

UN2005/UN2401 '18 -- Lecture # 14 -- RNA & Protein Synthesis

(c) 2018 Mowshowitz Department of Biological Sciences Columbia University New York, NY. Last revised 10/24/18

Handouts: You will need 14A -- code table & tRNA structure
 14B -- Protein Synthesis
 13B -- DNA synthesis vs RNA synthesis

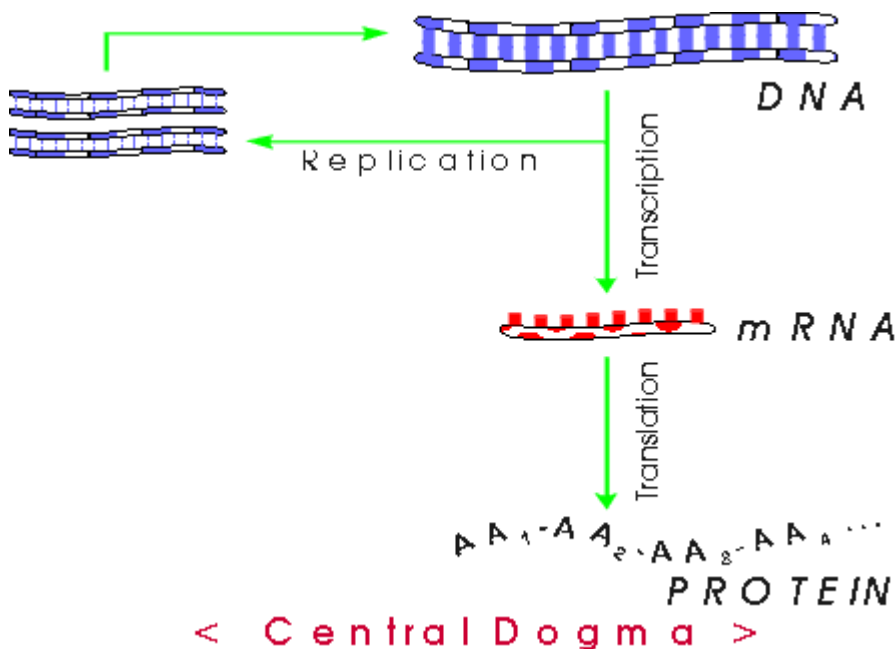
Paper copies of the handouts will be provided in class; PDFs are linked to the Handouts page. Note that Dr. Mary Ann Price has made videos of all class demonstrations; see the videos link on the Course Menu. (We also have 2 older videos of protein synthesis; links are below.)

1 . Central Dogma -- How does DNA do job # 1?

A. Big Picture. So we have a big DNA that includes a particular gene = stretch of DNA coding for a single peptide; how will we make the corresponding peptide?

Note: gene usually means a stretch of DNA encoding 1 polypeptide, but there are complications as we'll see later.

1. *Basic idea* -- See picture below; or see Becker fig. 18-1 (21-1) or Sadava fig. 14.2:



2. Terminology:

a. Replication = DNA synthesis using a DNA template.

b. Transcription = RNA synthesis using a DNA template.

c. Translation = Protein Synthesis. This term has two possible meanings (we will stick to the first):

(1) Usual meaning = protein synthesis using an RNA template (RNA → protein).
Used in contrast to transcription (DNA → RNA).

(2) In some contexts, translation can mean the entire process (DNA → RNA → protein).

II. Why RNA, especially mRNA?

A. Basic idea: mRNA = Working, disposable copy vs DNA = archival, permanent master copy. DNA = big fat comprehensive reference book or complex web site. mRNA = Xerox of one (book) page or print out of one web page with information you need for a particular assignment. Book stays safe in library; web site remains unchanged. Xerox goes to your room, is actually used, gets covered with coffee stains, smudged, and thrown away.

