

# Chapter 18. Weaver, 2/e. Mol Biol X107A.

## Ch 18. The mechanism of translation. II. Elongation and termination.

A. Introduction.....	1
B. Genetic code: overview.....	2
C. Genetic code: degeneracy .....	2
D. Genetic code: (near) universality.....	4
E. tRNA recognition: mRNA codons; wobble.....	4
F. The ribosomal sites.....	5
G. Elongation factors; GTP revisited .....	6
H. Peptide bond formation .....	6
I. Translocation.....	7
J. Terms (sense-related).....	7
K. Suppression.....	8
L. Termination .....	9
M. Breaking the rules (“Recoding”).....	10
N. The final steps: folding and transport -- and degradation.....	11
O. Further reading.....	12
P. Computer resources.....	18
Q. Errata.....	18
R. Homework .....	18
S. Partial answers.....	23

Clark & Russell. See Ch 17 handout.

### A. Introduction

- The genetic code, including use of tRNA.
  - The genetic code, Fig 18.6. It is substantially universal.
  - Recognizing the codon; degeneracy of the code: iso-accepting tRNAs, wobble.
- Unusual aspects.
  - Unusual versions of the code. (minor topic)

- Context dependent effects (“recoding”): frameshifting, readthrough, selenocysteine incorporation, hopping.
- The elongation part of the process includes binding of the new aa-tRNA, peptide bond formation, and translocation.
- Elongation factors and factor mimicry. GTP.
- Suppression.
  - The general idea of suppression.
  - Suppressor tRNAs, nonsense, missense and frameshift.
- Termination: Special codons and factors.

### B. Genetic code: overview

The story of how the genetic code was deciphered is interesting, but not critical for understanding the code. We will discuss some of this history, with an emphasis on the logical points. These can now be understood with the code at hand.

Logical argument: genetic code must have (at least) 3 nucleotides/amino acid (nt/aa), so that 4 nt can code for 20 aa.

Frameshift mutations: certain combinations of 3 frameshift mutations restore gene function. Fig 18.3. This work made a triplet code likely.

Use of poly(U) and other synthetic mRNAs [both random and known (block) sequence]. This revealed some codons, and some patterns. Fig 18.4.

Direct binding of tRNA to single-codon messages. This revealed the entire code. Fig 18.5.

You should be able to read the genetic code table, and to see the effect of single base change mutations. (A copy of the genetic code will be provided on tests when needed.)

Hayes (1998) and Smith (1999) discuss this story from unusual viewpoints. Knight & Landweber (2000) explore how the genetic code might have evolved.

### C. Genetic code: degeneracy

There are 1-6 codons for each amino acid. The existence of multiple codons for an amino acid raises two questions:

- How are they read?

- Are the multiple codons equivalent, or do they play different roles in some way?

#### How multiple codons are read.

- First, there are multiple tRNA species with different codon specificities. (These are called isoacceptors, p 573.)
- Second, an individual tRNA species can read multiple codons, because of “wobble” (Figs 18.7 & 8, and Sect E, below).

#### Role of multiple codons in the cell [not in book]

Are the multiple codons for one amino acid all equivalent? Yes, but... The cell contains different amounts of tRNA for different synonymous codons. If there is more tRNA for a codon, that codon will “work” faster, and perhaps with less error. Given that some codons are “easier” to translate than others, because there is more tRNA, it is reasonable that those codons might be used more frequently. This might be especially important for proteins that are expressed at a high level. This seems to be true.

The pieces of the story fit together reasonably logically: multiple codons, unequal translation ability of the multiple codons, most frequent use of easily translated codons. However, there are different views about cause and effect. For example, one view is that use of a smaller set of tRNAs (i.e., using a small number of preferred codons) allows faster protein synthesis and hence faster growth.

A consequence is that certain mutations that might naively appear to be neutral are not necessarily neutral. A change from one proline codon to another might seem neutral; however, if the change is from a codon that is easy to translate to one more difficult to translate, it might well slow down the production of the protein.

The preferred codons are different in different organisms. As a result, the issue of preferred codons may be a significant point to consider if you are cloning a gene from one organism into another, and want to express it at high level. There is disagreement about the practical importance of the effect. One possible resolution is that the principle is correct, but that it is a practical problem for only a very few codons. There is also evidence that it is more important for codons near the beginning of a protein.

Clouthier et al (1998) is an example of work on rate-limiting tRNAs; they show that rare tRNAs may have natural regulatory roles.

(Those interested in expression of foreign proteins should read the story of rate-limiting tRNAs in depth.)

Sect P lists a source of data on codon usage patterns.

Nakayashiki & Inokuchi (1998) explore why there are so many leucine tRNAs in *E. coli*. This may interact with the story of how multiple codons are used. Saks et al (1998) examine the broader issues of tRNA evolution.

Karlin & Mrazek (2000) use the relationship between codon usage and expression level as a tool in genome analysis -- to predict highly expressed genes.

[There are other reasons, not relating to tRNA availability, why a change between synonymous codons might not be as neutral as expected from just reading the code. A base change may affect the mRNA 2° structure, and this may affect translation. (Ch 17, especially hw.) Another possibility would be that the change affects a splicing signal (Ch 14).]

#### D. Genetic code: (near) universality

The major lesson is that the genetic code is substantially universal. The genetic code is part of a complex, interconnected system. It would be hard to change a codon assignment from one amino acid to another (at least for a major codon).

The rest of this section is “minor”.

However, there are exceptions. Weaver notes some of these in Table 18.1. Some specific code differences are found in mitochondria, and in a few “unusual” organisms. These need not concern us.

Kannan & Baseman (2000) discuss an example of cloning from an organism with a nonstandard genetic code.

Another interesting example of genetic code variability is the coding of the unusual amino acid selenocysteine (Sec) by the UGA stop codon. Within the same organism, UGA can code for both “stop” and Sec. More about such “flexible” coding in Sect K & M, below.

Hohsaka et al (1999) and Böck (2001) discuss developing expansion of the genetic code in the lab.

#### E. tRNA recognition: mRNA codons; wobble

The tRNA reads the message, by base pairing. But A-U & G-C pairs are not the whole story. The tRNA-mRNA interaction is set up physically so that the 3rd base (of the codon) can form acceptable pairs that are different from standard pairs. This is the basis of the so-called wobble.

There are two major contributions to the common wobble rules:

1. Use of G-U pairing (Fig 18.7 & 8).

2. Use of the anticodon nucleotide I (inosine); it recognizes A, U or C (Fig 18.7 and p 573). In one case, even a purine-purine pair, I with A, is allowed. [I is the product of deaminating G.]

The following table summarizes the “standard wobble” pairs that are possible, based on those two phenomena: (Modified from Watson et al, Molecular Biology of the Gene, 4/e, 1987, Table 15-5.)

Base in anticodon, 1st position.	Base(s) in codon, 3rd position, that can pair.	
	normal pairs	novel pairs
G	C	U
C	G	
A	U	
U	A	G
I		A, U, C

You should understand the general idea of the wobble phenomenon, which allows one tRNA to “properly” read more than one codon.

Inosine in the anticodon is an example of tRNA base modification. More about this in Ch 19.

We have already seen how the (bacterial) initiator tRNA can use a variety of codons for initiation (Ch 17). This can be thought of as a wobble-type phenomenon. But it is a distinct phenomenon.

