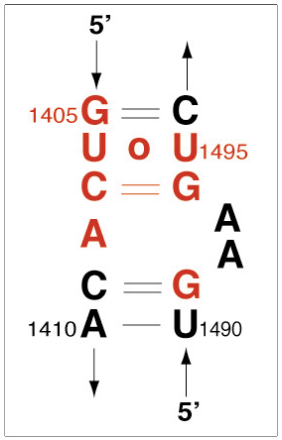
1. Calculated energies of the AntiCodon-Codon mini-helix







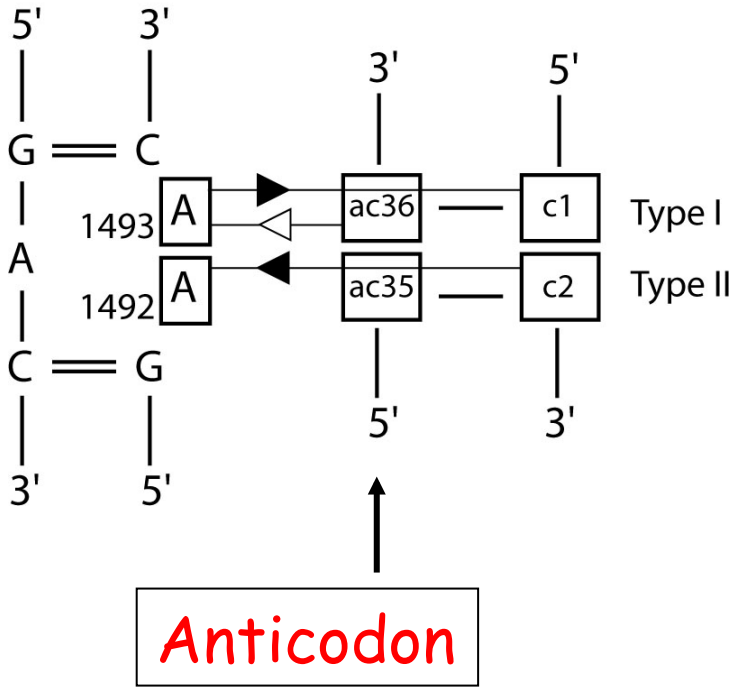
* < means often modified



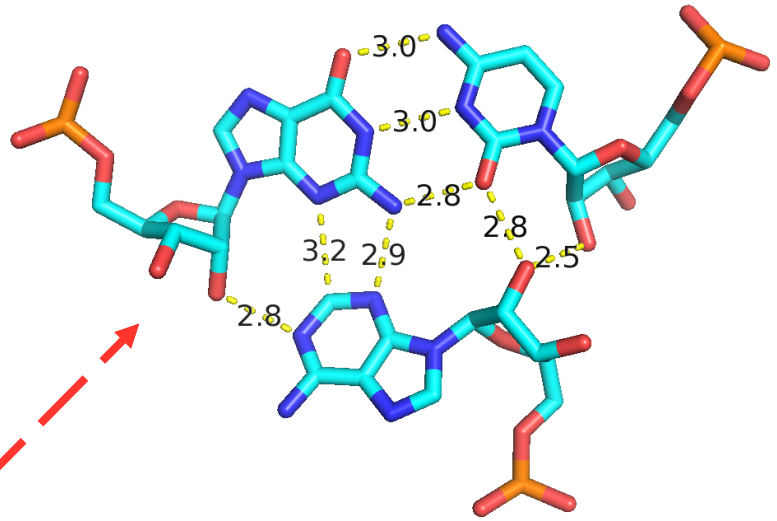
A site

Codon

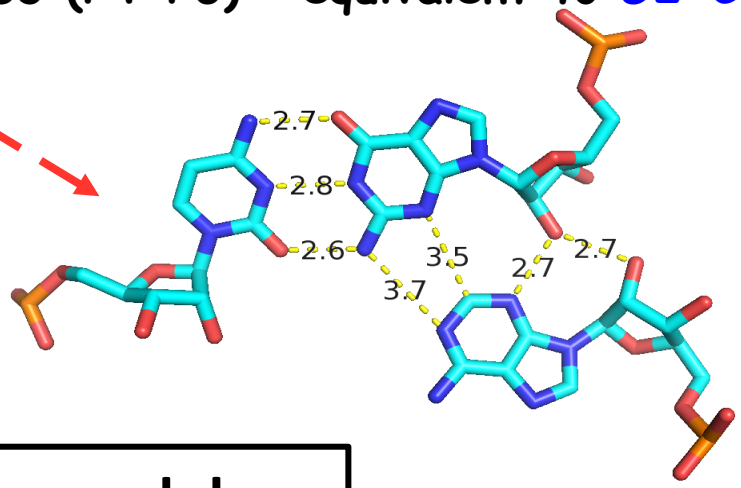
Anticodon/Codon minihelix



A184 (P4-P6) equivalent to C1=G36

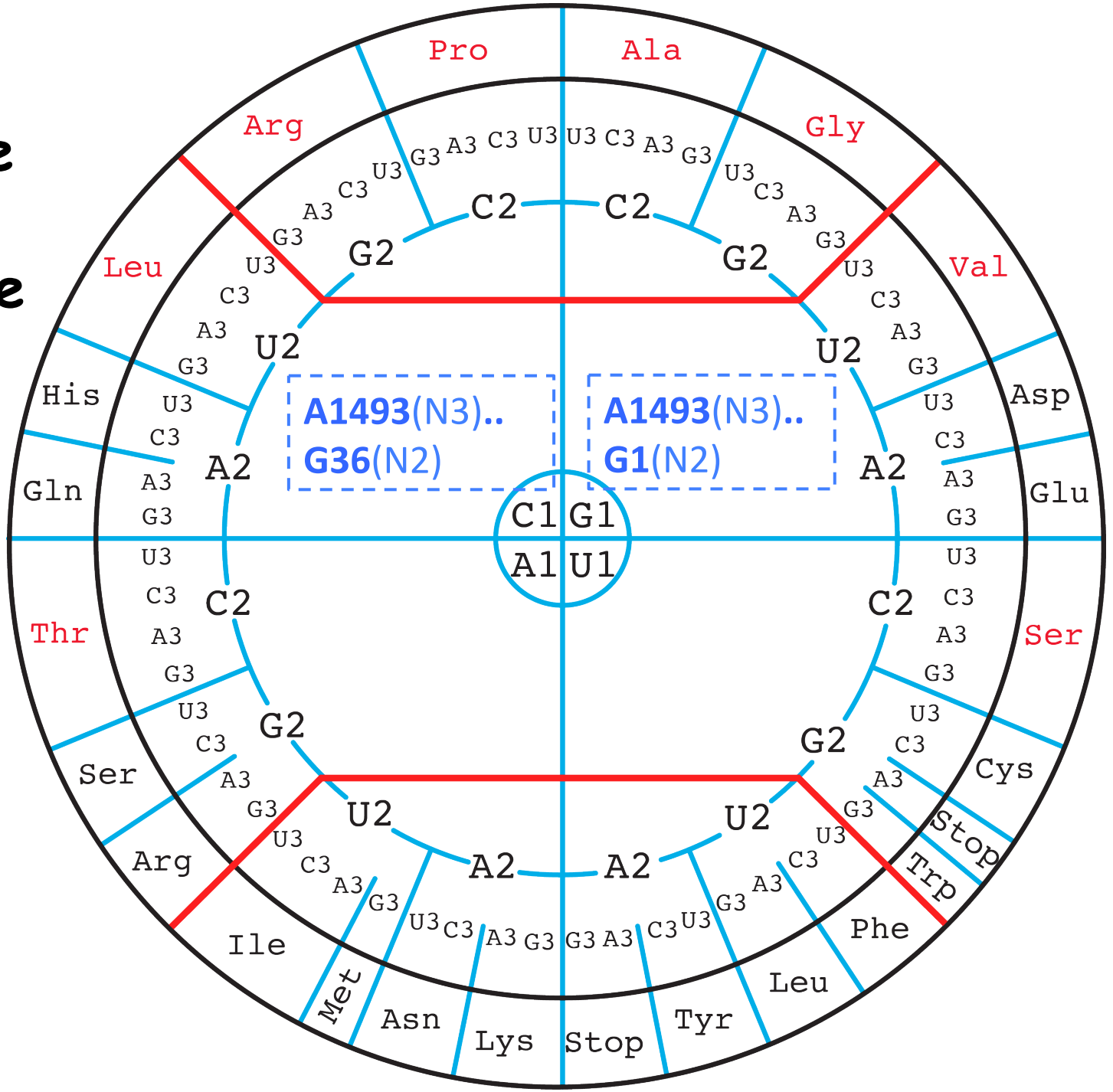


A183 (P4-P6) equivalent to G2=C35



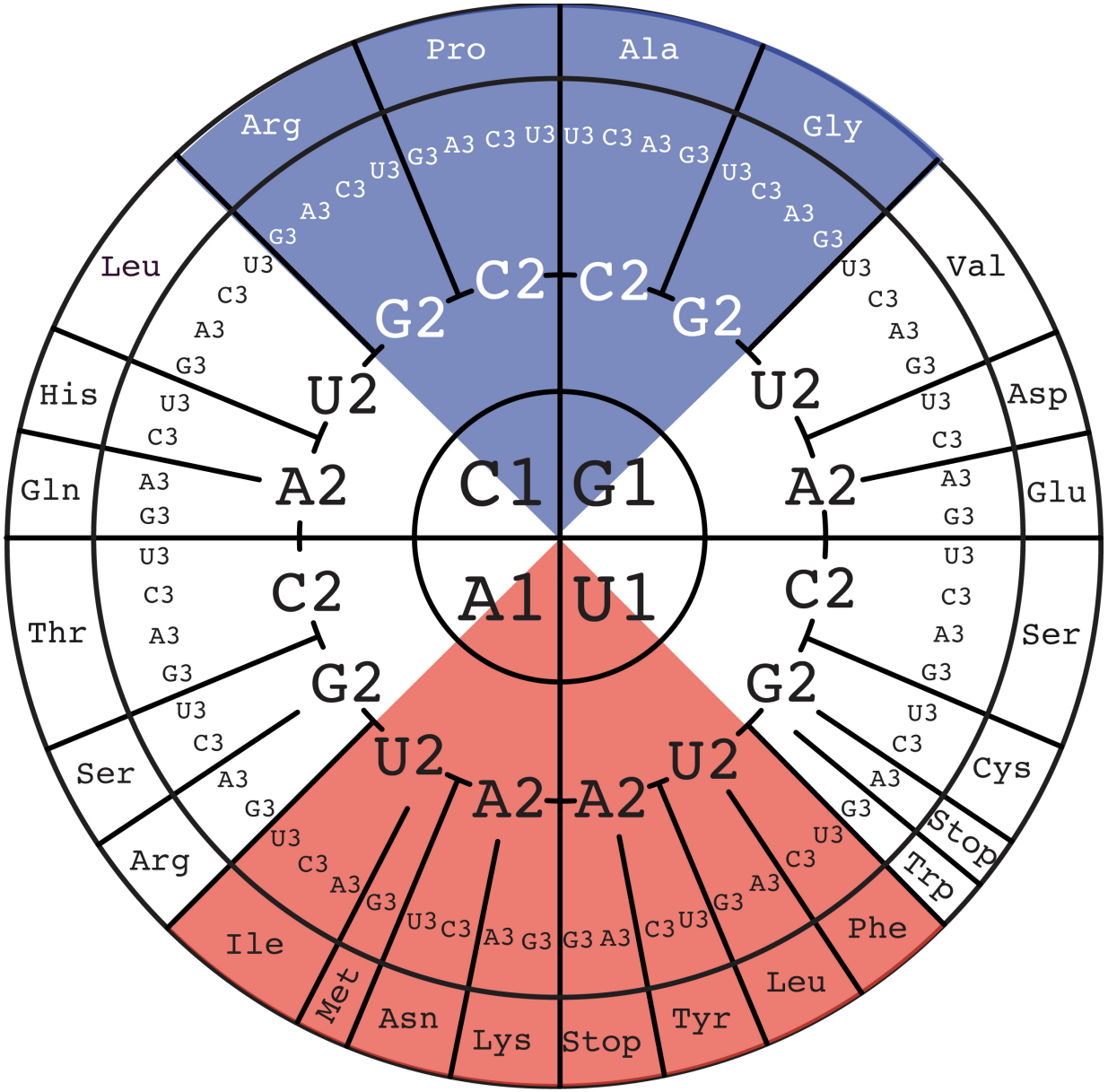
The most common RNA-RNA interaction module