B. Why mRNA?

1. *Convenience.* Small size (1 or a few peptides' worth) is much more convenient than many genes' worth. Xerox of one page much more convenient to work with than big fat book.
2. *Preserve Master.* Using mRNA to make protein saves wear and tear on master -- no coffee stains on the archival copy (DNA).
3. *Flexibility.* Different amounts of mRNA can be made when cell needs to make different amounts of different proteins. More on this when we get to regulation (operons).

C. Summary: How does RNA make protein?

1. "RNA makes protein" means two things:

a. Need mRNA (info goes DNA → RNA → protein)

b. Need several kinds of RNA to make protein -- See Sadava Fig. 14.2

- mRNA to act as template -- determines order of amino acids
- tRNA to carry the amino acids to the template, and line them up
- rRNA (in ribosomes) to align the tRNA's carrying the amino acids and hook the amino acids together
- Of course you need additional proteins (enzymes and other factors) to make protein

2. *Hardware vs. Software.* rRNA and tRNA are the hardware or tools or machines; mRNA is the software or working instructions or MP3/ tapes/CDs/punchcards. Cells use same old hardware and constantly changing, up to the minute, supply of new software.

D. How does RNA compare to DNA in structure, chain growth & function? Overview

1. *Structure:* See Sadava table 4.1 & fig. 4.3 and or Becker table 3-4 & fig. 3-17 for comparison of DNA and RNA.

- RNA is single stranded (although sections may double back on themselves → double stranded regions)

- RNA has U not T, ribose not deoxyribose and is generally shorter, but otherwise like DNA.
- RNA molecules are usually shorter than DNA molecules.
- RNA molecules don't last as long (in cells) as DNA.

These differences are summarized in a table in lecture #13.

2. *Some Important Consequences of Structural Differences*

- RNA is less stable than DNA -- more easily damaged (because of reactive OH on ribose and because a single strand is more exposed)
- RNA is less easily repaired (because there is no 2nd strand to use to correct mistakes on first strand).
- DNA is also more easily repaired because it has T not U, so damaged C's (which are oxidized to U) can be recognized and removed.

3. *Types.* There are 3 major types of RNA involved in translation: messenger RNA (mRNA), transfer RNA (tRNA) and ribosomal RNA (rRNA). The roles of the different types of RNA are outlined above and will be explained in detail below.

4. *Synthesis – Chain Growth.*

a. RNA grows just like DNA by adding nucleoside triphosphates (XTP's) to the 3' end of a growing chain. For RNA, enzyme for elongation is called RNA polymerase, XTP's are ribo (not deoxy) and U replaces T.

b. Synthesis of RNA requires a DNA template.

More details on RNA synthesis below.

Reminder: Any kind of RNA -- tRNA, mRNA, rRNA, or any other type of RNA -- is made in the same way from a DNA template. The product of transcription does not have to be an mRNA -- it can be a tRNA, rRNA, etc. RNA is NOT used as template to make more RNA.

III. DNA synthesis vs RNA synthesis. How is RNA made? The easiest way to go over RNA synthesis, given that we've discussed DNA synthesis at length, is to compare DNA and RNA synthesis. **This is a run through of handout 13-B.**

A. What is the same?

1. *Energy input from ATP:* Use nucleoside triphosphates (XTP's, not XMP's) & split off PP_i; use pyrophosphatase. (Use XTP's with ribose not deoxyribose, but mechanism same as for DNA. See below.)

2. *Direction of Chain Growth:* Chain grows 5' to 3' by addition to 3' end.

3. *Template & Base Pairing:* Need anti-parallel DNA template, put in complementary bases -- A (in template) pairs with U not T in growing chain, but otherwise same [{Q&A}](#)

4. *Template is always DNA:* **All RNA chains are made from a DNA template -- that includes tRNA's and rRNA's, as well as mRNA's.** tRNA and rRNA molecules are **not** made from an "mRNA" template.

5. *Binding Site for Polymerase is Double Stranded.* The polymerase (DNA or RNA polymerase) binds to a specific double stranded sequence in the DNA. (The sequence is different for the two enzymes.)