#### F. The ribosomal sites

We refer to two major sites on the ribosome where tRNA is bound (p 576):

- A site. The aminoacyl-tRNA site. (Sometimes called the acceptor site.) This is where the incoming charged tRNA binds.
- P site. The peptidyl-tRNA site. (Sometimes called the donor site.) Prior to peptide bond formation, this site holds the tRNA with the growing peptide chain.

Both sites overlap both the large and small ribosomal subunits, as suggested in Figs such as 18.10. They contact the mRNA on the small ribosomal subunit. The growing peptide chain and the incoming amino acid come together on the large subunit.

These sites were defined logically, before any details of their location or how the growing chain moves from one site to another were determined.

During protein synthesis the growing peptide chain is transferred from the peptidyl-tRNA in the P site to the incoming amino acid on its tRNA in the A site. At some point, the now-uncharged tRNA remaining in the P site leaves the ribosome, and the growing chain attached

to the new tRNA moves to the P site. More about these steps, called translocation, in Sect I, below. We will also mention the E (exit) site there.

### G. Elongation factors; GTP revisited

Recall Ch 17 handout Sect I for an introduction to the role of GTP as an energy source in protein synthesis.

The three major elongation factors (in bacteria) are EF-Tu, EF-Ts, and EF-G. EF-Tu and EF-G use GTP, and EF-Ts is involved in the GTP cycle. The EF-Tu/EF-Ts system has been the focus of much of the work on G proteins.

(EF-G is the key factor involved in translocation; Sect I, below.)

EF-Tu, with GTP bound, “delivers” charged tRNA to the ribosome. Hydrolysis of the GTP causes release of the EF-Tu from the ribosome. In fact, it may well be that this GTP hydrolysis is actually what triggers peptide bond formation to occur; thus the rate of GTP hydrolysis plays an important role in determining the accuracy of protein synthesis. Weaver notes the competition between GTP hydrolysis and dissociation of incorrect aminoacyl-tRNAs (p 585). It is also possible that establishment of a good codon-anticodon interaction triggers GTP hydrolysis.

The initiation factor IF2 (Ch 17) and the elongation factor EF-Tu (Ch 18) play “similar” roles of delivering the tRNA. Of course, they have different specific roles.

The protein EF-Ts helps remove the GDP from EF-Tu. EF-Ts is a nucleotide exchange factor. Such exchange factors are common with mammalian G-proteins. Sprang & Coleman (1998). Also recall p 550.

### H. Peptide bond formation

Peptide bond formation is catalyzed by an enzyme called peptidyl transferase -- by definition. The inability to isolate peptidyl transferase free of ribosomes caused considerable frustration over the years. It is now clear that this “enzyme” is an activity of the 50 S ribosomal subunit. Further, it is virtually certain that peptidyl transferase is an activity of the ribosomal RNA (Figs 18.24 & 25). That is, the ribosome is a ribozyme. Nissen et al (2000) have recently shown that there is no protein within 18Å of the active site of peptidyl transferase, thus eliminating any remaining doubt about the pre-eminent role of RNA as the catalyst in this reaction.

This fits with our idea of the primordial RNA world. The first proteins must have been made by RNA catalysis; apparently, they are still made that way.

## I. Translocation

Translocation refers to the relative motion of the ribosome and the mRNA, one codon at a time. It is the step between rounds of amino acid addition. During this process, the peptidyl tRNA moves from the A site to the P site, and the ribosome moves one codon along the mRNA. The process requires the GTP-dependent factor EF-G. The “discarded” tRNA leaves, apparently through the E site (p 579-80).

It is important to start with a reasonable understanding of the overall process of translocation, before trying to understand any of its details. Bottom line is that the process is not well understood in detail.

Some issues...

We now know that the structures of the two elongation factors, which compete for the same site, are similar (Fig 18.29). Note that one domain of EF-G mimics the structure of the tRNA bound to EF-Tu.

More about proteins competing for binding sites in Sect L.

Coordination of the steps of elongation, peptide bond formation, and translocation is important.

The biggest mystery about translocation is how to move so many things at once. Much current thought suggests that translocation is actually a two-step process, with separate movements on the large and small ribosomal subunits. For example...

Peptide bond formation and translocation may be intertwined. We usually say that the growing peptide chain is transferred to the incoming aminoacyl tRNA in the A site. However, it is possible that the end of the incoming aminoacyl tRNA moves over to the growing peptide chain in the P site. This model is gaining favor (Wilson & Noller, 1998; Frank & Agrawal, 2000). An animation of the proposed translocation model from the latter paper is available at the Nature web site; see Sect P. However, Green et al (1998) place constraints on any model that obligatorily couples peptide bond formation and translocation.

Term. In protein synthesis, “translocation” refers to the process described above. The term also has other meanings in molecular biology and genetics. For example, a genetic rearrangement in which part of one chromosome moves to a different chromosome is called a translocation. The process of moving proteins across a membrane is also called translocation.

## J. Terms (sense-related)

The following terms refer to changes (mutations) in the “sense” of the mRNA.

Mutations that change one amino acid to another are called missense.

Mutations that create a stop codon are called nonsense. We now know that nonsense codons, while not coding for an amino acid, have a specific function: they code for the active process of chain termination (Sect L).

There are three nonsense, or stop, codons. UAG is called the amber codon, UAA is called ochre, UGA is (less commonly) called opal (p 595). An “amber mutant” is a mutant which carries a mutation to the amber codon. How do you know? Because the mutation is suppressed by strains carrying an “amber suppressor.”

### K. Suppression

[Expanded from book, to make the more general point about the nature of suppression. Weaver’s Glossary entries are helpful.]

Suppression is a general term, for a phenomenon, and does not imply mechanism. Suppression refers to a phenotype reversal which is due to something other than a true genetic reversal. Thus a suppressor mutation is a second mutation which (somehow) compensates for (“suppresses”) the first.

Consider a Lac<sup>-</sup> mutant, whose phenotype is due to a mutation in the gene for the enzyme β-galactosidase, *lacZ*. Now isolate revertants that are Lac<sup>+</sup> (phenotype). Possible causes:

- True revertant, reversal of original mutation, so that the revertant carries no mutation.
- Pseudo-revertant, which carries both the original mutation and a second mutation. Phenotype is wild type, Lac<sup>+</sup>, but genotype is doubly mutant. By definition, the second mutation is a suppressor mutation. It suppresses the phenotype of the first.

What is the nature of the suppressor? Possibilities:

- Intragenic. Two mutations in same gene may compensate, leading to an active product. Fig 18.3 is an example; the two mutations affect the reading frame, and they cancel out.

Intragenic suppressors are usually allele-specific. A secondary change that compensates for one defect is not likely to compensate for other defects. Thus, intragenic suppressors can provide information about a protein, as illustrated by Kloser et al (2001).

- Intergenic. A nonsense suppressor is an example. The original mutant carries a nonsense codon, the second mutation now provides a tRNA that can read the nonsense codon, insert an acceptable amino acid, and allow the chain to proceed.

The simplest type of suppressor tRNA has an altered anticodon, Figs 18.34 & 35. However, anticodon mutations are not the only way to make a suppressor tRNA. The real issue is anticodon specificity. I will show an example of a suppressor with a mutation elsewhere.