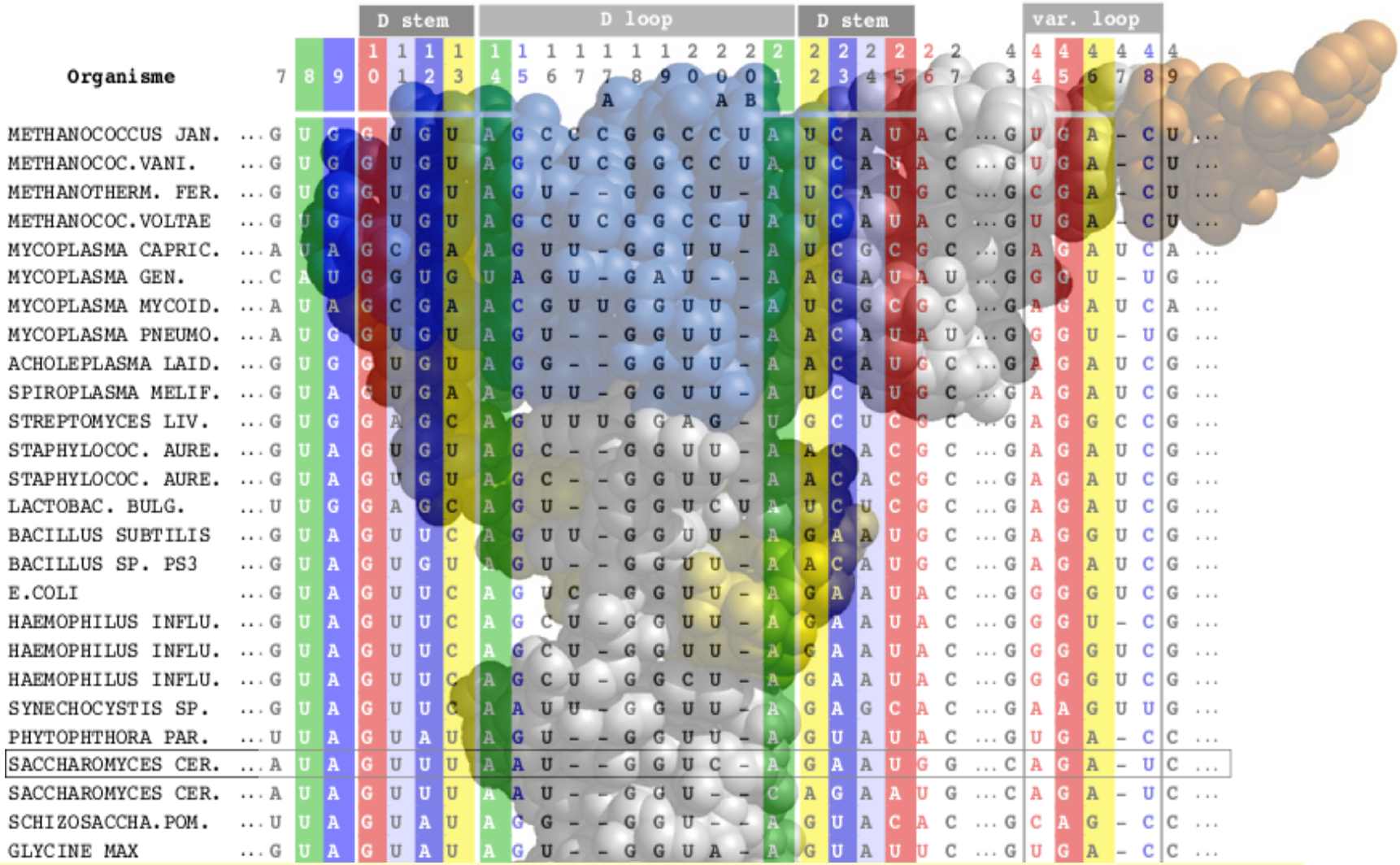
Most of the contacts are invariant except those



How to optimize the ribosome « rhythm »?
How to optimize protein synthesis and folding?

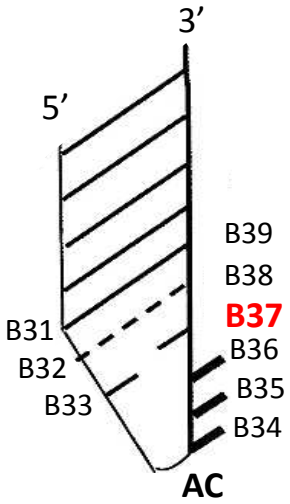
Increasing stability of network interactions



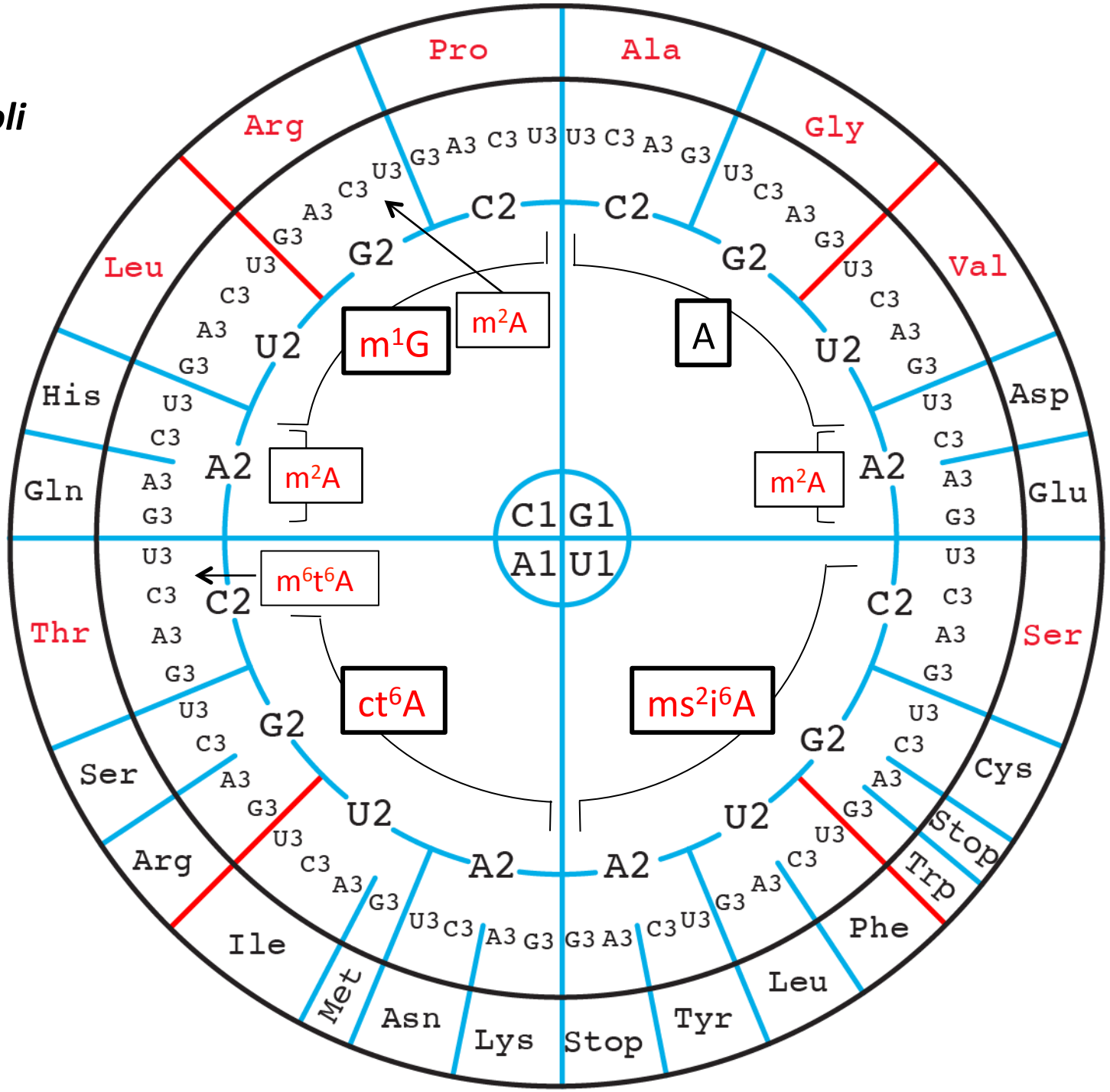


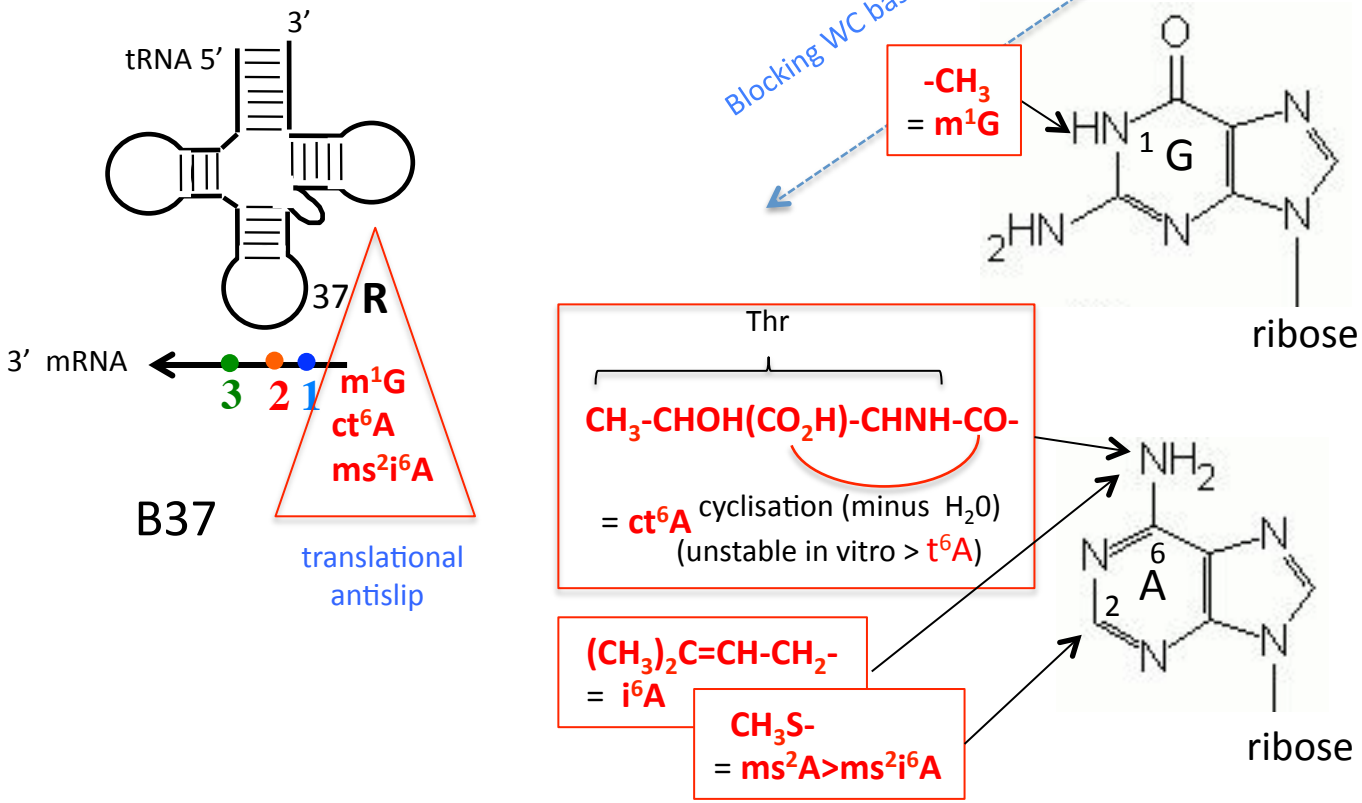
tRNAs are very active partners during ribosomal translation