See problems 7-1 & 7-2.

B. What's Different?

1. Enzymes Used

- Growth of DNA chain is catalyzed by DNA polymerase (and associated enzymes)
- Growth of RNA chain is catalyzed by RNA polymerase.
- DNA polymerase can proof read but cannot start chains; RNA polymerase does not proof read but can start chains. See the section on Proof Reading below.

2. *Choice of Substrate.* If you put all 8 XTP's in a test tube, what do you get, DNA or RNA? Enzyme (DNA vs RNA pol) is responsible for which nucleotides used.

- RNA pol. uses ribonucleoside triphosphates (containing U, not T).
- DNA pol uses deoxyribonucleoside triphosphates (containing T, not U).

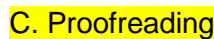
3. Products

- DNA is long and double stranded
- RNA is short and single stranded

4. Choice of which part of Template to use

- Template = short section, one strand at a time (for RNA synth.) vs all of both strands (for DNA synth.)
- Why? Because starts and stops are different. Starts & stops = sequences in DNA recognized by the enzymes = places where replication or transcription starts (or ends). These must be different for the two enzymes.

- See problem 7-6**



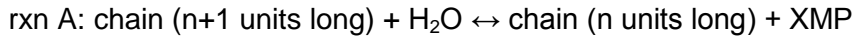
3. **What good is proofreading?** This allows the enzyme to 'back up' and remove nucleotides that were added in error by hydrolyzing the phosphodiester bond it has just made (if the wrong base was put in). The 3' to 5' exonuclease is important in correcting mistakes and maintaining a high accuracy during DNA synthesis.

4. The proofreading reaction (3' to 5' exo) is not the same as the reverse of the polymerization reaction, so a different catalytic center is needed for each one.

The polymerization reaction is :



The proof reading reaction is



Rxn B is the normal elongation (synthesis) reaction catalyzed by DNA polymerase = reaction 4 on handout 12-2.

Reaction B is not reversible because the PP_i is immediately hydrolyzed by pyrophosphatase. Note that reaction A is NOT the reverse of reaction B. You took a nucleotide off the 3' end, but you did NOT get the XTP back; you got XMP instead.

5. Proof reading and the ability to start chains are linked

DNA polymerase proofreads and cannot start new chains; RNA polymerase does not proofread and can start new chains. These properties are linked, because the binding site of DNA polymerase that allows proofreading prevents it from starting new chains. For DNA polymerase, there must be a previous phosphodiester bond to break -- if the polymerase needs to back up. For RNA polymerase, no previous bond is required, as the enzyme never backs up.

Detailed explanation (FYI): DNA polymerase checks twice that the base added is correct. After adding each base, the base to template match is "rechecked" before the chain grows any longer. If the last base added turns out to have been the wrong one (perhaps it was in the wrong tautomeric form temporarily and mispaired with the template?), then the enzyme backs up and removes the last base before trying to add another. Once the enzyme checks that the last base added is ok, it checks the match between the base to be added and the template. If there is a match, the enzyme catalyzes formation of the phosphodiester bond. So each base - template match is checked twice -- once when the base is about to be added to the growing chain and once before the next base is added to it.

RNA pol. only checks the pairing between the base to be added and its complement in the template. So if the last base put in was wrong, so be it. No backing up or corrections.

An example of proof reading (which you should be able to do) is in problem 6-14, part B-4.

D. Template Details

1. One Strand is Template for RNA polymerase. For any one gene or region, RNA polymerase uses Crick or Watson, but not both, as template. RNA that is made is complementary (and antiparallel) to the template strand. Note that an entire strand is not used as template throughout. The "Watson" strand of DNA is used as template in some sections and the "Crick" strand in others.