In some cases, tRNA suppressors can also be found for missense mutations and frameshifts. Hohsaka et al (1999) make use of frameshift-suppressing tRNAs, which read 4-base codons.

(There are other possibilities for intergenic suppression. For example, a quite different enzyme to hydrolyze lactose might appear, from quite another gene. Guigueno et al, 2001, explore an example involving the toxicity of unfolded proteins.)

Björkman et al (2000) explore suppression in a practical context.

## L. Termination

An interesting feature of termination follows from the way energy is provided for peptide bond formation.

In nucleic acid synthesis, the high energy bond comes in with the base being added. In protein synthesis, the high energy bond comes from the chain being extended. No great principle involved; both ways work. However, there is an implication for the termination reaction. A growing peptide chain always carries a tRNA at its growing end; thus, there is a special termination reaction which removes the terminal tRNA.

Puromycin (Fig 18.12) is a commonly used inhibitor of bacterial protein synthesis. It resembles an aminoacyl-tRNA; it can bind to the A site and can be transferred to the growing chain by the peptidyl transferase reaction. However, because it is so small, it falls off during translocation, taking the nascent chain with it.

Fig 18.12 may be misleading. The puromycin structure shown there is the entire molecule; the tyrosyl-tRNA has another 75 or so nucleotides (where it says "tRNA" near upper left).

We mentioned earlier that the structures of EF-Tu and EF-G, which compete for a common binding site, are similar (Fig 18.29). Release factors, which also compete for the same site, have a similar structure (p 598 and Fig 18.39). The simple view is that RF-3 is a co-factor for both RF-1 and RF-2, and is not itself codon specific.

Kannan & Baseman (2000) deal with the competition between a release factor and a tRNA.

Weaver discusses the importance of the ribosome cycle, at the start of Ch 17. But the overall process is more complex than he suggests. Karimi et al (1999) clarify details of how ribosomes are recycled following termination. Their "Fig 6" is attached, p 22. Selmer et al (1999) show that RRF (= ribosome release factor) is also a tRNA mimic. Also see Ohnishi et al (1999).

Spahn et al (2001) describe yet another protein that competes for the same site on the ribosome. Burley & Roeder (1998) discuss a different example of protein-nucleic acid mimicry, from Ch 11.

M. Breaking the rules (“Recoding”)

[not in book]

[There is an amusing ambiguity of word usage. Weaver has a section subheading called “Breaking the code”; here, “break” means “to solve”. I have a section on “Breaking the rules”; here, “break” means “to violate”. These are both proper uses of the word -- along with dozens of other meanings!]

At this point the “rules” for making proteins are clear enough. DNA → RNA → protein. And the final step involves sequential translation of base triplets to amino acids via the genetic code. However, it doesn’t always work that way. In some cases, the mRNA intermediate is spliced or otherwise edited (Ch 14, 16). In other cases, the mRNA is not translated according to the “rules.” Instead, the simple linear information in the message is “recoded” to a nonstandard meaning during translation.

These recoding events are quite rare, but some are natural; some essential proteins depend on them. Further, the events are “predictable” (even if we usually do not understand enough to predict them); they seem to depend on specific “context”, such as mRNA secondary structure and tRNA availability.

Many types of recoding events have been identified. My intent here is to make you aware of this area, with little detail. Here are some nonstandard translational events...

Frameshifting. Sometimes translocation occurs by a distance other than three bases. A two base translocation results in a “-1” frameshift; a four base translocation results in a “+1” frameshift. Frameshifts are “common” at certain sites, presumably due to some combination of mRNA secondary structure and perhaps an odd tRNA. Mejlhede et al (1999) discuss an example of natural frameshifting. Also see Clouthier et al (1998). Hohsaka et al (1999) make use of frameshift suppressors.

Readthrough of termination codons (natural suppression). UGA codons are the least efficient natural stop codons. Most commonly, UGA readthrough inserts Trp.

[One way or another, retroviruses typically require a rule break for proper translation. “gag-pol” is a fusion protein, which is cleaved to yield gag (a coat protein) and pol (polymerase). Translation of the gag-pol message always yields gag, but yields pol only when a rule is broken to avoid normal termination. Since pol is essential, this infrequent event is essential to the virus. The rule break is a readthrough for some viruses, a frameshift for others.]

Selenocysteine. The coding of the unusual amino acid selenocysteine by the codon UGA is a special case of readthrough. In bacteria, SeCys incorporation requires a special factor, which appears to be an analog of EF-Tu. This SelB protein can deliver only the tRNA<sup>SeCys</sup>, and it is

recruited by binding to the mRNA at a specific 2° structure. The system in eukaryotes is probably similar but more complex (Atkins & Gesteland, 2000).

[Another interesting aspect of SeCys incorporation is that the tRNA is charged with Ser; the Ser is subsequently modified to SeCys while on the tRNA. This is unusual, but not completely unprecedented. Some organisms charge tRNA<sup>Gln</sup> with Glu, then amidate the Glu while on the tRNA (e.g., Tumbula et al, 2000).]

Hopping. Sometimes a ribosome will “hop” from one codon to another identical (or similar) codon elsewhere in the mRNA. The new codon may or may not be in the original reading frame. Most hops are one to a few bases along the mRNA. However, one case is well characterized where the synthesis of a normal protein requires skipping 50 bases in the mRNA. This hop event is influenced by mRNA secondary structure, and also by the folding of the nascent peptide (Maldonado & Herr, 1998).

Vogel (1998) describes another unusual way to make proteins.

True & Lindquist (2000) suggest that some yeast prions are an evolutionary strategy to allow flexible readthrough.

#### N. The final steps: folding and transport -- and degradation

⇒ You are not responsible for this section.

In the steps discussed so far, the gene sequence has been translated to produce the corresponding protein chain. Production of a functional protein requires some additional steps. These always include folding of the protein, perhaps aided by chaperones. The final steps may also include transport of the protein to the proper cellular (or extra-cellular) location, and modifications of the protein, such as glycosylation or cleavage. Actually, the final final step is degradation -- perhaps because a protein was made or folded wrong, or because it has served its purpose.

These topics are not in Weaver (except for occasional mention in passing), and are beyond this course. (Some were mentioned briefly in Ch 3 handout, along with some relevant FR.)

Bukau et al (2000) is a good introduction to protein folding, and chaperones.

Song et al (2000; Ch 17 FR) discuss a key part of secretion in eukaryotes, the role of the signal recognition particle.

Schubert et al (2000) discuss how much protein degradation is a part of protein synthesis.

Orr & Zoghbi (2000) discuss folding and its corollary, degradation, in the context of a human disease.

Johnson & Haigh (2000) and Pelham & Rothman (2000) review aspects of how proteins are targeted in eukaryotic cells.

If you would like to fold proteins yourself, see Sect P.

The 1999 Nobel Prize in Medicine was awarded to G Blobel, of Rockefeller University, for key proposals and discoveries on how proteins are sorted to get to their proper cellular location. See, for example, *Science* 286:666, 10/22/99.

#### O. Further reading

B Hayes, Computing science: The invention of the genetic code. *Amer Sci* 86:8, 1/98. Discussion of the history of deciphering the genetic code -- from the viewpoint of a computer scientist.

