Complete Student Notes for BIOL2202

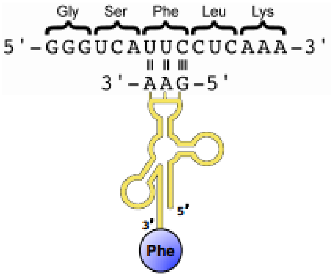
Revisiting Translation & the Genetic Code

Overview

- How tRNA molecules interpret a degenerate genetic code and select the correct amino acid
- tRNA structure: modified bases, functional domains, common fit
- Experimental evidence for the triplet code
- Degeneracy (redundancy) and variations of the standard genetic code

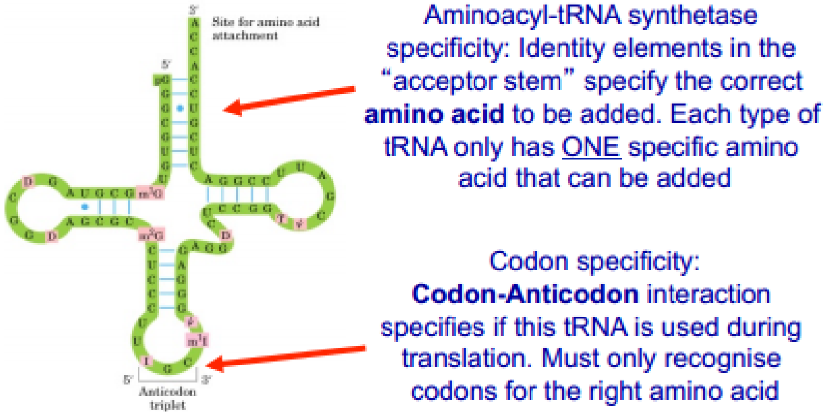
A, P, E site of the Ribosomes

Transfer RNAs (tRNAs)

- Adapters between amino acids and mRNA codons
- Interface of nucleirc acids and proteins
- Transcription→nucleic acids of DNA to RNA (same language but different dialect)
- Translation→RNA to proteins
- tRNA are the translators
- Anticodon of the tRNA base pairs with codons of mRNA
- "Cloverleaf" secondary structure
- Phe is the protein component
    - Covalently attached
- Each amino acid has at least a tRNA
- The amino acid is covalently attached to the 3' end of the tRNA
- Each amino acid has at least one (often more) tRNA gene that encodes tRNAs used specifically for that amino acid
- 

Specificity of tRNAs

- tRNA molecules exhibit TWO levels of specificity that are necessary to accurately adapt a trinucleotide sequence in mRNA to the correct amino acid



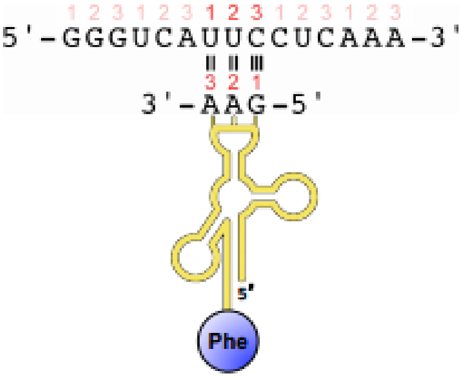
Aminoacyl-tRNA synthetase specificity: Identity elements in the "acceptor stem" specify the correct **amino acid** to be added. Each type of tRNA only has ONE specific amino acid that can be added

Codon specificity: **Codon-Anticodon** interaction specifies if this tRNA is used during translation. Must only recognise codons for the right amino acid

- If you bring the wrong aminoacid then it will be almost the same as a mutation →that is why the specificity is needed
- Anticodon loop →contains the anticodon, three nucleotides of the RNA that recognizes at the A side of the codon