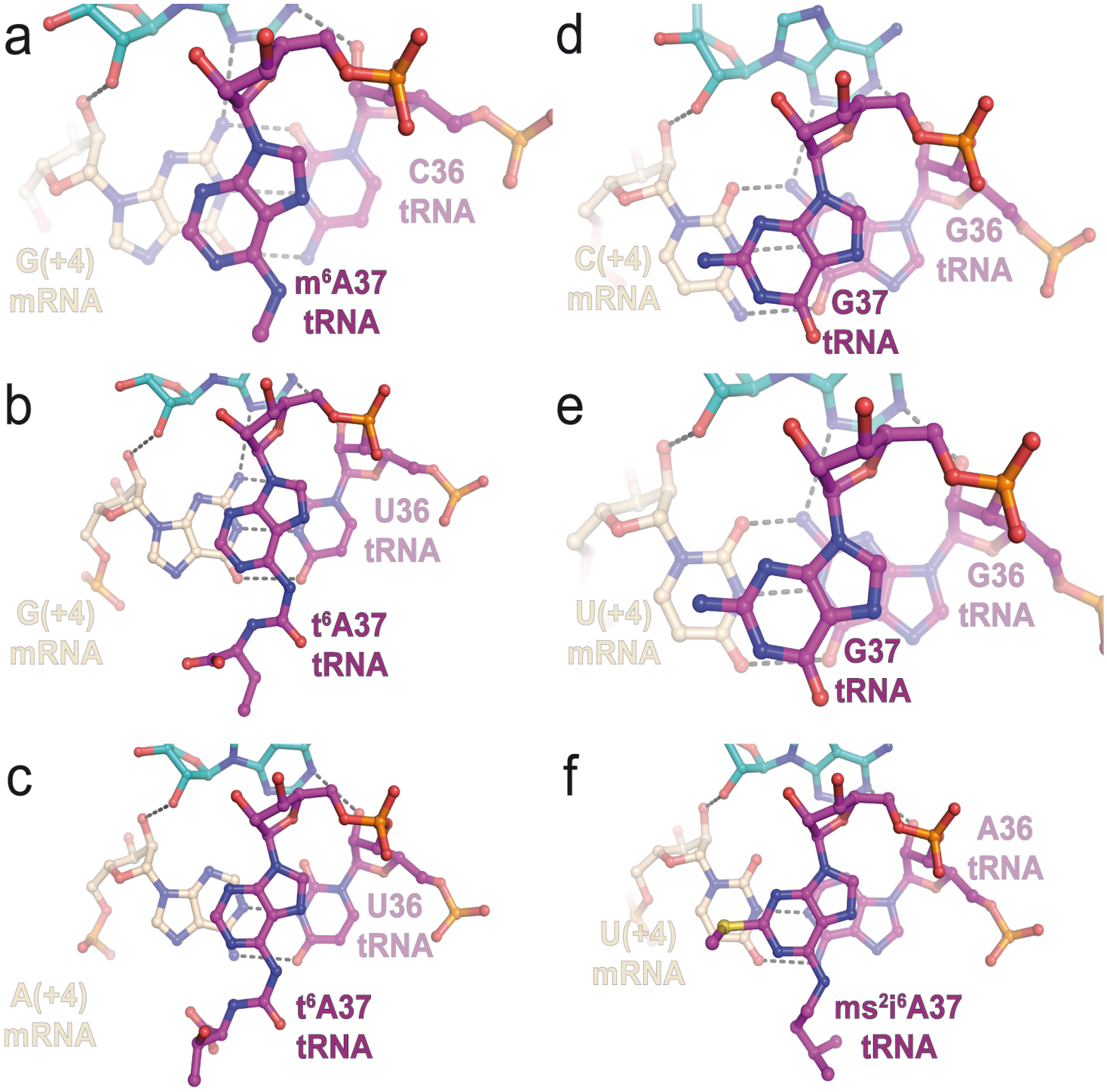
Escherichia coli

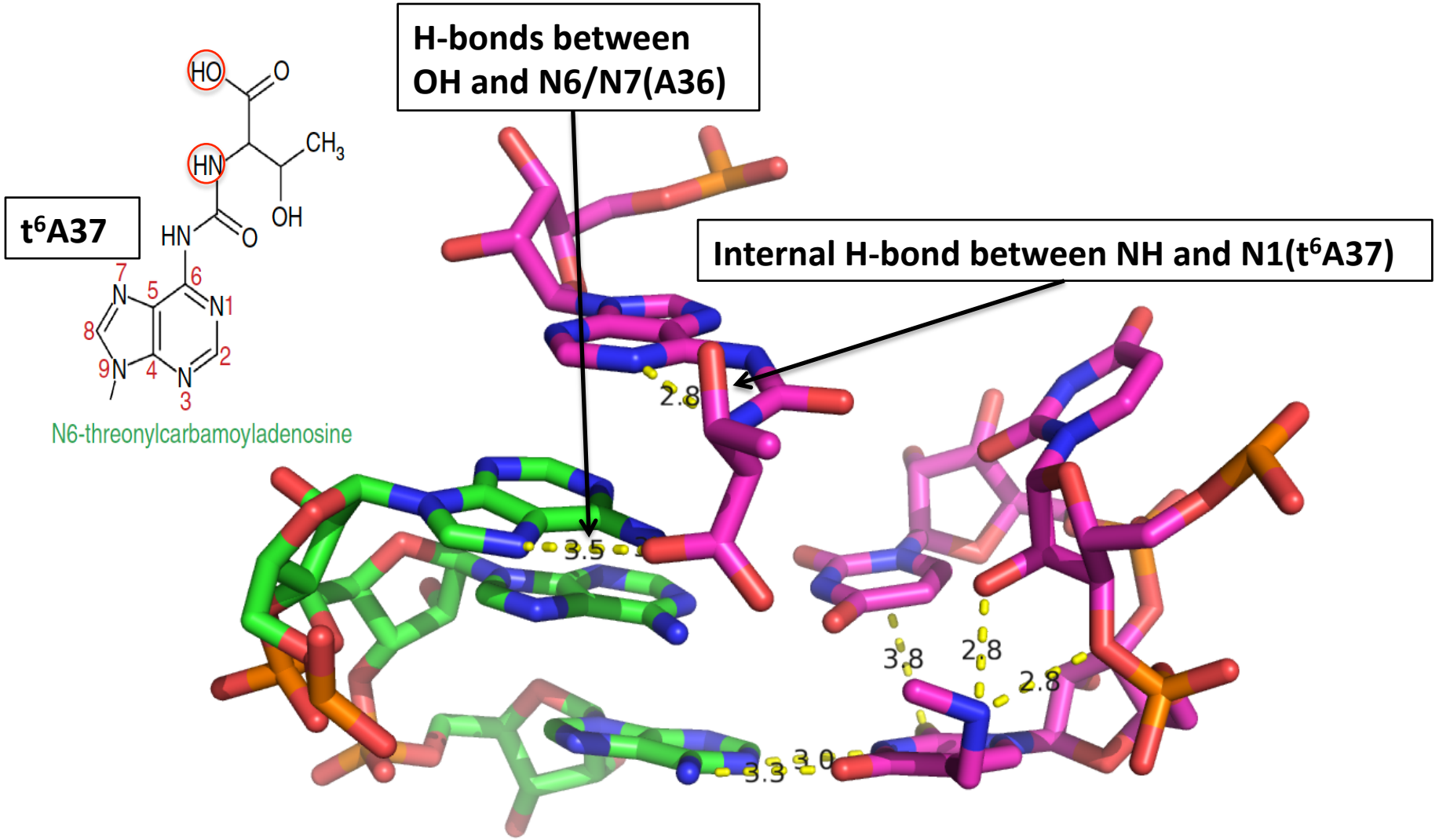


B37



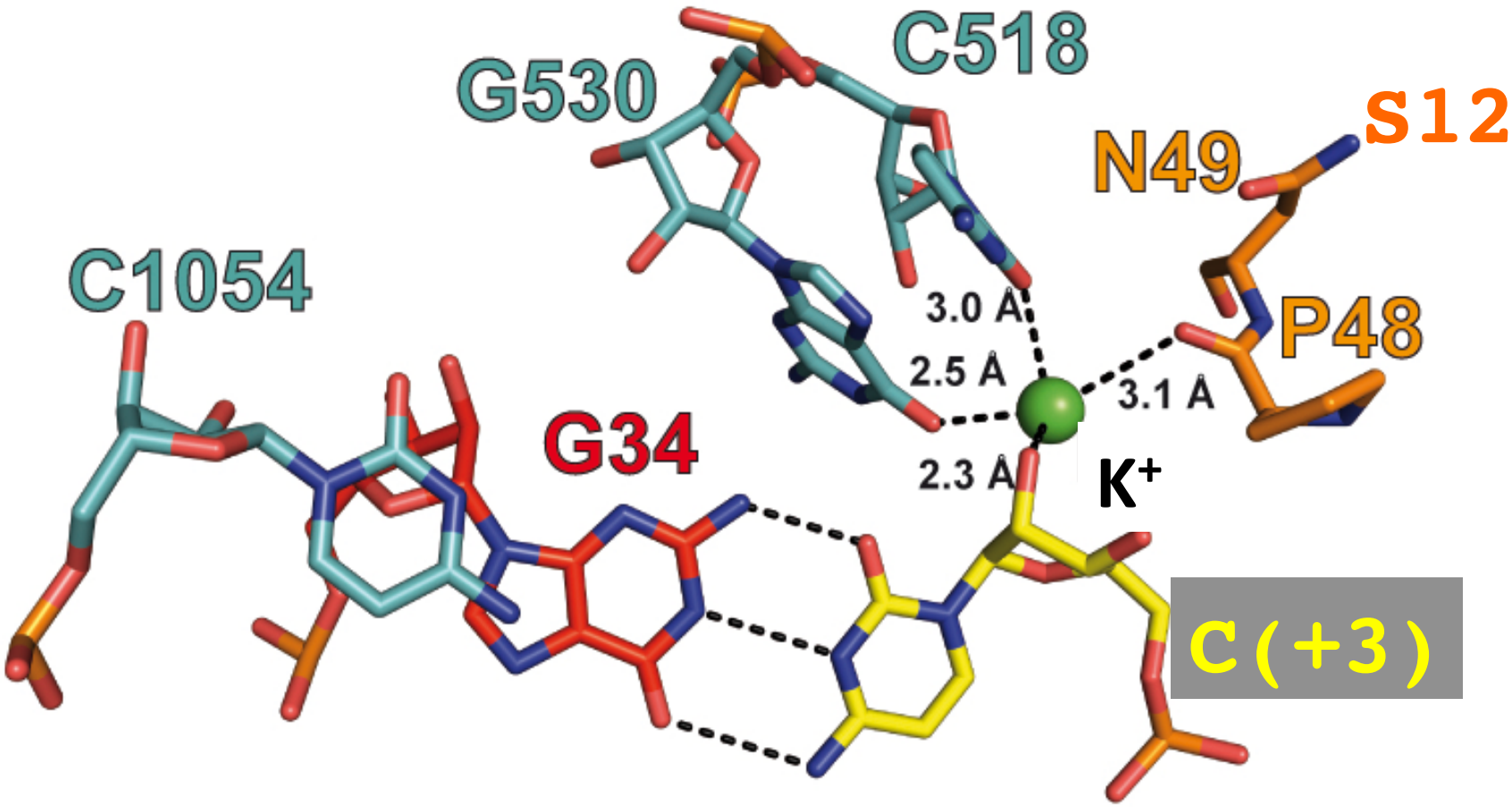

3’
tRNA 5’
37
R
3’ mRNA
3 2 1
m¹G
ct⁶A
ms²i⁶A
B37
translational
antislip
Blocking WC base-pairing
-CH₃
= m¹G
HN
1
G
2HN
N
O
ribose
Thr
CH₃-CHOH(CO₂H)-CHNH-CO-
= ct⁶A cyclisation (minus H₂0)
(unstable in vitro > t⁶A)
(CH₃)₂C=CH-CH₂-
= i⁶A
CH₃S-
= ms²A>ms²i⁶A
NH₂
6
A
2
N
ribose

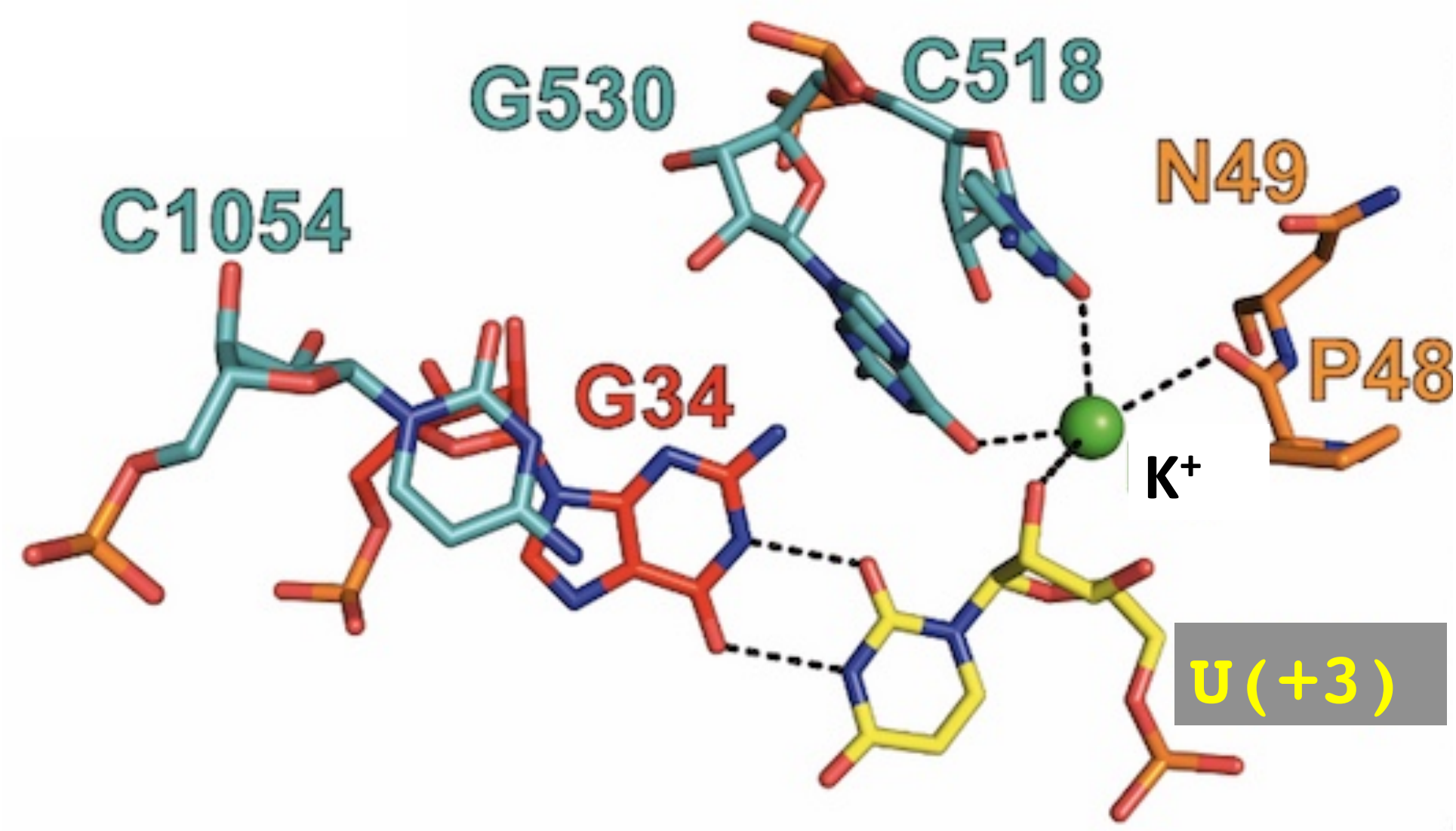

a
G(+4)
mRNA
C36
tRNA
m⁶A37
tRNA
b
G(+4)
mRNA
U36
tRNA
t⁶A37
tRNA
c
A(+4)
mRNA
U36
tRNA
t⁶A37
tRNA
d
C(+4)
mRNA
G36
tRNA
G37
tRNA
e
U(+4)
mRNA
G36
tRNA
G37
tRNA
f
U(+4)
mRNA
A36
tRNA
ms²i⁶A37
tRNA



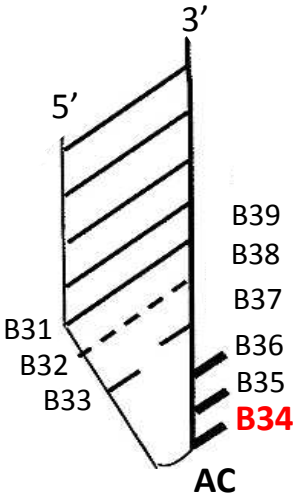
Rozov, et al. Nature Comm. 7, 10457(2016)