K S Wilson & H F Noller, Molecular movement inside the translational engine. *Cell* 92:337, 2/6/98. A major review, with considerable emphasis on translocation and the elongation factors. (This issue of *Cell* contains several reviews on the broad theme of macromolecular machines. Other topics includes nucleic acid polymerases, splicing, protein folding and degradation, and nucleocytoplasmic transport.)

S C Clouthier et al, tRNA<sup>Arg</sup> (fimU) and expression of SEF14 and SEF21 in *Salmonella enteritidis*. *J Bact* 180:840, 2/98. The idea that rare tRNAs may be rate-limiting for protein synthesis is “intuitive”, but generally backed by little specific evidence. It’s now clear that one or two very rare Arg codons (AGG & AGA) in *E. coli* do limit translation. The details of the story vary somewhat among test systems, but the effect may be more pronounced when the rare codons are early in the gene or are clustered. A simple model is that the rare tRNA is so rare that virtually all of it is sequestered on ribosomes that are using it. The resulting tRNA starvation causes stalling, and may also cause misreading (recoding) events such as frameshifts. Here they show that a rare Arg tRNA is involved in regulation of cell surface structures in pathogenic bacteria. (In other cases, Leu and Ile tRNAs have been shown to be limiting.)

M E Saks et al, Evolution of a transfer RNA gene through a point mutation in the anticodon. *Science* 279:1665, 3/13/98. The availability of complete genome sequences allows them to look at the relationships among gene sequences for the entire complement of tRNAs. This allows them to develop or test models of tRNA evolution. One type of event is an anticodon change, which also changes the amino acid specificity. But clearly the big picture is much more complicated.

R Green et al, Ribosome-catalyzed peptide-bond formation with an A-site substrate covalently linked to 23S ribosomal RNA. *Science* 280:286, 4/10/98. This work makes a number of interesting points about the peptidyl transferase reaction. First, it pins down the location of the peptidyl transferase center. Second, it shows that the peptidyl transferase center is activated (“created”??) by the presence of a peptidyl-tRNA in the P site. Third, that the peptidyl transferase reaction can occur with the A-site substrate covalently bound to the ribosome places constraints on any model in which peptide bond formation is obligatorily coupled to translocation. From Noller at UCSC.

R Maldonado & A J Herr, Efficiency of T4 gene 60 translational bypassing. *J Bact* 180:1822, 4/98. They show that the efficiency of bypass -- of 50 bases -- is about 50%, representing a competition between release and bypass.

A L Arkov et al, An rRNA fragment and its antisense can alter decoding of genetic information. *J Bact* 180:2744, 5/98. They screen a library of rRNA fragments, to find one that has a specific effect on protein synthesis. More in the hw.

T Nakayashiki & H Inokuchi, Novel temperature-sensitive mutants of *Escherichia coli* that are unable to grow in the absence of wild-type tRNA<sup>Leu</sup><sub>6</sub>. *J Bact* 180:2931, 6/98. There are 6 codons for leucine. *E. coli* has 5 tRNAs for leucine, some of which recognize more than one codon. In lab strains of *E. coli*, leucine tRNA #6 (yes, #6 is one of the 5) seems to be non-essential. Here, they isolate mutant strains in which #6 is now essential. One set of these is defective in a tRNA modification, another set is defective in a cell division function. What this all means is unclear. However, it has been reported that one pathogenic strain of *E. coli* requires this tRNA for virulence.

G Vogel, Protein chemistry: A two-piece protein assembles itself. *Science* 281:763, 8/7/98. News. An odd variation of an odd variation of protein synthesis. In some cases, proteins contain intervening sequences that are cut out of the synthesized protein; these regions are known as inteins (cf introns, in RNA; Ch 14). In the case reported here, an enzyme is made in two pieces, which are then joined together. How can that happen? Each half contains half an intein. The two half-inteins interact, creating a functional intein. The intein is removed (as normal for inteins), thus fusing the two pieces together.

S K Burley & R G Roeder, TATA box mimicry by TFIID: Autoinhibition of Pol II transcription. *Cell* 94:551, 9/4/98. Minireview. Protein-nucleic acid mimicry is becoming recognized in more and more cases. In this chapter we noted how part of EF-G mimics the tRNA that would be bound to EF-Tu. Here, they discuss a key eukaryotic transcription factor (Ch 11). Part of it apparently mimics the binding site for the factor on the DNA, thus providing auto-inhibition. (More specifically, part of a TAF inhibits the DNA-binding site of TBP.)

S R Sprang & D E Coleman, Invasion of the nucleotide snatchers: structural insights into the mechanism of G protein GEFs. *Cell* 95:155, 10/16/98. Minireview. EF-Ts is the bacterial example of a GEF = guanine nucleotide exchange factor.

M Ohnishi et al, Molecular cloning, sequencing, purification, and characterization of *Pseudomonas aeruginosa* ribosome recycling factor. *J Bact* 181:1281, 2/99. RRF is a factor that aids in the release of the ribosomes from the mRNA after termination. It is an essential protein in bacteria, though probably not in eukaryotes (where it may be found only in the prokaryote-like organelles); thus the authors suggest that drugs targeted to RRF might be useful antibacterial agents. Also see Karimi et al (1999), below.

N Mejlhede et al, Ribosomal -1 frameshifting during decoding of *Bacillus subtilis* cdd occurs at the sequence CGA AAG. *J Bact* 181:2930, 5/99. An example of "breaking the rules" in translation. In this case, a frameshift occurs about 15% of the time, avoiding a stop codon. As a result, two proteins are made from the same gene. The two proteins have the same activities,

so far as they can tell. As often the case with such natural frameshifts, they find a Shine-Dalgarno sequence near the frameshift site; it seems to be involved in the frameshift event. It is also the rbs for the downstream, overlapping gene; thus the frameshift event would seem to affect this downstream gene.

R Karimi et al, Novel roles for classical factors at the interface between translation termination and initiation. *Mol Cell* 3:601, 5/99. One of the most poorly understood parts of translation is between “termination” and “initiation”. Here, they study the role of RRF in dissociating ribosomes; this process also requires EF-G and GTP hydrolysis. They also show that IF3 helps remove the old tRNA from the released 30S subunit, a prerequisite for the next round of initiation. Their “Fig 6” is attached, p 22.

J M Smith, Too good to be true. *Nature* 400:223, 7/15/99. An essay on the aesthetics of the genetic code -- and how the code was revealed.

B A Neilan et al, Nonribosomal peptide synthesis and toxigenicity of cyanobacteria. *J Bact* 181:4089, 7/99. Some unusual peptides, such as peptide antibiotics, are made by dedicated enzyme complexes. Aside from their natural role, these are of interest in developing novel products.

M Selmer et al, Crystal structure of *Thermotoga maritima* ribosome recycling factor: A tRNA mimic. *Science* 286:2349, 12/17/99. RRF provides another example of protein-RNA mimicry.

T Hohsaka et al, Incorporation of two different nonnatural amino acids independently into a single protein through extension of the genetic code. *J Am Chem Soc* 121:12194-12195, 12/29/99. Not surprisingly, people have attempted to “alter” or even “expand” the genetic code in the lab. One goal is to allow the protein chemist to explore some novel modifications. One approach involves modifying the amino acid on the tRNA, thus incorporating novel amino acids using old codons. Another approach involves using “non-essential” codons, and arranging for them to incorporate unusual amino acids. Here, Hohsaka et al use tRNAs that read four base codons (frameshift suppressors). By using two of these, they are able to incorporate two novel amino acids into a protein. Among the amino acids they used: 2-naphthylalanine and a nitrophenylalanine.