The Anticodon Loop

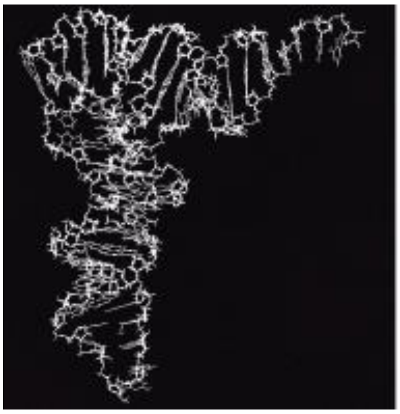
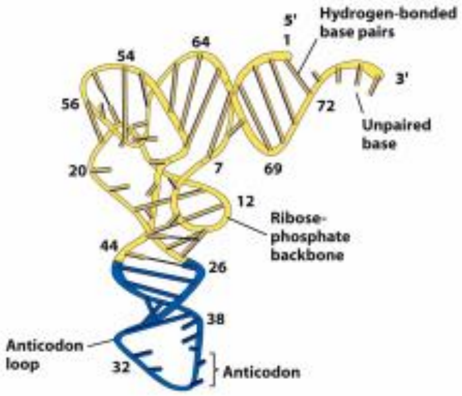
- The **anticodon** reads the genetic code in mRNA
- 3 bases complementary to mRNA codons, written 3'→5'
  - Most nucleotides sequences are written 5'→3'
- 3rd position on codon = 1st position on anticodon
- Trna codon is written from 3' to 5'
- **Interact via hydrogen bonds**
- Often contains a modified nucleotide in or immediately adjacent to anticodon



Abnormal Bases in tRNAs

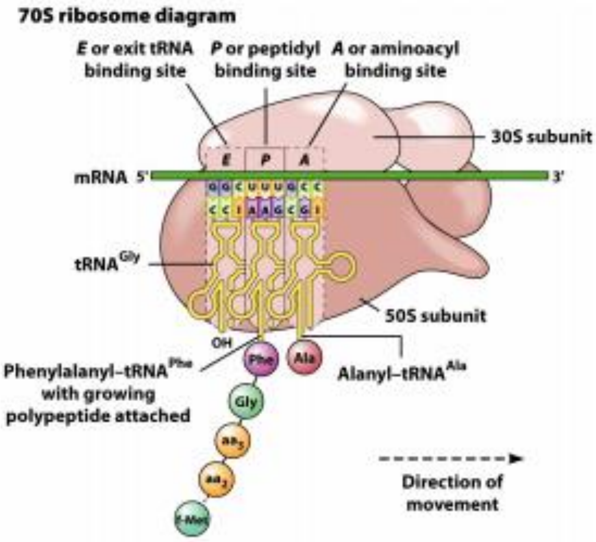
- tRNAs contain many abnormal bases/nucleotides that are created via post-transcriptional modification
- Modified bases allow tRNA greater flexibility of function that possible with only adenine (A), uracil (U), guanine (G) or cytosine (C)
- ~80 known
- Best known is inosine (I)
  - Created by deanimation of adenine into hypoxanthine
  - Forms hydrogen bonds wih C, A, and U, but not G → (**H**)
- Trna is critical for living things to live
- Aminoacids mutate the tRNA and introduce abnormal letter to the Trna molecule

Tertiary tRNA Structure

- A common structure enables tRNAs to fit into ribosomes
- Alternative bases changed by aminoacids
- Modified bases assist tertiary structure formation
- 
  

- Bacterial tRNA are very similar to eukaryotes, they have a selective pressure despite the evolution facts → this is because the production of proteins is vital