2. Continuous vs. discontinuous synthesis.

- DNA synthesis: Replication fork moves down DNA making complements to **both** strands; one new strand is made continuously and one discontinuously. Ligase is needed for synthesis of lagging strand.
- RNA synthesis: RNA polymerase moves down DNA making complement to one strand **or** the other (in any particular region). Therefore RNA synthesis is always continuous. Do you need ligase?

3. Terminology

a. Transcribed Strand of DNA. Strand used as template is called the transcribed or template strand or the antisense strand (in that region). This strand is **complementary to** the RNA that is made.

b. Sense Strand of DNA. Strand that is **not** transcribed (in that region) is called the sense strand or coding strand. The base sequence of this strand is **identical to** the RNA that is made (except that the RNA has U and the sense strand of DNA has T).

c. An entire DNA strand (going the length of a whole molecule) is not all "sense" or "antisense." "Watson" may be sense in one section and "Crick" may be sense in the other (as in the picture on handout 13-B). The terms "sense" and "transcribed" strand are defined for each section of the DNA that is transcribed as a unit (usually a gene or small number of genes).

d. Sense RNA. The usual RNA transcribed from the DNA is said to be "sense." (Sense RNA matches the sense strand of the DNA.) The complementary RNA, if it exists, is said to be "antisense." Some practical uses of "antisense RNA" are below.

e. Why this terminology? The sense strand (not the template) actually contains the information used to line up amino acids to make proteins. (Assuming the gene codes for a peptide.) When a DNA sequence is published, it is usually the sense strand that is given. Why? If the gene codes for a protein, the amino acid sequence of the protein is much easier to figure out using the sense strand -- you just consult the code table (details below or next time).

f. Additional notes FYI on terminology:

(1). Sense Strand. Becker (and some others) call the sense strand the coding strand, meaning the "strand coding for protein." I prefer the term "sense strand" since coding strand could mean "coding for protein" or "coding for mRNA." (The term "coding strand" is almost always used the way Becker uses it, to mean "coding for protein.")

(2). Template Strand. The terms "template strand" or "transcribed strand" can also be interpreted in more than one way, but these terms are virtually always used to mean the strand acting as template for RNA synthesis (= the strand that is transcribed from, not the strand that is being made, during transcription). The template or transcribed strand is **not** the strand equivalent to the mRNA -- the template strand is the strand **complementary** to the mRNA.

4. *Directions:* Suppose you have a double stranded DNA template. If need to copy "Crick," RNA polymerase will go one way (say right to left -- actual direction will depend on which end of template is 5' end); if need to copy "Watson" RNA polymerase will need to go the other way (say left to right). What determines where RNA polymerases starts & which way it goes? This is discussed below.

See problems 7-3, 7-4, 7-8 & 7-9.

E. Details for Starts and Stops (see picture above or bottom of handout 13-B)

1. *Start sequences as binding sites.* A start signal for transcription or replication is a sequence in the DNA recognized by the appropriate polymerase = binding site for that polymerase (or complex).

2. *Names of start sequences*

a. **Starts for DNA synthesis** = Origins. DNA pol. recognizes (binds to) start signals for replication called origins (ori's).

b. **Starts for RNA synthesis** = Promoters. RNA pol. recognizes (binds to) start signals for transcription called promoters (P's).

3. *Promoter Details:*

a. **Promoters determine the direction of transcription.** Promoter and enzyme are asymmetric; therefore once enzyme binds, the catalytic end of RNA pol. is "facing" in one direction, and that determines the direction of transcription (and therefore which strand will be template).

b. **The promoter will be a double stranded sequence** at the end of the gene where RNA polymerase starts (= on 3' end of template strand = on 5' end of sense strand). Going along the sense strand, the way the gene is usually written (5' to 3', left to right) the promoter is "upstream" of the gene (that is, upstream of the protein coding part).

4. *How many starts?* There are more P's than ori's in prokaryotic DNA. (Only need one ori per prok. DNA; need one P per mRNA made.)