J Björkman et al, Effects of environment on compensatory mutations to ameliorate costs of antibiotic resistance. *Science* 287:1479, 2/25/00. (+ *News, Bull & Levin*, p 1409.) Antibiotic-resistant bacteria often have growth disadvantages. One can select mutants that grow better; these often retain the antibiotic resistance but have acquired secondary (suppressor) mutations, either intragenic or intergenic. Here they explore how the nature of such suppressors depends on the growth conditions.

H T Orr & H Y Zoghbi, Reversing neurodegeneration: a promise unfolds. *Cell* 101(1):1, 3/31/00. Minireview. Huntington’s disease (HD) involves an expanding triplet repeat in the gene, which codes for a polyglutamine tract in the “huntingtin” protein (recall Ch 2). The polyglutamine tract apparently makes the protein neurotoxic. Here they review recent work in a mouse model system showing that continuous expression of the polyglutamine protein is necessary for progressive neurodegeneration. If relevant to the human disease, this obviously has therapeutic implications -- for HD and possibly for other diseases that may involve

accumulation of defective protein (Alzheimer's, Parkinson's, ALS, prion diseases ???). They discuss the possible mechanisms; one idea is that the defective protein clogs the protein degradation machinery.

L R Cruz-Vera et al, Molecular basis for the temperature sensitivity of Escherichia coli pth(Ts). *J Bact* 182:1523, 3/00. The peptidyl-tRNA hydrolase discussed here is not part of the normal translation process, but is apparently a "scavenger" enzyme to recycle tRNAs. It is an essential enzyme in E. coli.

U Schubert et al, Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature* 404:770, 4/13/00. (+ News, Schild & Rammensee, p 709.) The heading for the news article is a good introduction: "When making proteins, cells opt for high speed at the cost of large amounts of waste. But the waste-recycling process has a happy side effect - it gives the immune system a head start in the battle against viral invaders." Estimates are that 30% or so of newly synthesized proteins must be degraded, presumably because of faulty primary sequence or faulty folding.

B Bukau et al, Getting newly synthesized proteins into shape. *Cell* 101:119, 4/14/00. Minireview. Compares prokaryotic and eukaryotic protein folding.

T R Kannan & J B Baseman, Expression of UGA-containing Mycoplasma genes in Bacillus subtilis. *J Bacteriol* 182:2664-2667, 5/00. In Mycoplasma the UGA codon codes for Trp. Cloning genes from organisms with that codon into "normal" organisms that read UGA as "stop" is obviously a problem. Fortunately, Bacillus subtilis has a low but significant level of UGA readthrough as Trp, and it can be enhanced by mutations that reduce the effectiveness of RF2. The issues discussed here are more important than the particular solution.

R D Knight & L F Landweber, The early evolution of the genetic code. *Cell* 101(6):569, 6/9/00. Minireview. Discusses data that suggests that the amino acids might originally have interacted with "codons" directly and specifically.

J Frank & R K Agrawal, A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature* 406:318, 7/20/00. They analyze the structure of the functioning ribosome at various stages of the elongation cycle, by cryo-electron microscopy. Analysis of the structures provides evidence for a ratchet-like rotation of one subunit relative to the other, in translocation. An animation of their proposed translocation model is available at the Nature web site; see Sect P.

P Nissen et al, The structural basis of ribosome activity in peptide bond synthesis. *Science* 289:920, 8/11/00. (+ News article, Cech, p 878, and related articles on p 905 and 947.) A detailed X-ray structure of the 50S ribosome is an incredible technical accomplishment, and provides much for discussion. Of particular interest... There is no protein within 18Å of the active site of peptidyl transferase. This rather persuasively precludes that peptidyl transferase is a protein enzyme, and thus strongly supports the argument that it is a ribozyme. The data also provide the best estimate yet of the size of the exit channel for nascent proteins, and suggest that it varies between 10-20 Å in diameter along its length. This was the basis of Ch 3 hw # 6.

D L Tumbula et al, Domain-specific recruitment of amide amino acids for protein synthesis. *Nature* 407:106, 9/7/00. They examine how the amide amino acids (Asn, Gln) are prepared for protein synthesis in eubacteria, archaea, and eukaryotes. Two general approaches are used: the amino acid is made and then activated (as you might expect), or the non-amide form of the amino acid is activated, then amidated while on the tRNA. Interestingly, they find distinctive variations in each of the three domains of life. Their results are consistent with the notion that Asp and Glu are among the most primitive of the amino acids.

A E Johnson & N G Haigh, The ER translocon and retrotranslocation: Is the shift into reverse manual or automatic? *Cell* 102:709, 9/15/00. Minireview.

H R B Pelham & J E Rothman, The debate about transport in the Golgi - two sides of the same coin? *Cell* 102:713, 9/15/00. Review.

J F Atkins & R F Gesteland, Translation: The twenty-first amino acid. *Nature* 407:463, 9/28/00. News. Discussion of how certain UGA codons get read as selenocysteine, in both bacteria and eukaryotes. The immediate news is discovery of the proteins involved in the process for eukaryotes.

H L True & S L Lindquist, A yeast prion provides a mechanism for genetic variation and phenotypic diversity. *Nature* 407:477, 9/28/00. (+ News, Partridge & Barton, p 457.) Yeast prions are serving an important role as a model system for the study of prions. But what are they there? The first yeast prion discovered was a protein involved in translational termination. Here they argue that the prion-forming capability of this protein is maintained because of its usefulness for adaptation to unusual conditions.

S Karlin & J Mrazek, Predicted highly expressed genes of diverse prokaryotic genomes. *J Bacteriol* 182:5238-5250, 9/00. If highly expressed genes tend to use the "preferred codons", then finding genes with a high percentage of preferred codons would suggest that they are highly expressed. They analyze available genome sequences, using this logic to find candidate highly expressed genes. They use the codon usage in the genes for ribosomal proteins, known to be highly expressed in general, as the basis for knowing what codons are preferred.

A W Kloser, Intragenic suppressors of an OmpF assembly mutant and assessment of the roles of various OmpF residues in assembly through informational suppressors. *J Bacteriol* 183:264-269, 1/01. They isolate suppressors by selecting for revertants of a temperature-sensitive mutant. Exploration of the pseudo-revertant suppressors provides information about the protein structure and assembly.

A Guigueno et al, Oversynthesis of a new *Escherichia coli* small RNA suppresses export toxicity of DsbA'-PhoA unfoldable periplasmic proteins. *J Bacteriol* 183:1147-1158, 2/01. They have a toxic protein in *E. coli*. The basis of the toxicity is not clear, but may involve clogging of protein export machinery by an unfoldable protein. If it is degraded by the DegP protease, toxicity is not seen. We might say, then, that DegP suppresses the toxicity of their protein. But what they do here is to use a DegP mutant, and select for strains that grow when the toxic protein is produced. By definition, they are looking for suppressors. Of particular interest, they find that a small RNA molecule, which is apparently not translated, can suppress



the toxicity of the original protein. At this point, they do not understand the reason for the suppression.