Polypeptide Chain Elongation



Stop codons when Serine is at the protein chain, then a water molecule is added

50s subunit is the large subnit

P site →where you get the petide bond

A site →anticodon and codon interactions

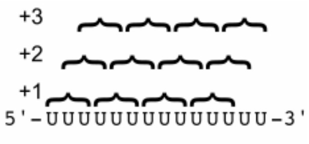
E site → exit site

Evidence of a Triplet Code: In Vitro Translation Studies

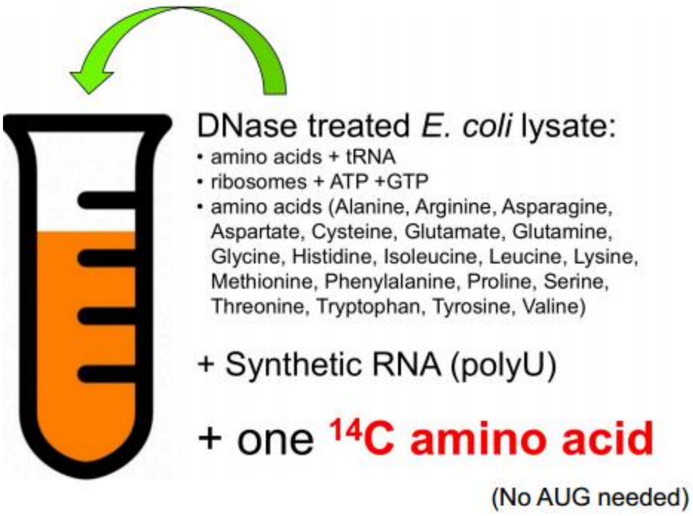
1. Trinucleotide binding of aminoacyl-tRNAs to ribosomes
2. Cell –free (in vitro) protein synthesis (Neiremberg & Matthaei)
    a. Protein synthesis requires
        i. Amino acids +tRNA
        ii. Ribosomes + ATP + GTP
        iii. mRNA template
    b. Grind E.coli cells with powdered alumina to give a cell lysate
    c. Centrifuge to pellet the cell wall
    d. Collect cell lysate (includes tRNA, ribosomes, enzymes and cofactors)
    e. Treat with DNase to stop de novo RNA synthesis
    f. Existing cellular RNA will degrade (RNA is relatively unstable)
    g. Ad synthetic RNA template + one ^{14}C labelled amino acid

Deciphering the Simplest Codons

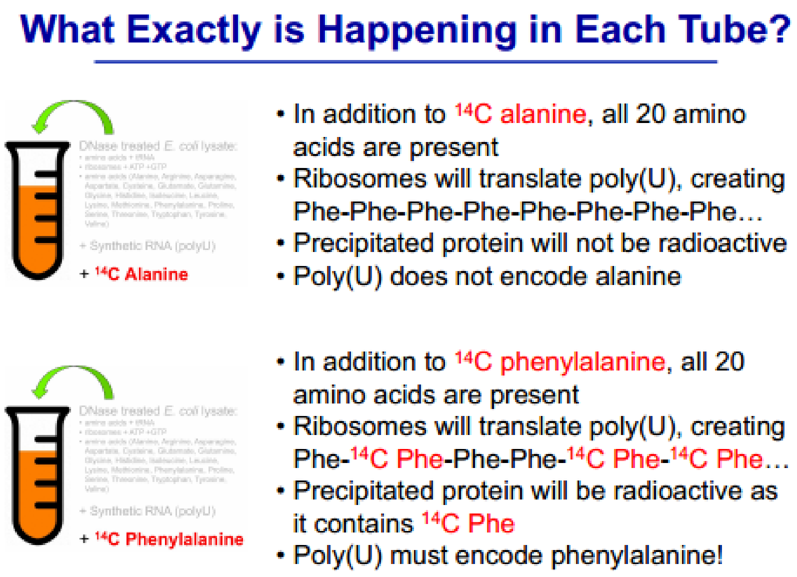
- Used poly(U) mRNA template with combinations of all 20 amino acids plus one ^{14}C radioactively labelled
    - 20 separate experiments conducted in parallel
- What combinations give ^{14}C incorporated into protein ?
    -



    -



-
- AUG start codon
- Out of all the 20 tubes, there was one that contained radioactivity and that UUU coded for the Phenylalanine
-



What Exactly is Happening in Each Tube?

DNase treated *E. coli* lysate:

+ Synthetic RNA (polyU)

+ ^{14}C Alanine

- In addition to ^{14}C alanine, all 20 amino acids are present
- Ribosomes will translate poly(U), creating Phe-Phe-Phe-Phe-Phe-Phe-Phe-Phe…
- Precipitated protein will not be radioactive
- Poly(U) does not encode alanine

DNase treated *E. coli* lysate:

+ Synthetic RNA (polyU)

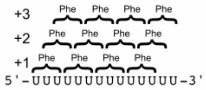
+ ^{14}C Phenylalanine

- In addition to ^{14}C phenylalanine, all 20 amino acids are present
- Ribosomes will translate poly(U), creating Phe-^{14}C Phe-Phe-Phe-^{14}C Phe-^{14}C Phe…
- Precipitated protein will be radioactive as it contains ^{14}C Phe
- Poly(U) must encode phenylalanine!