5. *Stop (Termination) Signals.* Special sequences in DNA may not be needed for DNA pol. -- enzyme may just go until it reaches the end. You do need some sort of mechanism to end synthesis of each RNA. In prokaryotes there are special sequences (often called terminators) that cause the end of transcription. The mech. for ending transcription is somewhat different in eukaryotes and prokaryotes. (We'll do euk. details next term.)

Important Notes:

(1) **Stop signals.** Stops signals for translation (stop codons) are different than the stop signals for transcription (terminators). See Sadava table (not fig.) 14.3. Translational stops are not recognized by the transcription (or replication) machinery. Each set of enzymes (for translation, transcription, or replication) recognizes only its own respective start and stop sequences. (More on this when we get to operons.)

(2) Starting and Stopping. The process of starting (initiation) and stopping (termination) macromolecular synthesis is always more complex than chain growth (elongation). We will skip many of the details of starting and stopping, which are covered in advanced courses. See texts for details.

IV. Sense & Antisense

A. Why use only one strand in any one region (to make RNA)?

1. The function argument: Messenger RNA must be single stranded to fit in a ribosome and be translated. If RNA complementary to mRNA were present, what would happen? The "sense" mRNA and the "anti-sense" complementary RNA would hybridize. The resulting double stranded **RNA** wouldn't be translated. So even though the gene was present, and transcribed, its protein product wouldn't be made. This is what would happen if both strands were transcribed to make mRNA. (Similarly, complementary RNA could bind to, and inactivate, tRNA and rRNA, which are basically single stranded.)

2. The evolutionary argument: If both strands are used to make RNA, you can't optimize one without messing up the other, and vice versa. If natural selection favors the sequence of one strand so that it has optimal function or coding activity, that automatically determines the sequence of the other strand. Natural selection can't simultaneously select for the optimal sequences of both strands (if each strand has an independent function).

B. Is there any use for anti-sense RNA?

1. What good is anti-sense RNA? Gene therapy (adding DNA) should allow you to replace a defective gene that is making an ineffective product. But what do you do about a gene that is making too much product, or making it when it shouldn't? In other words, how do you silence an over-active gene? This is an important question, because inappropriate or over expression of genes is thought to be a major factor in disease, for example, in allowing cancer cells to multiply when they shouldn't. Use of anti-sense technology should allow you to silence an over-active, or inappropriately active, gene.

2. How to get anti-sense RNA into cells? There are 3 ways to do it:

a. Antisense mRNA can be added to cells. Since RNA is easily degraded, modified RNA's, more resistant to hydrolysis, are used instead of ordinary RNA's.

b. Antisense mRNA can be made in the cell from a second copy of the gene. The second copy is added by genetic engineering methods; it is inverted (relative to the promoter), so that the second copy of the gene is transcribed in the opposite orientation from the original copy. Inverting a gene relative to its promoter is equivalent to moving the promoter to the opposite end of the gene (and turning it around) thereby reversing the direction of transcription. The original copy is transcribed from the usual template ("transcribed") strand to make mRNA; the second copy is transcribed from the complementary ("sense") strand to make anti-sense RNA. The two RNA's hybridize to each other and neither RNA is translated.

Third method on next page.

c. Double Stranded (ds) RNA can generate short antisense RNAs -- See Becker fig. 20-36 (23-33). Many eukaryotic cells have enzymes for generating short antisense RNAs from ds RNA. The short antisense RNAs interfere with translation of the corresponding sense mRNA. The short antisense RNAs can act as a defense against many viruses or as a mechanism of regulation during development of multicellular organisms. When double stranded RNA is used to generate short antisense RNAs that interfere with translation, the phenomenon is called RNAi (= RNA interference). More on this next term. (If you want the details now, see the notes of last year, lecture 14.)

To check your understanding of antisense, see problem 7-16, part C.

V. Translation (Protein Synthesis) -- roles of mRNA & tRNA.

RNA is not used as template to make more RNA, but mRNA is used as a template to "make protein." How does that work? That's the next question.