A Böck, Molecular biology: Invading the genetic code. *Science* 292:453, 4/20/01. News, to accompany two articles in this issue. The two articles show ways to get bacteria to incorporate non-standard amino acids. In one case, they modified a tyrosyl tRNA to read the amber codon, and modified the corresponding activating enzyme to use O-methyltyrosine. In the other case, they developed a mutant of the valine activating enzyme with relaxed proofreading of the amino acid, so that it now incorporated substantial amounts of aminobutyric acid.

C M T Spahn et al, Localization of the ribosomal protection protein Tet(O) on the ribosome and the mechanism of tetracycline resistance. *Molecular Cell*, 7(5):1037-1045, 5/01. Tetracycline inhibits protein synthesis by inhibiting binding of the aminoacyl tRNA. Both the Tet(M) and Tet(O) proteins promote tetracycline resistance by causing the tetracycline to be released from the ribosome -- in a GTP-dependent reaction. They now show that Tet(O) binds to the ribosome at a site that overlaps that for EF-G -- as had been shown previously for Tet(M).

T Wiegert & W Schumann, SsrA-mediated tagging in *Bacillus subtilis*. *J Bacteriol* 183:3885-3889, 7/01. tmRNA (also known as SsrA RNA) is an unusual RNA; the 'tm' refers to transfer and messenger. This RNA responds as a tRNA-like molecule if the ribosome reaches the end of a broken message, but translation then shifts to reading the tmRNA as a mRNA. As an mRNA it codes for a special tag sequence that marks the protein for degradation -- appropriate, since that protein was made from a broken message. It may be that the key role of the tmRNA is to free stalled ribosomes. Oddly, the tmRNA is found in all bacteria and maybe mitochondria, but is required only for some bacteria and some phages. Here, they explore how proteins tagged from tmRNA are degraded. See hw.

I S Gabashvili et al, The polypeptide tunnel system in the ribosome and its gating in erythromycin resistance mutants of L4 and L22. *Molecular Cell* 8(1):181-188, 7/01. They use sophisticated electron microscopy to explore the ribosomal structure, including from some mutants resistant to erythromycin ("Ery"; Fig 18.11). Interesting findings include...

- A network of channels (tunnels), not just a single exit channel. The role of these extra channels is unclear. It is possible that they represent alternatives for nascent protein export, but it is also possible that they are simply for water or ion flow.
- One Ery<sup>R</sup> mutant has a channel opening too small to allow the drug to enter.
- Another Ery<sup>R</sup> mutant seems to have a channel big enough that the presence of Ery does not block the protein.

F J Iborra et al, Coupled transcription and translation within nuclei of mammalian cells. *Science* 293:1139, 8/10/01. (+ News, Hentze, p 1058.) Translation in the nucleus? There have been hints of this for some time. Most of the required players show up in the nucleus, at some level. Perhaps most intriguingly, the finding that premature nonsense codons are detected in the nucleus, marking the mRNA for destruction, seemed inexplicable with the conventional view that translation is restricted to the cytoplasm. Here they show translation in isolated nuclei. Technical questions remain, but this is an intriguing paper.

P. Computer resources

For data on preferred codons in a variety of organisms, see the Codon Usage Database:  
<http://www.kazusa.or.jp/codon/>

A movie on translocation, showing the “ratchet”. This movie accompanies the article by Frank & Agrawal (2000). To get it, go to the Nature web site, find the journal issue, then the article; choose Supplementary Information, and you can download the movie file. It is about 2 Mb. You will need to register for the Nature web site, but it is free. You will need Quick Time 4 to view the movie. I will show the movie in class. The Nature web site:  
<http://www.nature.com>

A Stanford chemist, Vijay Pande, is soliciting help in testing models of protein folding. He wants to use the extra capacity of your home computer. I think you may find the site interesting, regardless of whether you want to participate. See

<http://foldingathome.stanford.edu>

(If you want to read about this, there are brief items in Nature 407:667, 10/12/00, and Science 290:1903, 12/8/00. The latter discusses some of the computing issues involved and how these differ among various of these distributed processes.)

Q. Errata

Ch 3 (but relevant to current Ch)

p 53, left col four lines before the summary. ‘deacetylated’ should be ‘deacylated’.

R. Homework

The homework is divided into sets, for convenience.

Set 1. Genetic code

1. Consider the peptide sequence Met Phe Leu Pro Lys.

How many possible nucleotide sequences are there that could code for this peptide? (You may use the abbreviations R, Y and N to mean any puRine, any pYrimidine or any base, respectively.) (Suggest: Write the peptide sequence on a line. Then write all possible codons for each amino acid below that aa.)

2. Consider the DNA segment 5'.....CACGTGCCTAACGGA.....

a. How many possible protein sequences could come from this region of the genome? Why, in terms of the DNA? (Assume that you have no other clues about the region except this partial sequence. And for now, don't worry about what any specific codons mean.)

b. If you had a more extended DNA sequence (and a computer), what would you look for to try to decide which of the possible proteins is the correct one?

3. a. What two amino acid sequences could result from the following bacterial mRNA sequence, assuming that the reading frame is as shown: 5'-... GUG ACG AGG ?

b. How might you tell which of those two possible sequences would actually result, if you had a more extended DNA sequence?

c. Why are there only two possible translation products in this question (rather than six as in the previous question)?

4. Can a single base change result in the charge on a protein changing by 2 charge units? If so, give a specific example.

5. There are a few individuals who seem to be resistant to infection by HIV-1 (the AIDS virus). Some of them are homozygous for a defective CKR5 gene. The common mutation found is a deletion of 32 bases in the coding region.

a. What effect should this mutation have on the primary sequence of the protein? Be as specific as possible.

b. What phenotype would you expect for individuals homozygous for this mutation (besides HIV resistance)? Explain. (The wild type gene codes for a chemokine receptor.)

6. Consider a gene for a membrane protein. The first part of the gene contains several (10-20) consecutive repeats of the pentanucleotide sequence CTCTT. This part of the gene codes for the signal peptide, which is involved in targeting the protein to the membrane. After the targeting occurs, the signal peptide is removed and discarded.

For reasons that are not entirely clear, mutations occur at high frequency that result in a change in the number of repeats of this sequence. (Assume that the length and amino acid sequence of the signal peptide are not of concern here.)

a. What will be the effect of such mutations on the expression of the membrane protein? Consider mutations that change the number of repeats by 1, 2 or 3.

b. What amino acid sequence is coded for by the repeat of the given pentanucleotide? (The sequence given is the sense strand.)

c. Mutations of the type described may be initiated by DNA slippage, in which one strand of one repeat pairs with a different copy of the repeat on the other strand. Once such a wrong-repeat pairing occurs, it is easy to imagine that replication and/or repair processes could “fix” the mutation (i.e., make it permanent). For the original slippage event to occur, it might help if the repeats in question were transiently single stranded. Which form of DNA might promote this slippage event? Does the DNA sequence of the repeat in this case support this suggestion? (Review Ch 2 handout for the forms of DNA; not in book.)

7. What feature(s) of the genetic code, or its usage, facilitate the transfer of genes (by genetic engineering -- or naturally) from one organism to another? What feature(s) may complicate such transfer?

### Set 2. Other

8. We have suggested that a rare tRNA can limit the rate of protein chain elongation (Sect C). We have also suggested that this can result in reducing the amount of protein made. However, there is an argument that this should not be so. The amount of protein made is determined by the number of initiation events, and should not be affected by how long it takes to make a protein once it has been initiated. Why, then -- considering current/recent chapters -- is there sometimes an effect of rare tRNAs on the amount of protein made?