Around the third base pair



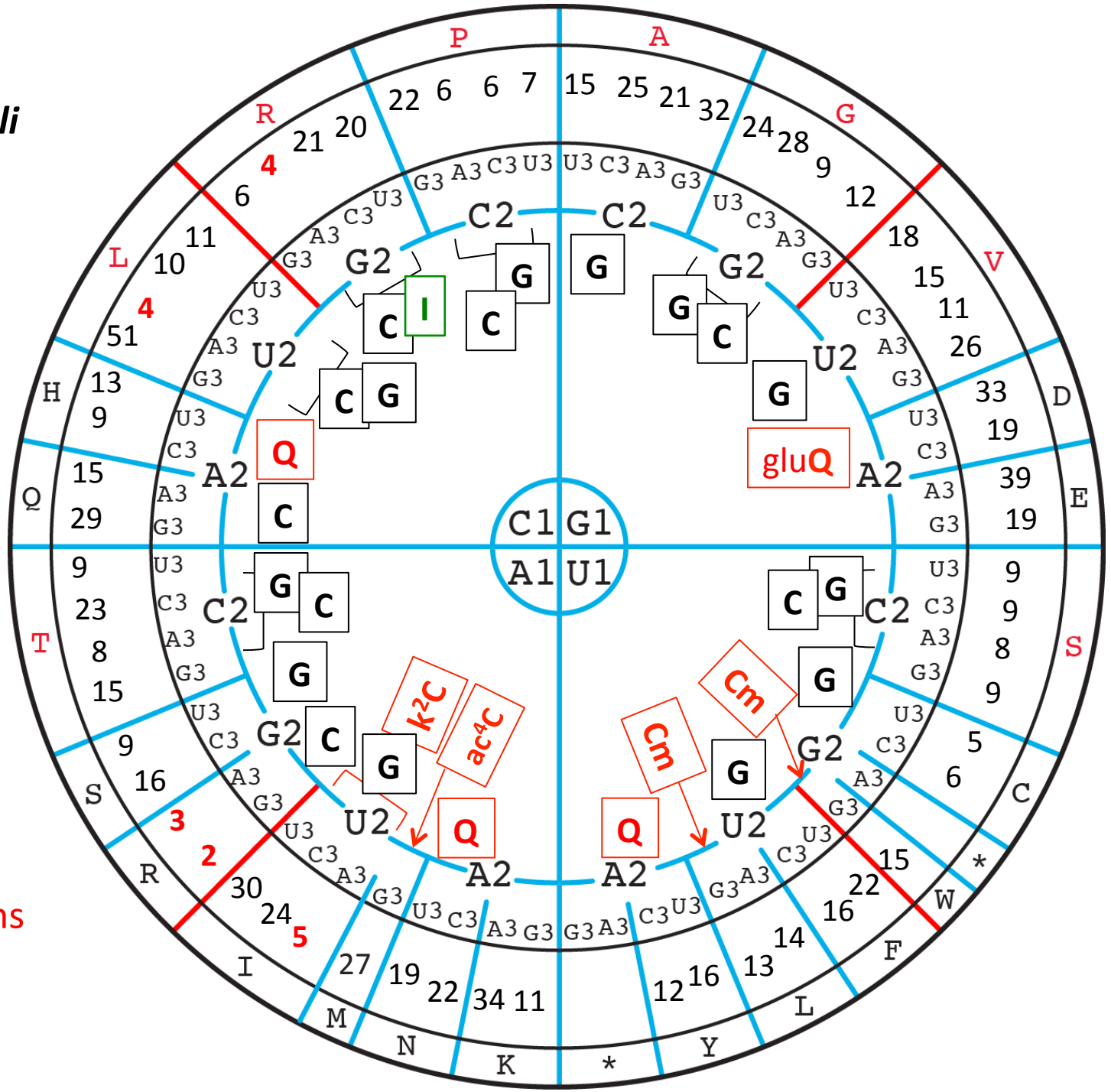

G530
C518
N49
C1054
P48
G34
K+
U(+3)

Escherichia coli



B34
except
(U34)

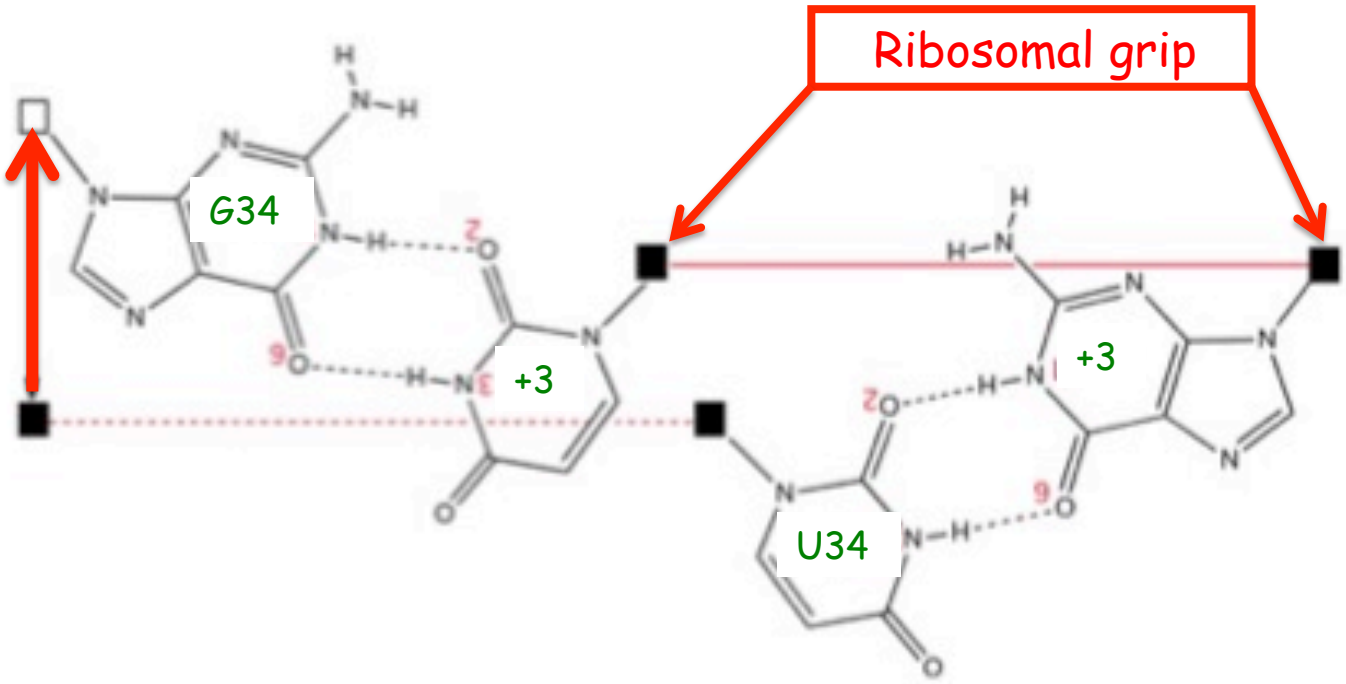
Y3-ending codons



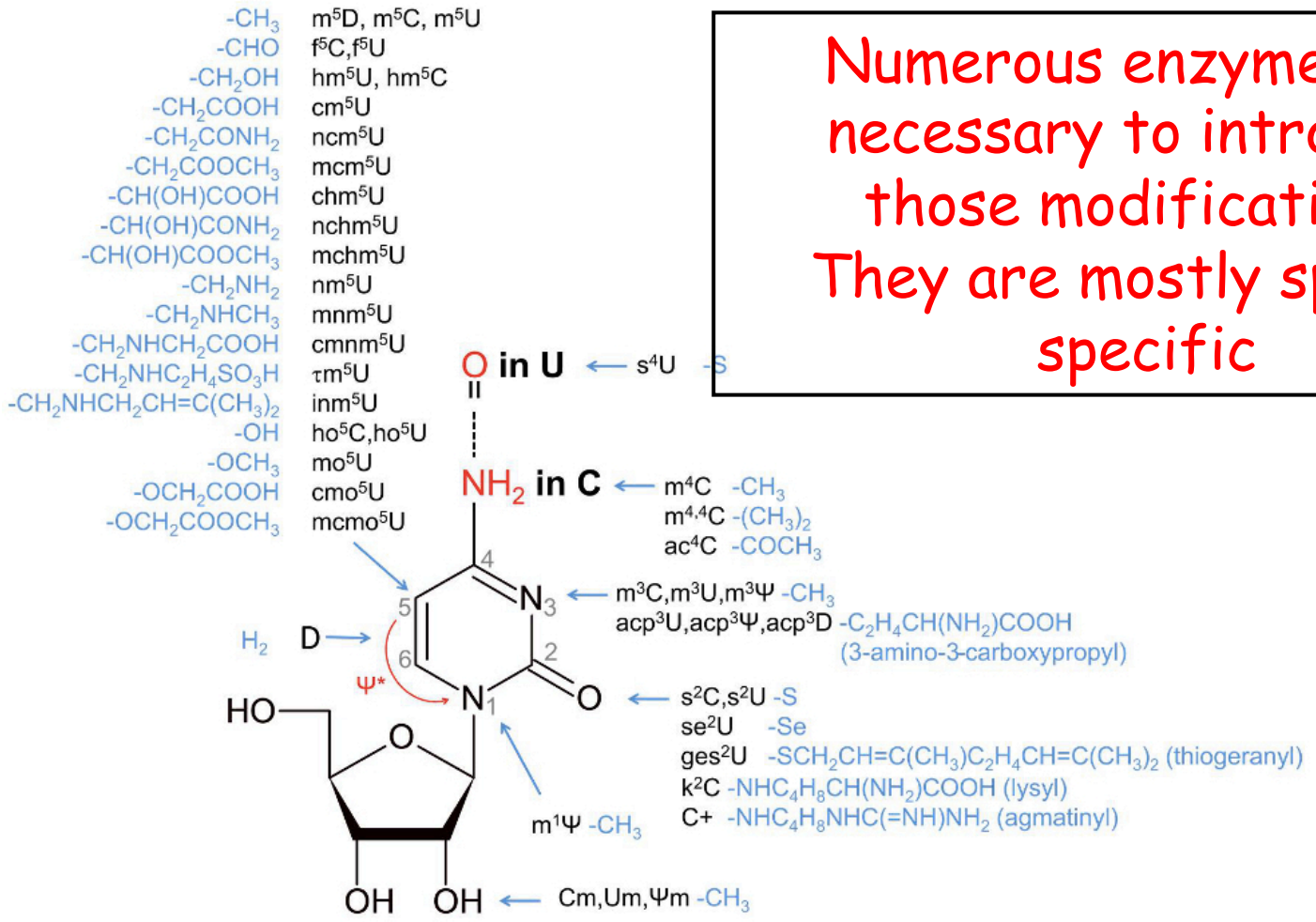
Uridine in the mRNA G34oU(+3)

versus

Uridine in the anticodon U34oG(+ 3)



More than 110 modified nucleotides in transfert RNA



Numerous enzymes are necessary to introduce those modifications. They are mostly species specific

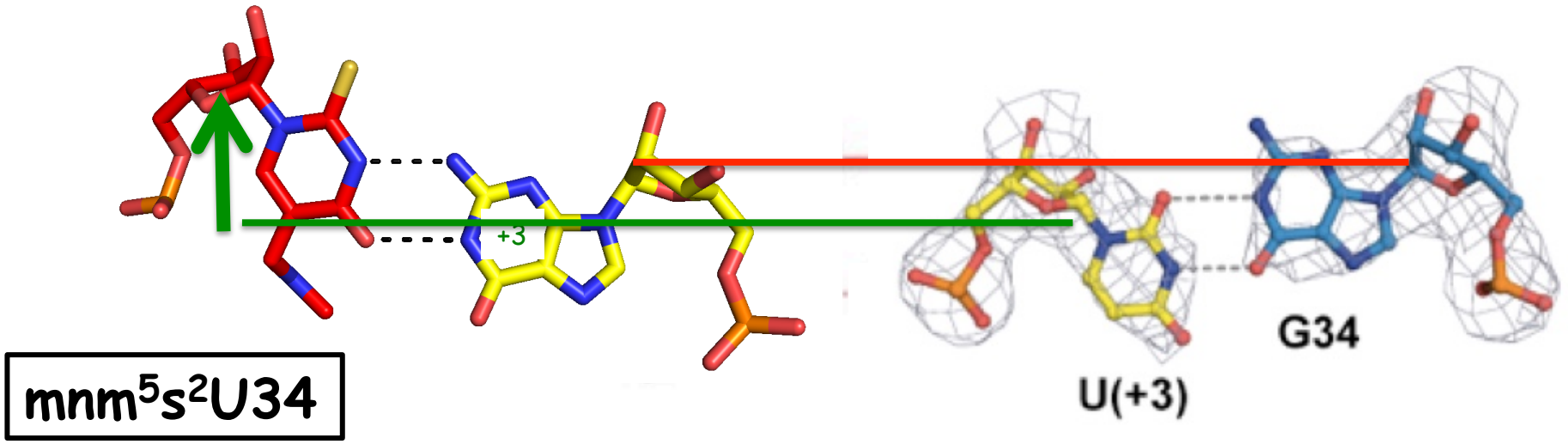
From Modomics database 2013
(http://modomics.genesilico.pl/)