What are the big issues in protein synthesis? Same as for all non repeating polymers = Order, energy and enzymes!! We'll focus on order first.

A. How is mRNA read?

1. *It's read in triplets going 5' to 3'.* Reading starts at a fixed point and then mRNA is read one triplet or codon at a time in the 5' to 3' direction.

2. *Code table* See handout 14A or texts for code table. Note that table lists codons = triplets found in the mRNA (NOT complements of codons) and corresponding amino acids. One codon specifies one amino acid. For example, CUA means leucine; UUU means phenylalanine, AUG means methionine. DO NOT memorize the code table. If you need it to solve a problem, you can always consult it.

3. *Punctuation.* Note that some codons signify "stop", not an amino acid. AUG does double duty as both "start" and "methionine." Translation starts at an AUG, and ends when it reaches the first stop codon after the AUG. How the proper AUG is chosen is different for prokaryotes and eukaryotes. (See the texts if you are interested in the details.) More specifics on stops & starts below or next time.

4. *Leaders & Trailers.* The region before the first AUG is not translated. It is called a leader, or 5'UTR (un-translated-region) or 5'UTS (un-translated sequence). Translation generally stops before the end of the mRNA (at a stop codon -- UAG, UAA or UGA). The untranslated region after the stop codon is called a trailer, or 3' UTR or 3' UTS.

5. *Reading Frames.* There is more than one way to read a nucleic acid sequence in non-overlapping groups of three, depending on where you start. The different ways are called different reading frames. If you start with the first, 4th, or 7th.... base you get one reading frame; if you start with the 2nd, 5th, or 8th.... you get the second, and so on. There are 3 possible reading frames.

To be sure you understand how to use the code table, try problem 7-12, parts A & B.

B. Structure/Function of tRNA For a video of the class demonstration see video (windows media file) by Peter Sloane at <http://www.columbia.edu/cu/biology/courses/c2005/lectures/tRNA.wmv> or see videos on Canvas site made by Dr. Mary Ann Price.

1. Adapter Function -- how does cell know AUG is met and CUA is leu? You have the text or handout with the code table, but cell doesn't.

a. Transfer RNA (tRNA) = adaptor. Cell uses tRNA to match the codon in the mRNA (say AUG or CUA) with the corresponding amino acid (met or leu, respectively).

b. Loading Enzymes. Adaptor must carry the correct amino acid. Cell uses loading enzymes to put the correct amino acids on to their respective tRNA's. More details on loading enzymes below or next time.

2. Structure of tRNA (see handout 14A & texts for pictures)

a. Size: About 75 bases long (relatively small). Consists of RNA chain folded back on itself.

b. Many different ones. Actual number of dif. tRNA's is more than 20 (#of dif. amino acids) and less than 64 (# of dif. codons). More exact estimate of # of different tRNA's to follow below or next time.

c. Two headed molecule: tRNA has 2 critical parts

- one part (in middle of chain) is complementary to codon (= anticodon)
- one part (on 3' end) is acceptor end -- picks up the appropriate amino acid with the help of the appropriate enzyme.
- when tRNA is folded in 3D, acceptor end and anticodon are at opposite ends of molecule

d. General features of structure

Secondary Structure: Each tRNA molecule is doubled back on itself to form a cloverleaf with double stranded sections. Sequences of different tRNA's differ, but all are self complementary in certain regions. Every tRNA molecule has same basic "secondary structure" = cloverleaf.

Tertiary Structure: Cloverleaf is folded into an L shaped "tertiary" structure, which has anticodon at one end and acceptor for its amino acid at the other. (See handout, Becker fig. 19-3 (22-3), or Sadava fig. 14.10 (14.11), for secondary and tertiary structures.) The final folded tRNA molecule is about one codon wide. That way two tRNAs can attached to neighboring codons without bumping into each other.