-

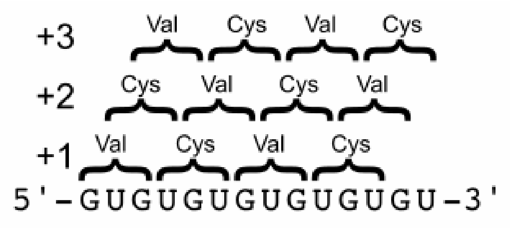
Deciphering the Simplest Codons

- Used poly(U) mRNA template with combination of all 20 amino acids plus one ^{14}C radioactively labelled
    - 20 separate experiments conducted in parallel
- What combinations give ^{14}C incorporated into protein?
- Only poly (^{14}C Phe) produced, so UUU= phenylalanine
-



Deciphering the More Difficult Codons

- Use randoms heteropolymers such as random U+ G, but data difficult to interpret 2^3 possible codons
- Used a range of synthetic defined RNA heteropolymers such as $(GU)_n$ – 2 possible codons, but how to interpret?
-

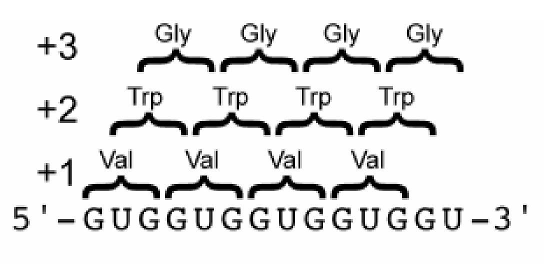


^{14}C Val incorporated, so either **GUG** or UGU = valine
^{14}C Cys incorporated, so either GUG or **UGU** = cysteine

-
- Adding GU and two codons where produced

So is valine encoded by GUG or UGU?

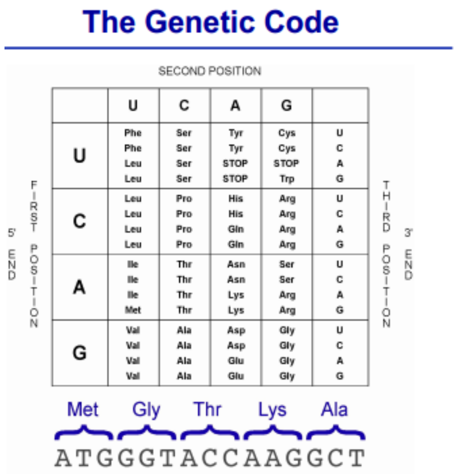
- Use some different synthetic defined RNA hetero-polymers. Let’s try $(GUG)_n$ – 3 possible codons



^{14}C Val incorporated, so either **GUG**, UGG or GGU = valine
^{14}C Trp incorporated, so either GUG, **UGG** or GGU = tryptophan
^{14}C Gly incorporated, so either GUG, UGG or **GGU** = glycine

Combine the data

- Experiment 1
  - GUG = valine or cysteine
  - UGU =valine or cysteine
- Experiment 2
  - GUG= valine, tryptophan or glycine BUT NOT cysteine
  - UGG= valine, tryptophan or glycine BUT NOT cysteine
  - GGU= valine, tryptophan or glycine BUT NOT cysteine
- When the two data sets are combined, logical conclusions can be reached
  - GUG does not encode cysteine
  - Therefore, GUG must encode valine
  - Furthermore, UGU must encode cysteine



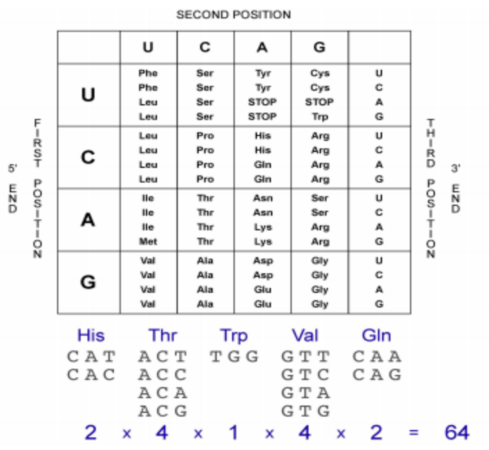
The Genetic Code

SECOND POSITION

5' END FIRST POSITION

	U	C	A	G	
U	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr STOP STOP	Cys Cys STOP Trp	U C A G
C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	U C A G
A	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	U C A G
G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly	U C A G

THIRD POSITION 3' END



SECOND POSITION

5' END FIRST POSITION

	U	C	A	G	
U	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr STOP STOP	Cys Cys STOP Trp	U C A G
C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	U C A G
A	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	U C A G
G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly	U C A G