9. In the Ch 17 hw, #2, we discussed why two sites are used to determine initiation sites. We raised the question of the relevance to termination. So, now is the time... Compare and contrast the way sites are used to define initiation and termination.

10. You find that a specific fragment of rRNA interferes with termination at UGA codons, and seems to have little or no other effect. Assume for now that the effect is due to the fragment mimicking a region of the rRNA in the ribosome, which normally binds a specific protein.

a. Suggest a protein that is likely to be involved in this binding. Choose from among the various "factors" that Weaver introduces. What is the basis of your choice, and what are your reservations in making this choice?

b. Interestingly, not only does the particular rRNA fragment promote readthrough at UGA codons, but so does the anti-sense RNA from the same region. That is, both the rRNA fragment and its complement have essentially the same effect. How can this be? (Discussing this does not depend on knowing which protein is involved, so this part is substantially independent of part a.)

11. a. Why, in certain cases, does an amber (UAG) suppressor fail to restore function of a gene carrying a UAG mutation? (There are two fairly distinct possible reasons. Assume that the mutation is within the normal coding region.)

b. Is part a a question about genotype or phenotype? Explain.

12. Try to do this question by referring only to the genetic code and the wobble rules.

a. What is the anticodon in tRNA<sup>Trp</sup>?

b. This tRNA can mutate so that it (also) suppresses the nonsense codon UGA. The tRNA<sup>Trp</sup> UGA suppressor that has an anticodon mutation can read both the original trp codon and UGA. What is the anticodon in this suppressor tRNA?

c. The tRNA<sup>Trp</sup> can also mutate to become an amber (UAG) suppressor, by an anticodon change. What is the anticodon in this suppressor tRNA?

d. Ordinary *E. coli* strains carrying the tRNA mutation discussed in part c are not viable. Suggest two possible reasons for this result. If you made a diploid, containing one wild type allele and one allele carrying this mutation, how would this distinguish the two reasons?

13. The stop codons are three of the four possible codons of the form URR (R = purine). Analysis of the structure of a region of bacterial RF1 and RF2 shows a critical tripeptide in the region that determines stop codon specificity. In this region, RF1 has ProAlaThr and RF2 has SerProPhe. The data suggest that the Pro of RF1 precludes G from being read in the 2nd position of the codon, and the Phe of RF2 precludes G from being read in the 3rd position of the codon. (Weaver discusses this on p 600, and Fig 18.38.)

a. Make a little chart to show how this explains the known specificity of these RFs.

b. What specificity of stop codon recognition would you expect from artificial RFs that contain ProProPhe or SerProThr in this position?

14. The top part of the Fig (at right) shows a model of how EF-Tu binds GTP. The lower part shows the modified base xanthine (X). Design a modified EF-Tu that will bind XTP instead of GTP.

15. Consider a “broken” mRNA, which lacks a termination codon for protein synthesis. What will happen when a ribosome reaches the end of the message?

Try to propose a couple of reasonable possibilities, and note their advantages and disadvantages.

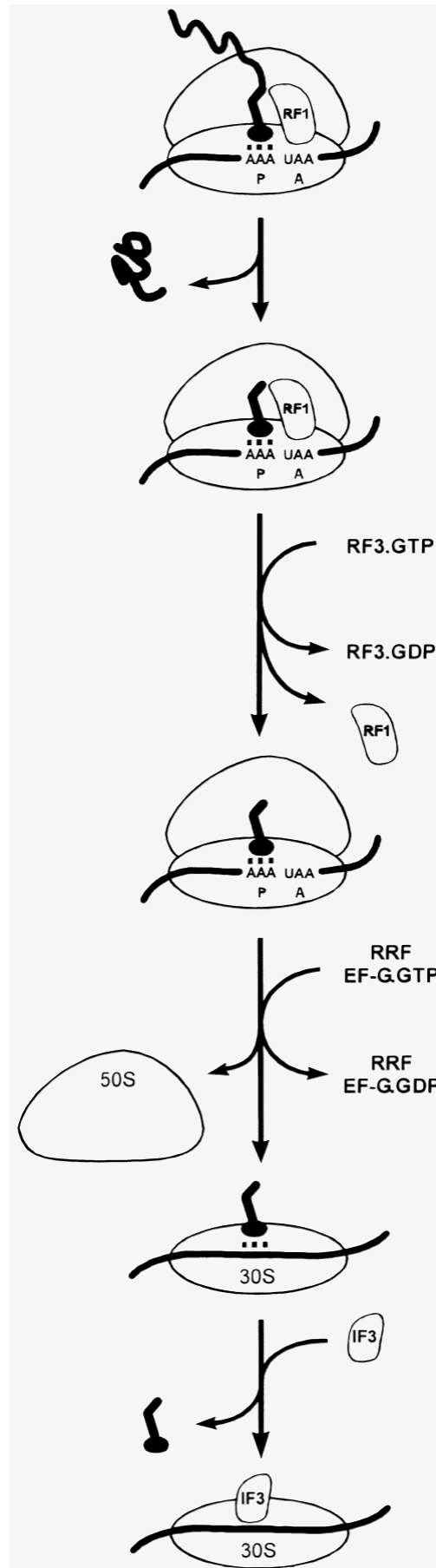
Hint... AANDENYALAA. ☺

Fig 6 from Karimi et al (1999), re Sect L.

Figure legend from the paper:

Figure 6. Model for the Events following Peptide Release.

Release factor RF3 catalyzes the dissociation of RF1 (or RF2) from the A site after peptide release (Freistroffer et al., 1997). RRF and EF-G bind to the posttermination complex and in a GTP-requiring reaction provoke the dissociation of the 50S subunit. Initiation factor IF3 displaces the deacylated tRNA from the 30S:mRNA complex.



S. Partial answers

1.

Met	Phe	Leu	Pro	Lys
AUG	UUY	UUR	CCN	AAR
		CUN		
1	2	6	4	2

The last line shows the number of possible codons for each position. Multiplying all these together gives the number of possible coding sequences for the peptide: 96.

2. a. There are six possible protein sequences from the region. (If you didn't get six, think about it some more before reading on.)

There are 3 reading frames in the direction shown. These would start with the codons CAC, ACG, CGT. But the region might be translated from the other strand, in the opposite direction. The complementary DNA strand is 5'.....TCCGT...., and so would code reading frames starting with TCC, CCG, CGT.

b. Sequences that code for proteins have some nonrandom features. They include a long stretch of sense codons, from an initiation codon to a stop codon. (Remember that about 1 codon in 20 is a stop codon.) They also should be appropriately associated with transcriptional start and stop signals.

[This question explores a practical problem in molecular biology. The logic should be clear, but it is not always easy in practice. Proteins vary in length, some have odd coding features -- and sometimes published DNA sequences contain errors. Thus, the "rules" proposed here are generalities.]

3. a. GUG ACG AGG = Val/Met Thr Arg. That is, there are two possible ways to translate the first of these codons.

b. GUG is normally a valine codon. The Met would result if this GUG is used as an initiator codon. Look for ribosome binding sites, other initiator and terminator codons, i.e., look for an "ORF" -- an Open Reading Frame.

c. There is more information here, thus restricting the possibilities. In the previous question we did not know which strand was translated, or which reading frame was used. In this case, we know both of those, because the mRNA sequence is given and the reading frame is given. Normally, this case would yield only one translation product. There are two here, because of the ambiguity about initiation.