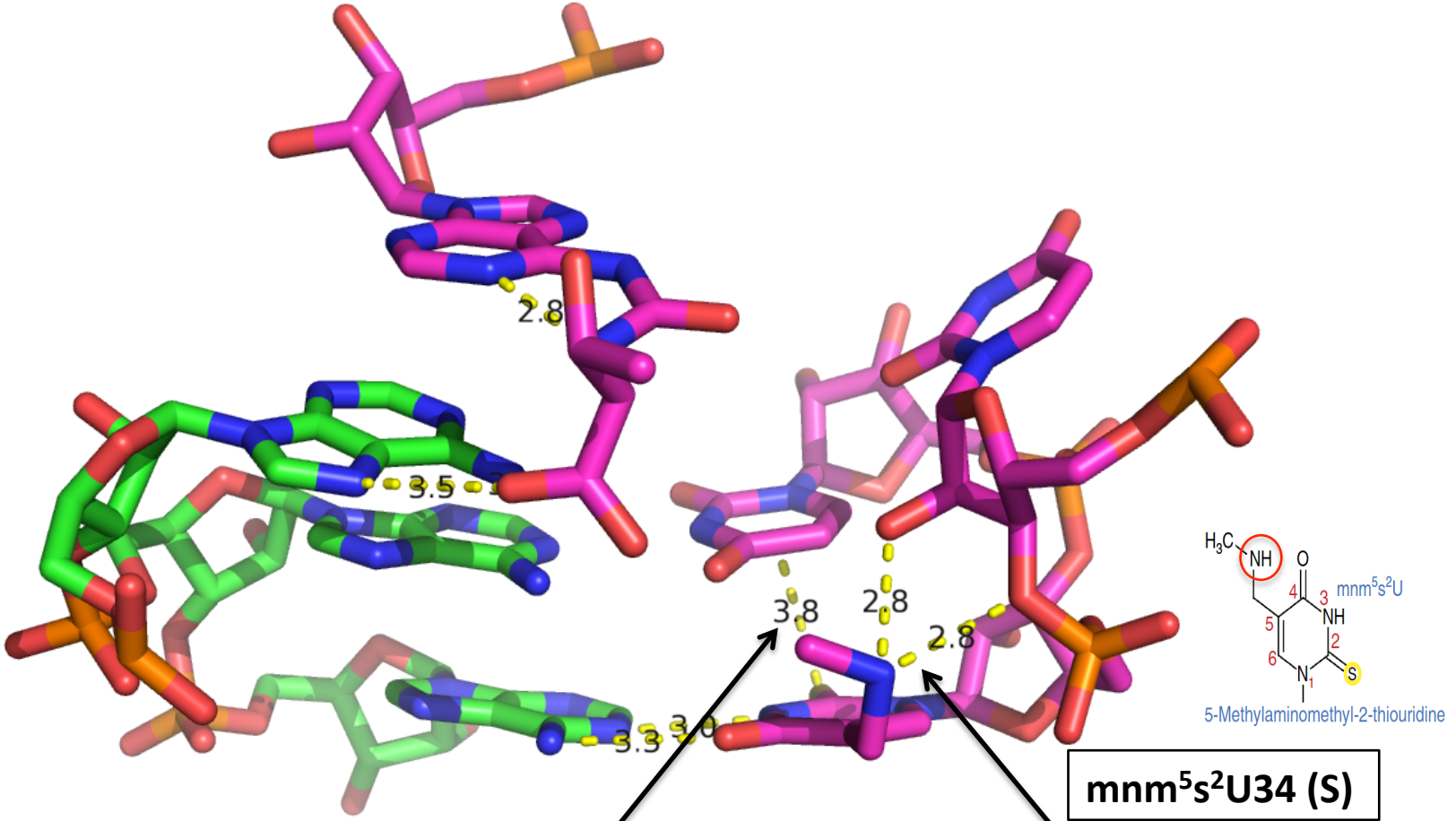
Modified Uridine mnm^{5}s^{2}U34oU(+3)

versus

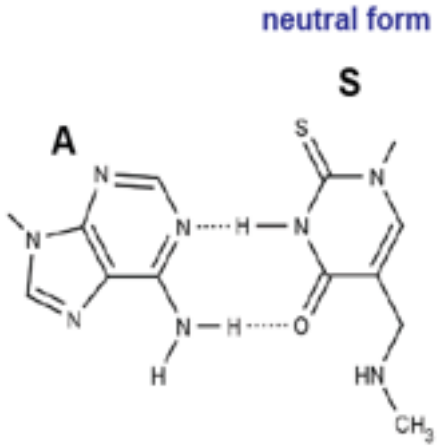
Non-modified Uridine U34oG(+ 3)

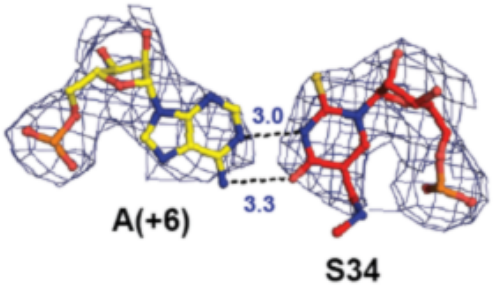
in the anticodon

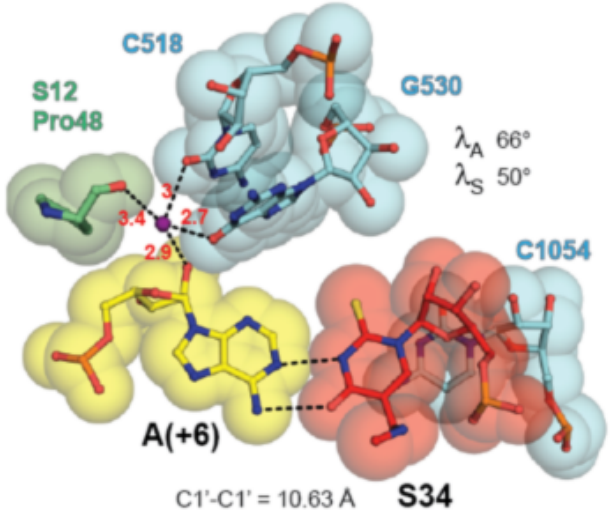




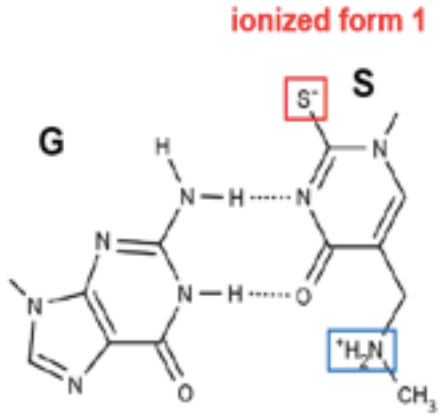
Rozov, et al. Nature Comm. 7, 10457(2016)

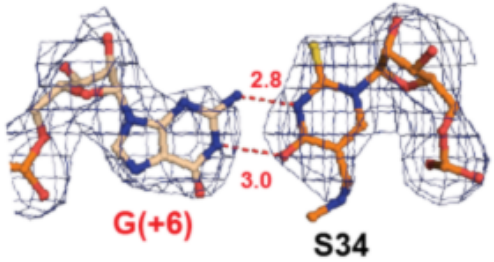


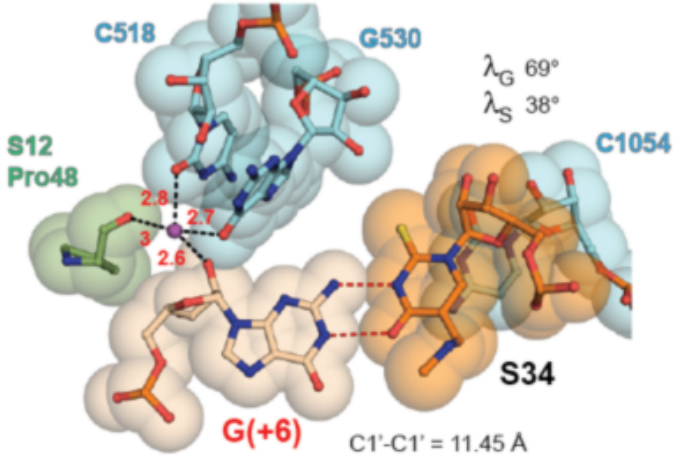


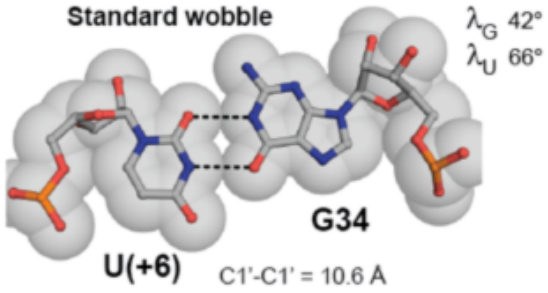


mnm^{5}s^{2}U

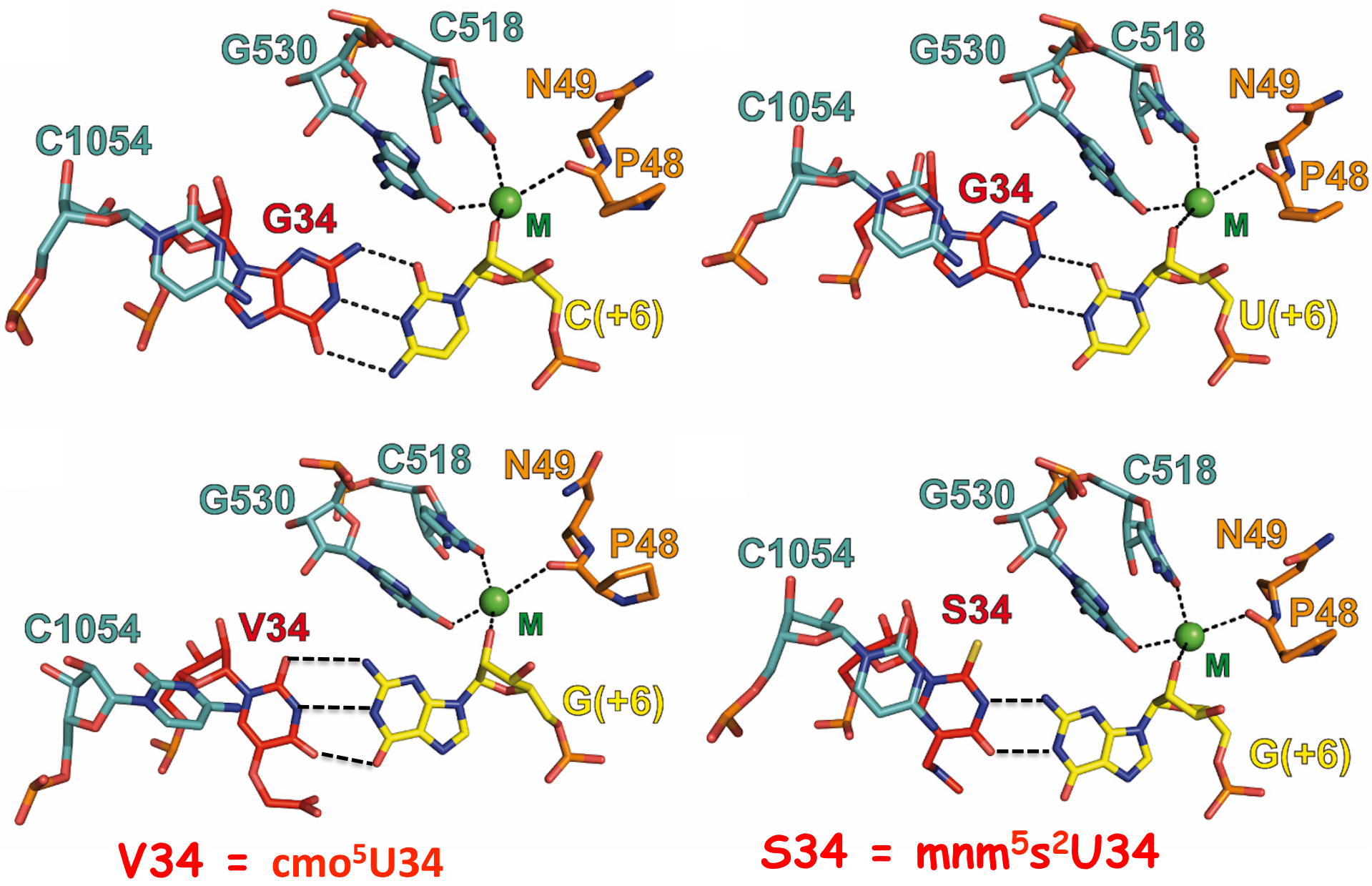








Rozov, et al. Nature Comm. 6, 7251 (2015)

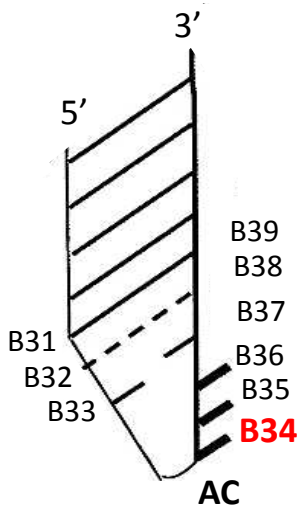


V34 = cmo^{5}U34

S34 = mnm^{5}s^{2}U34

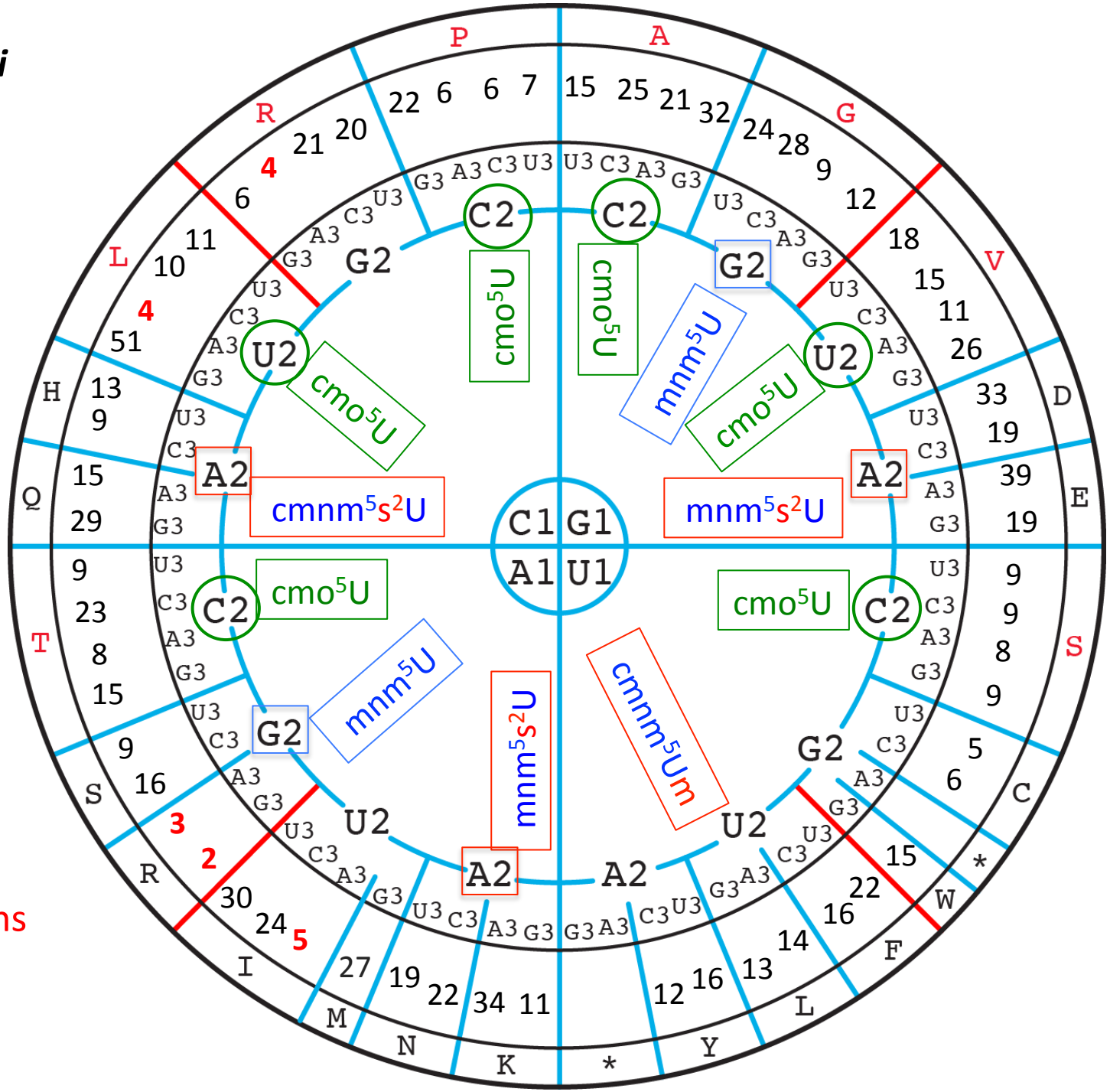
Trends in Biochemical Sciences, September 2016, Vol. 41, P. 798