THIRD POSITION 3' END

- Learn what protein is
- Use table

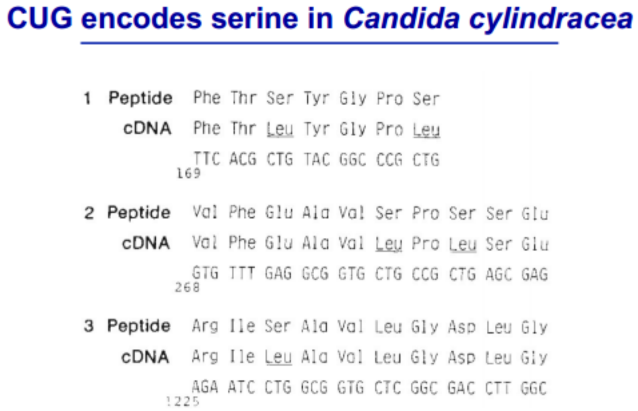
About the Genetic Code

- Initiation and termination codons

  - Initiation (start) codon: AUG as start has a dedicated tRNA. Met elsewhere uses a different tRNA molecule
  - Termination (stop) codons: UAA, UAG, UGA
- Degeneracy: most amino acids are specified by two to four codons. Most degeneracy is in the 3rd position
- Similar amino acids specified by related codons
- Comparison of nucleotide and protein sequence reveals the genetic code is nearly universal; with minor exceptions, the 64 triplets are interpreted in the same way in all organisms

Deviations from the Standard Genetic Code

- **Prokaryotes can use GUG and UUG as alternatives to AUG as a start (fMET) codon, but only at start**
- AUA encodes methionine instead of isoleucine in the mitochondria of vertebrates, *Drosophila* and some fungi
- UGA encodes tryptophan instead of STOP in the mitochondria of vertebrates, *Drosophila*, some fungi and trypanosomes
- AGA and AGG encode serine in *Drosophila* mitochondria and STOP in the mitochondria of vertebrates and protozoa instead of arginine
- CUG encodes serine instead of leucine in some Candida species including Candida albicans
-



CUG encodes serine in *Candida cylindracea*

1	**Peptide**	Phe	Thr	Ser	Tyr	Gly	Pro	Ser			
	cDNA	Phe	Thr	Leu	Tyr	Gly	Pro	Leu			
	169	TTC	ACG	CTG	TAC	GGC	CCG	CTG			
2	**Peptide**	Val	Phe	Glu	Ala	Val	Ser	Pro	Ser	Ser	Glu
	cDNA	Val	Phe	Glu	Ala	Val	Leu	Pro	Leu	Ser	Glu
	268	GTG	TTT	GAG	GCG	GTG	CTG	CCG	CTG	AGC	GAG
3	**Peptide**	Arg	Ile	Ser	Ala	Val	Leu	Gly	Asp	Leu	Gly
	cDNA	Arg	Ile	Leu	Ala	Val	Leu	Gly	Asp	Leu	Gly
	1225	AGA	ATC	CTG	GCG	GTG	CTC	GGC	GAC	CTT	GGC

-

How many tRNA Genes are Needed?

- Each amino acid must have its own dedicated tRNA gene
- Standard genetic code has 64 codons
- 61 of the specify the 20 amino acids. Does this mean that there must be 61 tRNA genes, each encoding a tRNA that recognises only one of these codons?
- E.coli genome encodes ~40 tRNA genes
- S.cerevisiae genome encodes 275 tRNA genes
- Some tRNA must be able to recognise more than one codon
  - How is this possible?

Base-Pairing Involving the 3rd base of the codon is less stringent

- The “Wobble” Hypothesis

    - Proposed by Francis Crick in 1966 to explain how one tRNA molecule could recognise one degenerate codon
    - The hypothesis: the "last base in anticodon (position 1) can "wobble"
    - Extra base-pair options compared to standard A=U and G=C

<u>Wobble Rules</u>

- "Wobble" consequences
    - Minimum 32 anticodons to read 61 codons (31 + 1 start tRNA)
    - ~40 tRNA species in E.coli (8 more than required by Wobble rules)
    - Mitochondria (22 tRNAs) have even more wobble/tRNA
- Use of wobble is limited by the fact that tRNAs must only recognise codons for a single amino acid, otherwise the genetic code would be compromised
-