4. A charge change of 2 would be from a basic amino acid to an acidic amino acid (or vice versa). The codon sets for Lys and Glu are in the same column of the genetic code, hence differ only at the first position.

5. a. The deletion is 10  $\frac{2}{3}$  codons, thus results in a frameshift. Not only are 10 aa deleted, but the entire rest of the protein is incorrect. (In fact, the protein is truncated, because of a stop codon generated by the frameshift.)

b. Well, can't predict phenotype without knowing something about what the gene does. If your intuition is that the gene sounds important, fair enough. But the individuals seem "perfectly normal". One possibility is overlap of chemokine function, so that no one hormone is critical.

This question is based on R Liu et al, Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. Cell 86:367, 8/9/96.

6. a. Assume that we start with N repeats, and that the membrane protein is made normally. Since the repeat contains 5 bases, addition (or deletion) of 1 or 2 repeat units will cause a frameshift; the membrane protein sequence will be out of phase, and thus it will not be made. Addition of 3 repeats will leave the membrane protein sequence unchanged. The signal peptide will be 5 amino acids longer in this case. However, by assumption (and real data), this change in length of the signal peptide is of no consequence. (Further, the signal peptide is cleaved off of the membrane protein, so there is no effect on that protein.)

This is based on G L Murphy et al, Phase variation of gonococcal protein II: regulation of gene expression by slipped-strand mispairing of a repetitive DNA sequence. Cell 56:539, 2/24/89. The gonococcus (*Neisseria gonorrhoeae*) contains several similar genes for a membrane protein (actually, an outer membrane protein). Each of the genes contains the repeat sequence, which does change in number quite rapidly. As a result, a single strain rapidly changes which members of the family are being expressed. This is one of the mechanisms that the gonococcus uses to change its antigenic face.

b. Since the repeat is 5 nucleotides, we must consider three consecutive repeats to get a sequence which will give an amino acid sequence that repeats. Write CTCTT (= CUCUU) 3 times; the resulting 15-mer codes for LeuPheSerSerLeu.

c. H-DNA contains SS regions. If the region in question is shifting back and forth between B and H forms, one can imagine such slippages occurring. The given repeat sequence is polypyrimidine (with polypurine on the other strand, of course). This is the type of sequence that favors the H form.

Strand slippages, such as the case discussed here, are being invoked to explain a number of phenomena. The slippages may involve consecutive repeats in the same direction, as in this case, but might also involve nonconsecutive and/or inverted repeats. Murphy et al introduce some of these phenomena.

7. Positive: universality. Negative: occasional non-universality, occasional recoding, codon usage patterns.



8. Think about the density of ribosomes along the mRNA (number of ribosomes per unit length of mRNA). What happens to this density as a result of the rare tRNA? (More or less by definition, a rare tRNA creates a pause in the elongation process.) How might this change in ribosome density (or ribosome “loading”) affect the amount of protein made?

As to the ribosome loading... We start with the simple picture that ribosomes are normally approximately evenly spaced along the mRNA. The density of that spacing is determined by the frequency of initiation (i.e., by the strength of the ribosome binding sites, and whatever other processes affect initiation). A rare tRNA creates a pause in elongation. As a result, ribosomes will back up “behind” (upstream, toward the beginning of the mRNA) the paused ribosome. On the other hand, ribosomes which have made it by the pause position can proceed freely. Thus the density of ribosomes behind the pause will be higher than normal, and beyond the pause it will be lower than normal.

So, how can this change in ribosome density affect the amount of protein made? There are at least two distinct possibilities. For class discussion.

10. a. RF2 is the only protein specifically involved in termination at UGA (p 598). But if the rRNA fragment works via RF2, why is there no effect on UAA termination, since RF2 acts at both UGA and UAA?? No real answer to this is available. However, UGA codons are the least effective termination codons -- easiest to read through. Mutations in rRNA that specifically affect termination at UGA are also known.

b. The simple model might be that the rRNA fragment interferes with termination by binding to the RF2, thus preventing it from being available for termination. The complement, then, would bind to the rRNA itself, thus preventing RF2 from binding to the ribosome.

This question is based on Arkov et al (1998).

11. a. 1) The amino acid may be unacceptable; that is, the protein containing the amino acid inserted by the suppressor may not be active. 2) The level of suppression may be inadequate, and not provide sufficient active product.

In the example of Fig 18.35, the question would be whether the substitution of Tyr for the original Gln is acceptable, and whether the level of suppression makes sufficient protein for normal function.

(Is it relevant here that the suppressor might cause harm by making some other protein inactive? That might depend on how you interpret the question. In the narrow sense, the question is about the effect on a particular gene. Of course, if you think about the question at the organismal level, an apparent failure to suppress could be due to a side effect. This possibility will come up again in a later question.)

b. Phenotype. The key word (issue) is “function.” The genotype is “carries suppressor” and “has suppressible mutation”, but we don’t know whether the phenotype will be “it works”.

12. a. The one Trp codon is 5'-UGG-3'. Thus the anticodon is 5'-CCA-3'. Note the direction. Convention is to write nucleotide sequences 5'→3'. If you say the anticodon is CCA, that's ok, because the implied direction is correct. However, you cannot say that the anticodon is ACC, unless you explicitly note the direction: 3'-ACC-5'. Confusing? That's why it's always safe to show the chain direction explicitly. (Recall Ch 2 handout Sect B.)

b. We now have a single tRNA that recognizes both UGG and UGA. By the wobble rules, U in the anticodon can read both G and A in the 3rd position of the codon -- the first position of the anticodon. Thus the suppressor tRNA anticodon must be UCA.

(Since we looked at the wobble table to choose the third base in part b, shouldn't we do the same for part a? Yes. And the table will tell us that if the only codon base recognized is G, the proper anticodon base is C.)

(The Trp tRNA can also mutate to become a UGA suppressor with a non-anticodon mutation.)

c. CUA. This reads UAG -- only; it no longer reads the Trp codon.

d. Possible reasons for nonviability of strains carrying this suppressor: if there is only one copy of the "parental" tRNA gene, the mutant (suppressing) strain has no tRNA to translate the normal Trp codon; the Trp suppression of some UAG codon(s) causes lethality.

To distinguish them: make a diploid strain, carrying both the parental and mutant tRNA genes. When this is done, the strain is fine. This supports the first suggestion.

13. b. ProProPhe excludes G in both 2nd and 3rd positions, so should recognize only UAA. SerProThr does not have either exclusionary amino acid, so should recognize all codons of the form URR -- including UGG, which is not normally a stop codon. Experimental work has borne out these predictions.

14. EF-Tu D138N. That shorthand means EF-Tu with the D (aspartic acid) at position 138 replaced by an N (asparagine). The result... XTP bonding to the new protein is very much the same as GTP bonding to the original, but the positions of one -C=O and -NH<sub>2</sub> H-bonding pair have been switched.

15. For class discussion of the possibilities.

Wiegert & Schumann (2001) discuss Nature's answer. We will discuss trans-translation, tmRNA, and AANDENYALAA.

Cruz-Vera et al (2000) discuss another protein whose main job may be as a scavenger to clean up from such problems.