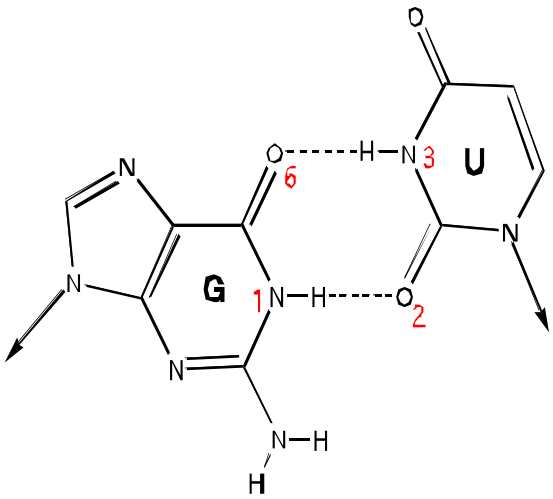
Escherichia coli



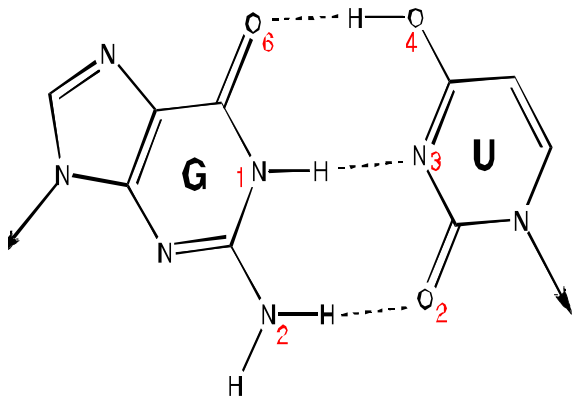
U34
only

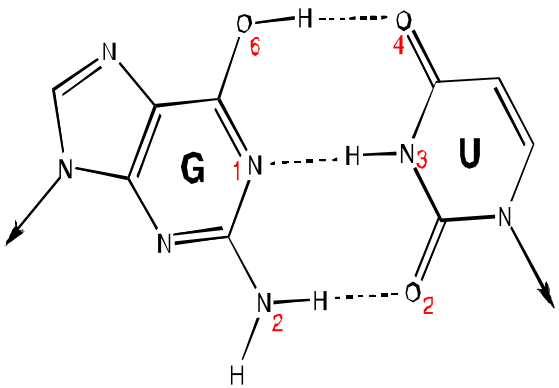
R3-ending codons



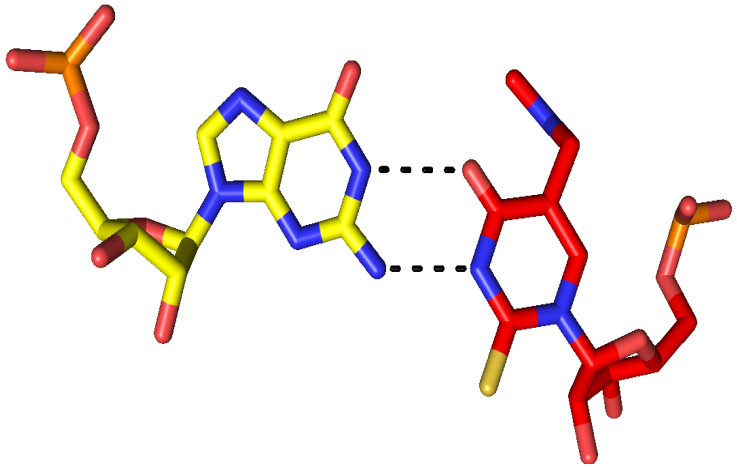


Usual GoU wobble pair
ONLY G34oU(+3)





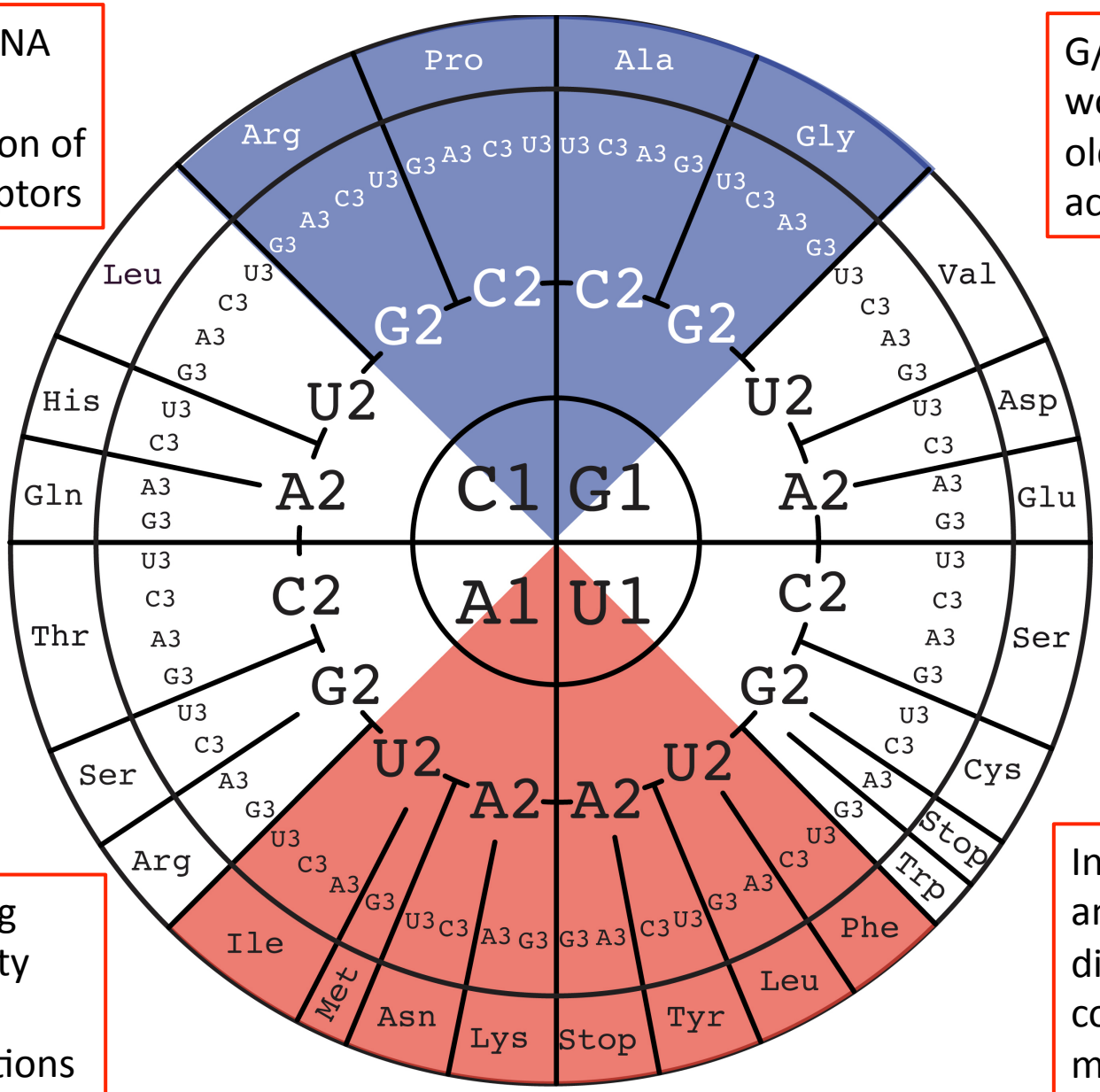
Watson-Crick-like
GoU tautomeric pair



WHEN G(+3)oU34*

Novel GoU wobble pair

Coevolution of the genetic code with metabolic pathways



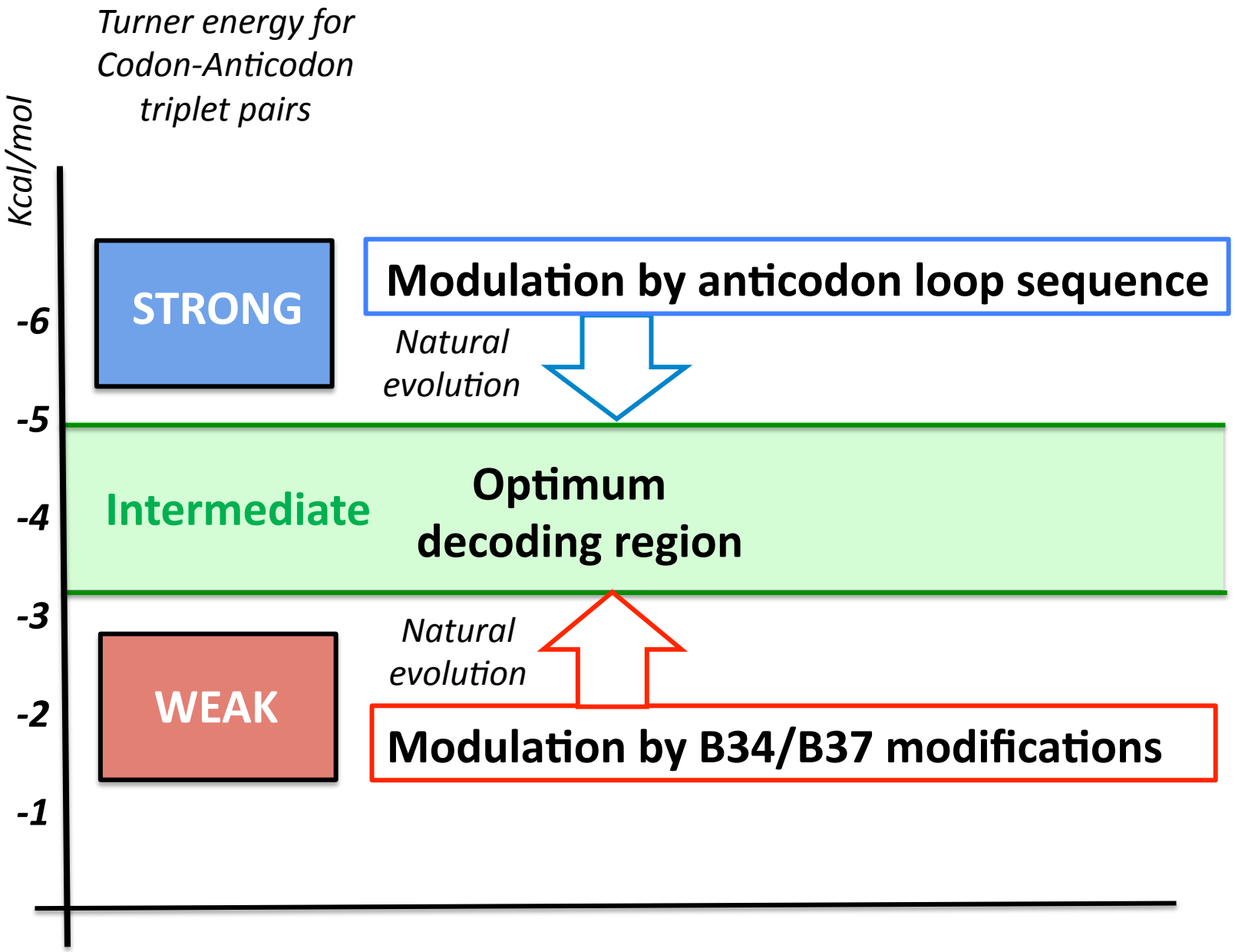
Increasing stability of network interactions