Important reminder: The code table lists the codons, NOT the anticodons. The anticodon in the tRNA is the **complement** of the triplet shown in the table.

See problem 7-18.

3. How is tRNA used to line up amino acids (AA)? 2 AA at a time are held in place by tRNAs (for forming peptide bond) -- see handout 14B. Why 2? Because a ribosome can hold only 2 **loaded** tRNAs at a time that are hydrogen bonded to mRNA. (Details on ribosomes next time.)

4. How are the tRNA & mRNA paired up? -- All nucleic acids pair in an antiparallel fashion. So if mRNA is written in usual way (5' → 3'), then tRNA is lined up in the opposite way, 3' → 5'. (With the amino acid or chain on its left, 3' end.) Anticodon is often written 3' → 5' to make this clear. For ex., if codon is CGG, anticodon is usually written 3' GCC 5' not CCG (or it is written upside down as on handout 14A).

5. How are the tRNA and AA connected? The AA is attached to the 3' end of its respective tRNA by an ester bond between the COOH end of the AA and the 2' or 3' OH on the final ribose (at the 3' end). This leaves the amino of the AA free.

6. Loading of tRNA. How do you get the right AA on the corresponding tRNA in the first place, and/or how do you reload the tRNA once it gives its AA away? Loading requires enzymes and energy -- we'll look at it carefully next time. For now we'll just assume each tRNA is loaded with its respective amino acid.

C. How does the new peptide chain grow? See handout 14B or Sadava fig. 14.14 (14.15) or Becker fig. 19-9 & 19-12 (22-7 & 22-10).

Note: The first figure in Becker is an overview; the second is a more detailed version of the chain elongation steps. (Note that we will be ignoring the roles of GTP, initiation factors (IFs) and elongation factors (EFs)).

For a video of the class demonstration see videos link on Main menu, or video by Peter Sloane at <http://www.columbia.edu/cu/biology/courses/c2005/lectures/translation.wmv> This video may take a while to load, but it does work (at least on a PC).

1. Chain adds to newest AA. When each peptide bond is made, the growing chain is transferred (from the tRNA that previously held it) to the next amino acid (still attached to its tRNA), not the other way around, for logistical reasons. The newest amino acid is not added to the free end of the chain. Instead, the chain is added to the newest amino acid. (The current system allows the translation machinery to slide down the mRNA reading 2 adjacent codons at a time. The other way doesn't.)

Catalyst for formation of peptide bonds is called peptidyl transferase because the growing peptide chain is transferred as described above. This catalyst is part of the ribosome.

2. Peptide chain grows amino → carboxyl. This follows because the amino acids are held down (attached to tRNA) by their COOH ends. So if chain must add to free end of next AA, must add to amino end of next AA.

Note for those who have had organic: From the point of view of mechanism, the electrons go the other way; the electrons of the amino attack the carboxyl.

3. Energy for peptide synthesis. The energy derived from splitting the tRNA~AA (really the tRNA~chain) bond drives peptide bond synthesis. In other words, the AA-tRNA connection is a high energy bond. How it is formed at the expense of ATP will be discussed next time. (Additional energy is required to bind the AA~tRNA and move the ribosome down the mRNA, but we will ignore the energy details of those steps, as well as the proteins needed to promote them.)

4. *Stops.* The peptide chain stops growing when the translation machine comes to a stop codon. There are no tRNA's for the stop codons, so there is no way that the chain can keep growing if a stop codon comes next. See Sadava fig. 14.15 (14.16) or Becker fig. 19-14 (22-11).

To review protein synthesis so far, and the role of tRNA, try problem 7-21.

Next time: Anything above we don't get to, and then more on ribosome structure and function. After we finish translation, then on to

- (1) what happens when macromolecular synthesis makes mistakes, and
- (2) how is protein synthesis regulated in prokaryotes?

(c) Copyright 2018 Deborah Mowshowitz Department of Biological Sciences Columbia University New York, NY