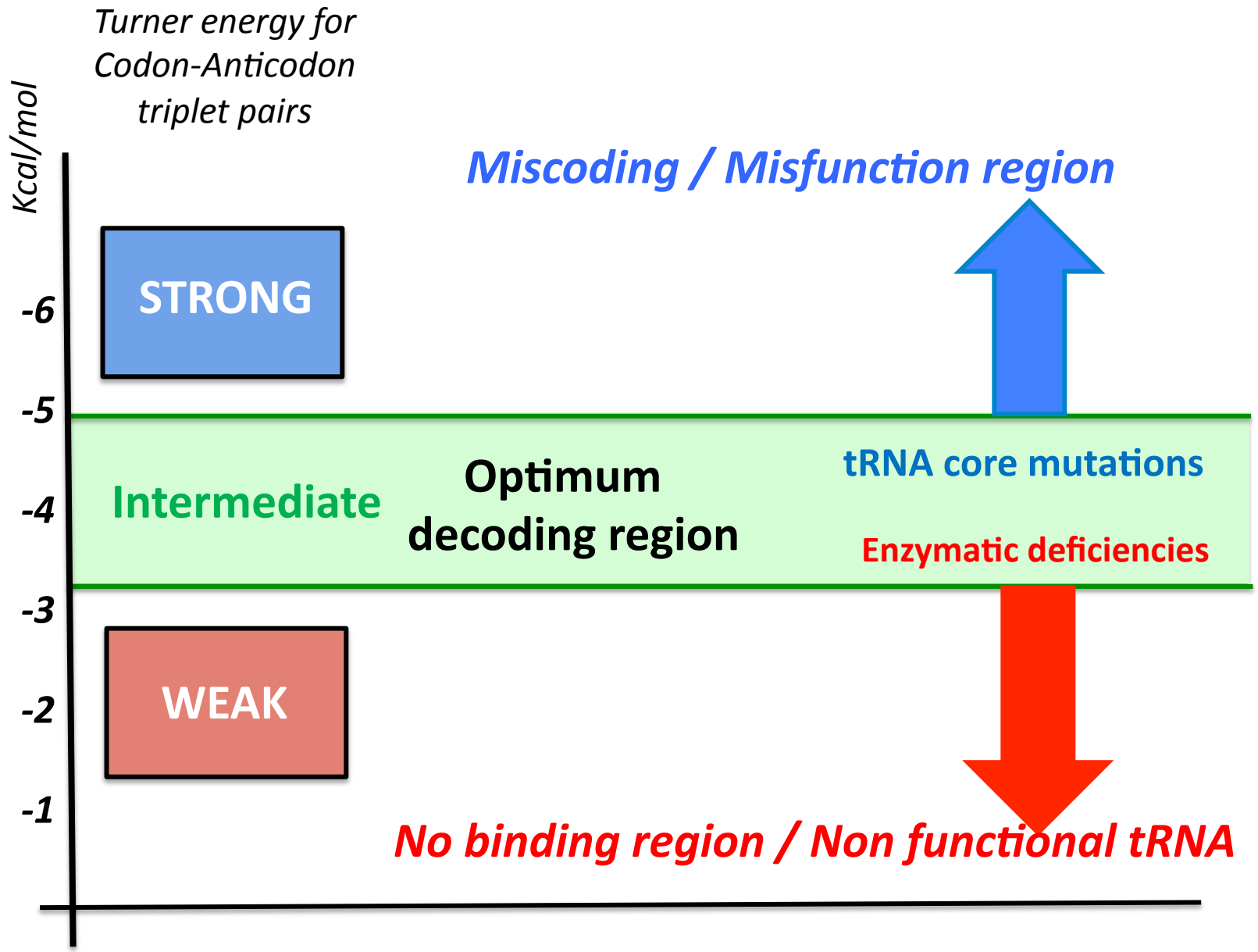
Strong tRNA binding
Modulation of helix adaptors

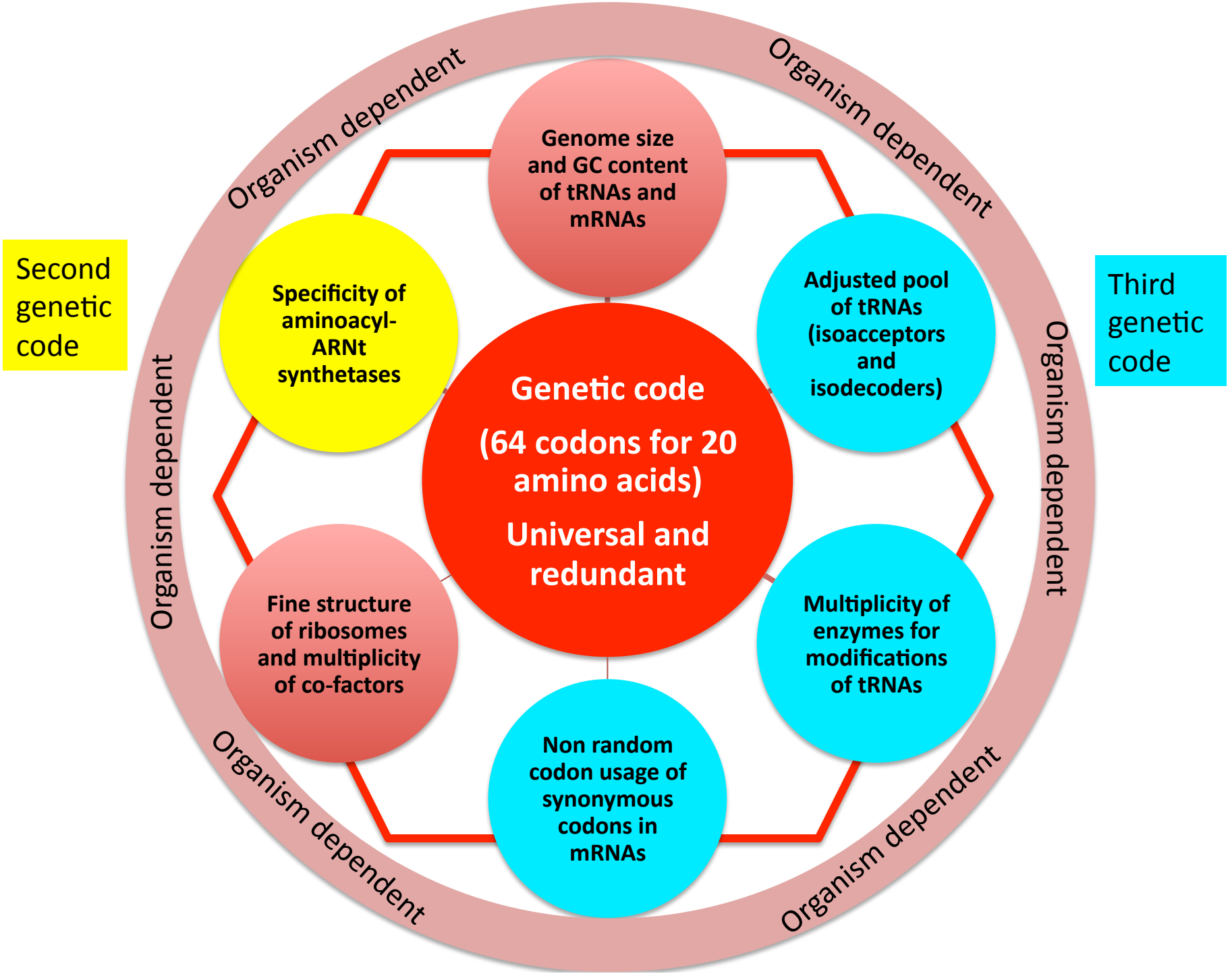
Increasing complexity of modifications at 34 & 37

G/C-rich RNA world with oldest amino acids

Increase in A/U and amino acid diversity coupled with metabolic enzymes


Turner energy for
Codon-Anticodon
triplet pairs
Kcal/mol
-6
-5
-4
-3
-2
-1
STRONG
Modulation by anticodon loop sequence
Natural
evolution
Intermediate
Optimum
decoding region
Natural
evolution
WEAK
Modulation by B34/B37 modifications


Turner energy for
Codon-Anticodon
triplet pairs
Kcal/mol
-6
-5
-4
-3
-2
-1
STRONG
Miscoding / Misfunction region
Intermediate
Optimum
decoding region
tRNA core mutations
Enzymatic deficiencies
WEAK
No binding region / Non functional tRNA


Second genetic code
Third genetic code
Organism dependent
Organism dependent
Organism dependent
Organism dependent
Organism dependent
Organism dependent
Genome size and GC content of tRNAs and mRNAs
Specificity of aminoacyl-ARNt synthetases
Adjusted pool of tRNAs (isoacceptors and isodecoders)
Genetic code
(64 codons for 20 amino acids)
Universal and redundant
Fine structure of ribosomes and multiplicity of co-factors
Multiplicity of enzymes for modifications of tRNAs
Non random codon usage of synonymous codons in mRNAs

Henri Grosjean

(Gif-sur-Yvette)

Crystallography:

Gulnara & Marat Yusupov

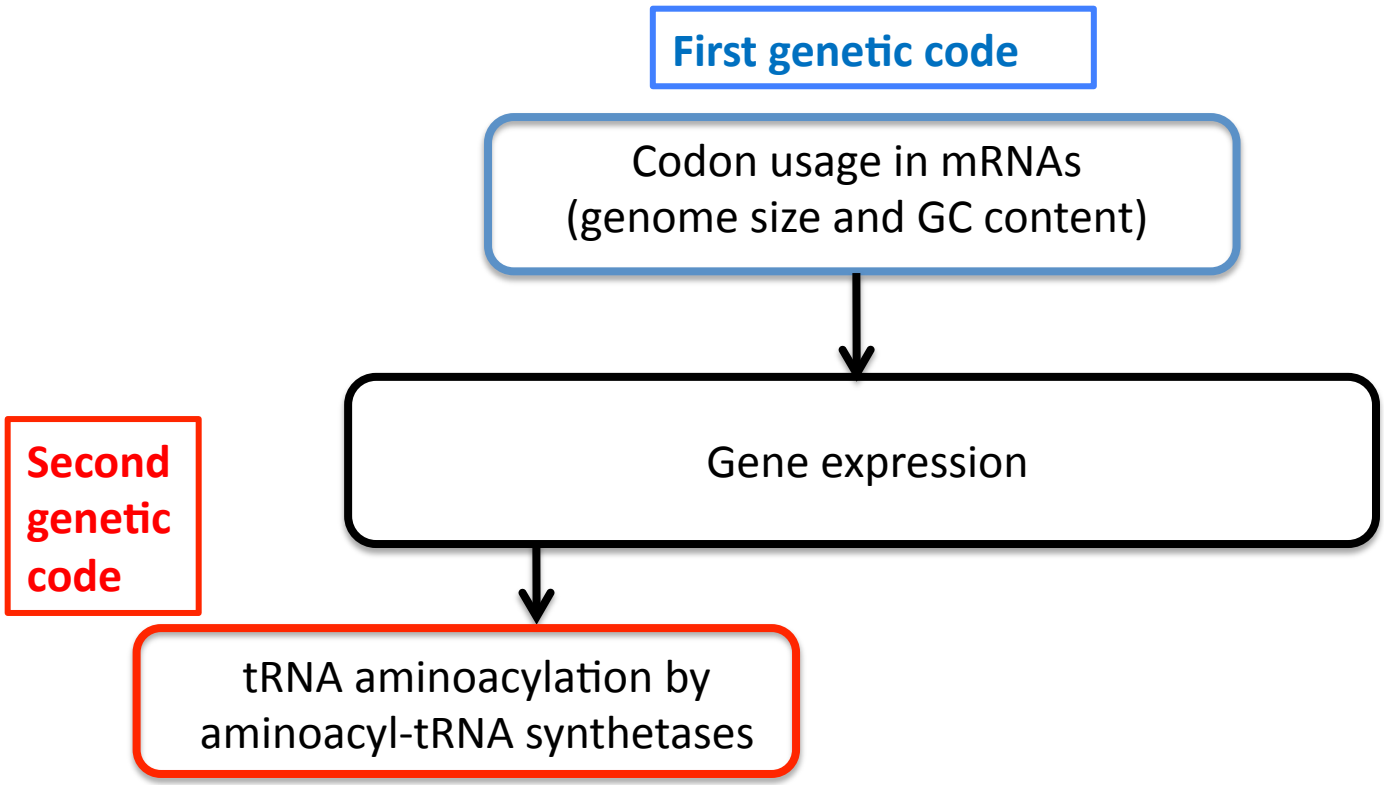
N. Demeshkina, L. Jenner, A. Rozov

(IGBMC-Strasbourg)

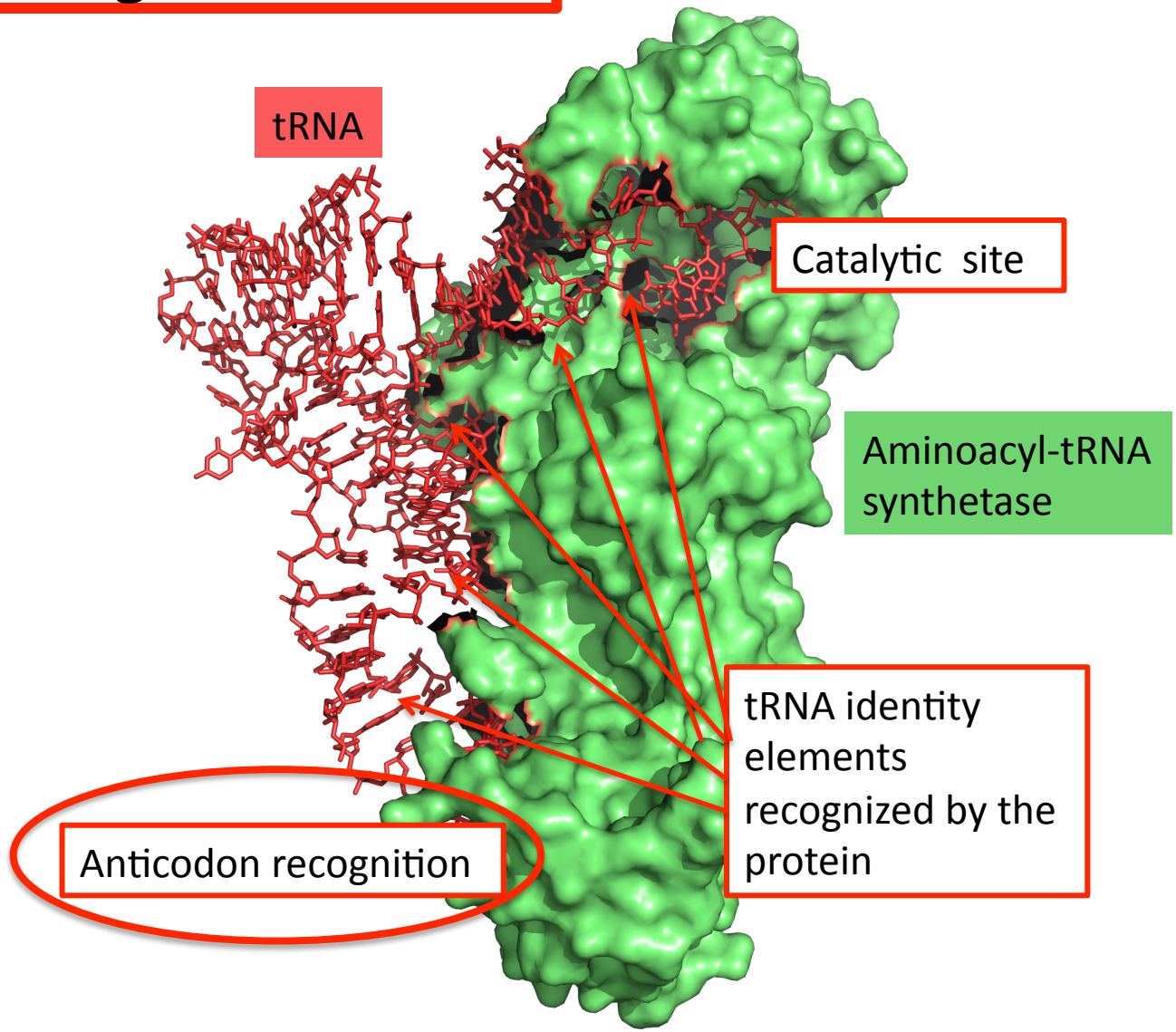
Henri Grosjean and Eric Westhof

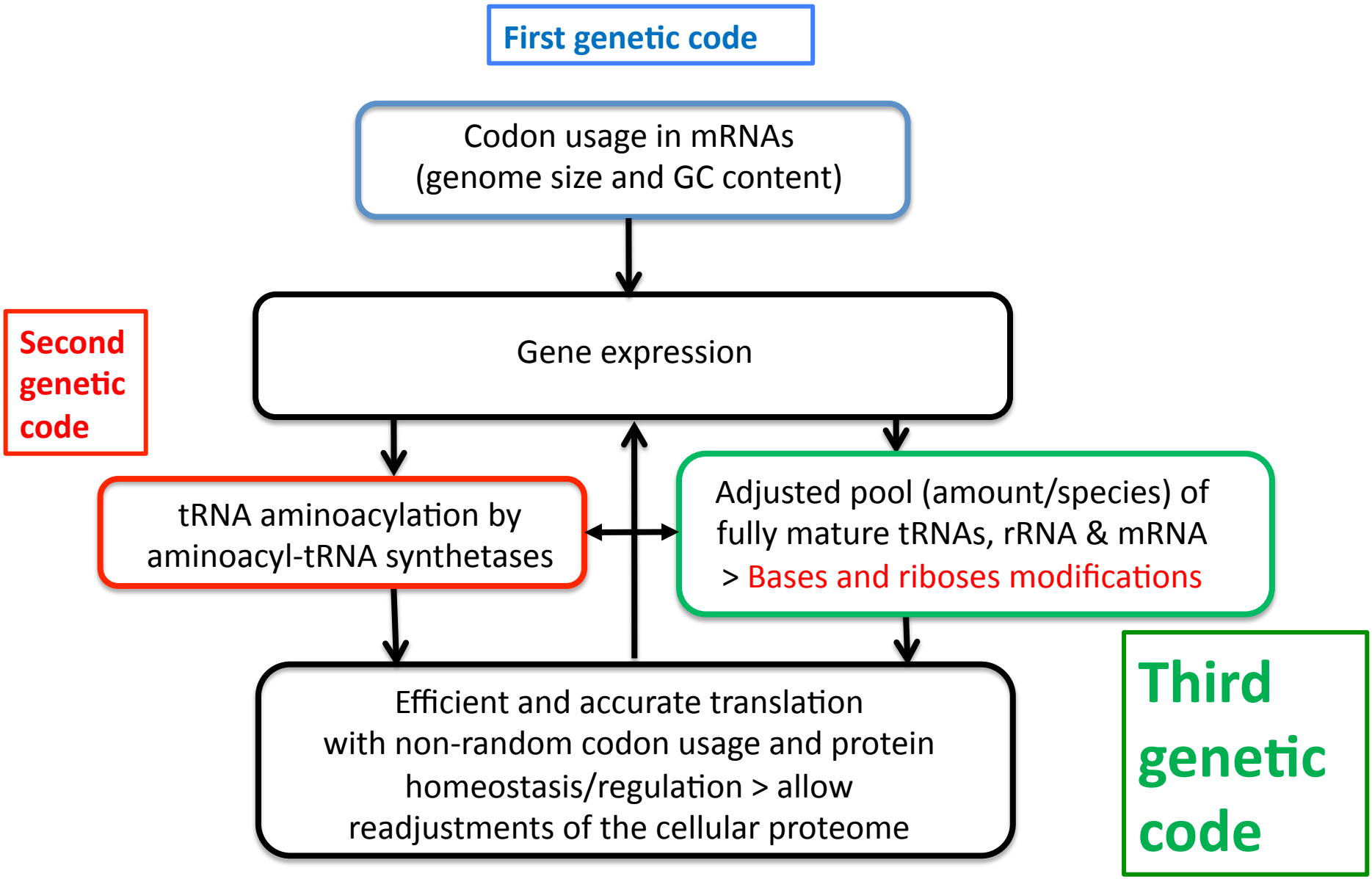
An integrated, structure- and energy-based view of the genetic code

Nucleic Acids Res. 44, 8020-8040 (2016).

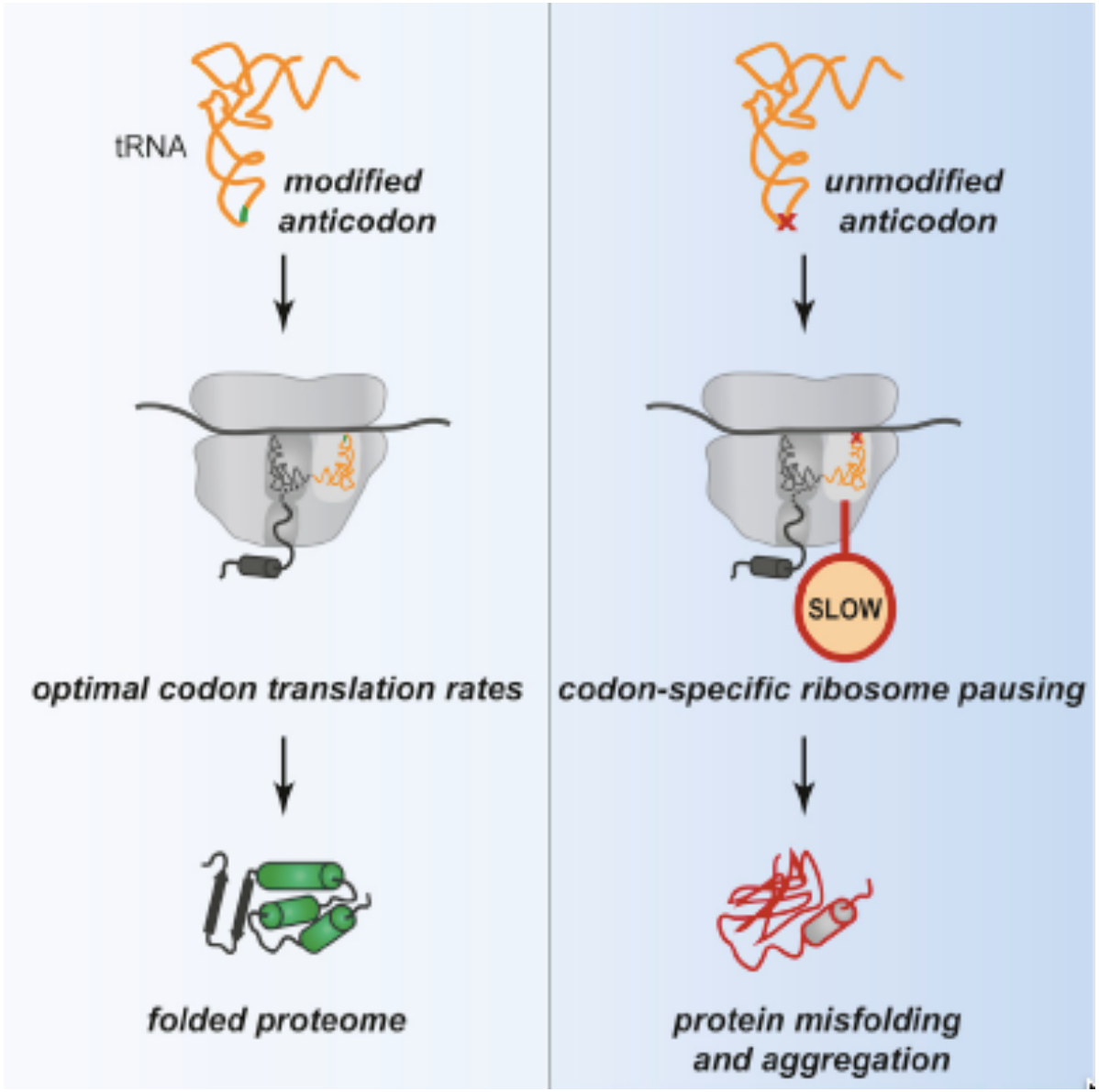

First genetic code
Codon usage in mRNAs
(genome size and GC content)
Gene expression
Second genetic code
tRNA aminoacylation by
aminoacyl-tRNA synthetases

The second genetic code



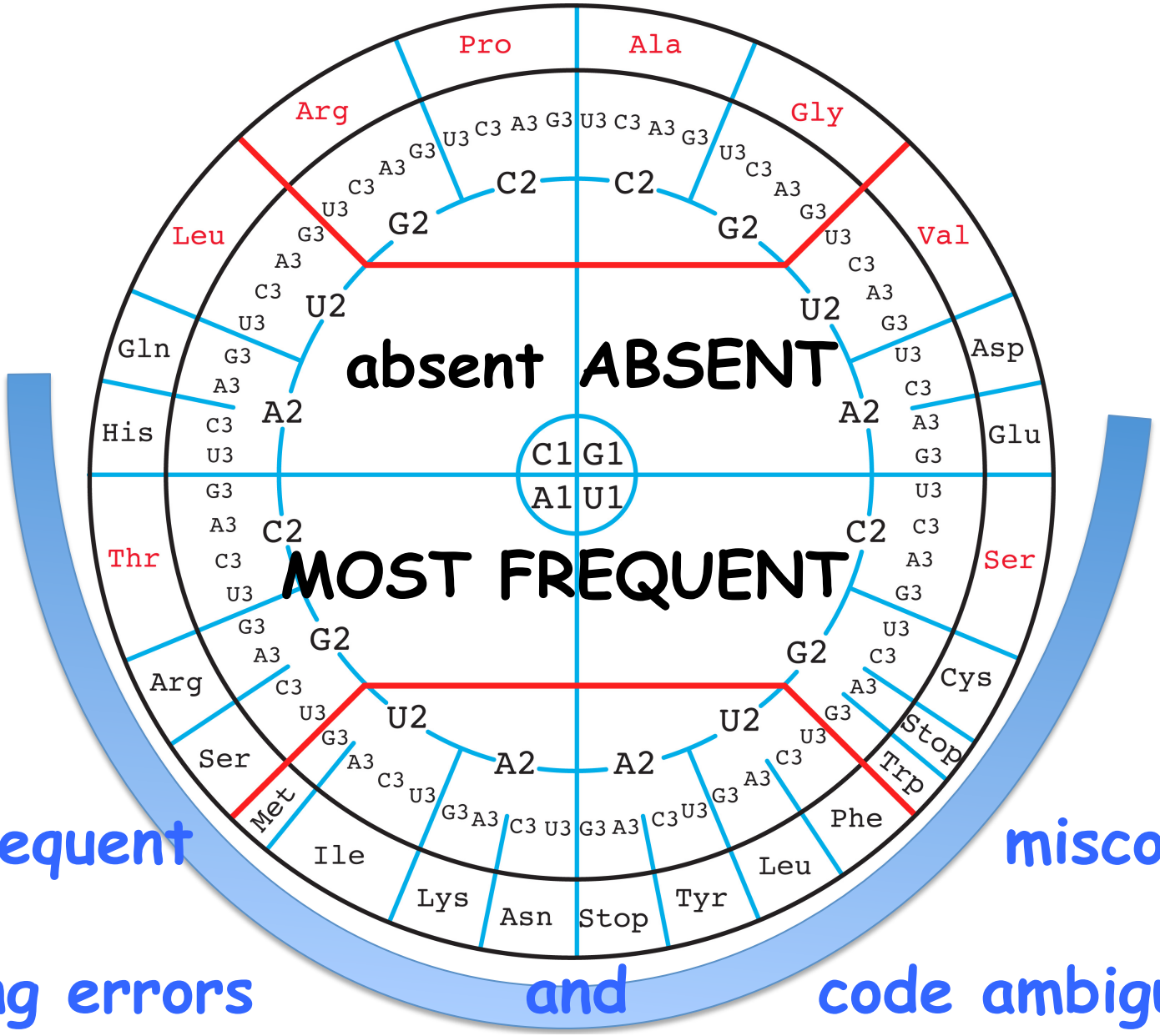

First genetic code
Codon usage in mRNAs
(genome size and GC content)
Gene expression
Second genetic code
tRNA aminoacylation by aminoacyl-tRNA synthetases
Adjusted pool (amount/species) of fully mature tRNAs, rRNA & mRNA
> Bases and riboses modifications
Efficient and accurate translation with non-random codon usage and protein homeostasis/regulation > allow readjustments of the cellular proteome
Third genetic code

Optimization of Codon Translation Rates via tRNA Modifications Maintains Proteome Integrity



Nedialkova & Leidel, 2015, Cell *161*, 1606–1618

Code deviations, recoding, codon reassignments



Most frequent miscodings

(decoding errors and code ambiguities)

UPR 9002 du CNRS,
Architecture et Réactivité de l'ARN,
Institut de Biologie Moléculaire et Cellulaire,
Université de Strasbourg

Henri Grosjean (Gif)

Gulnara & Marat Yusupov (IGBMC)

N. Demeshkina, L. Jenner, A. Rozov (IGBMC)

M.Ryckelynck, C. Rick, K. Pernod, F. Martin (IBMC)