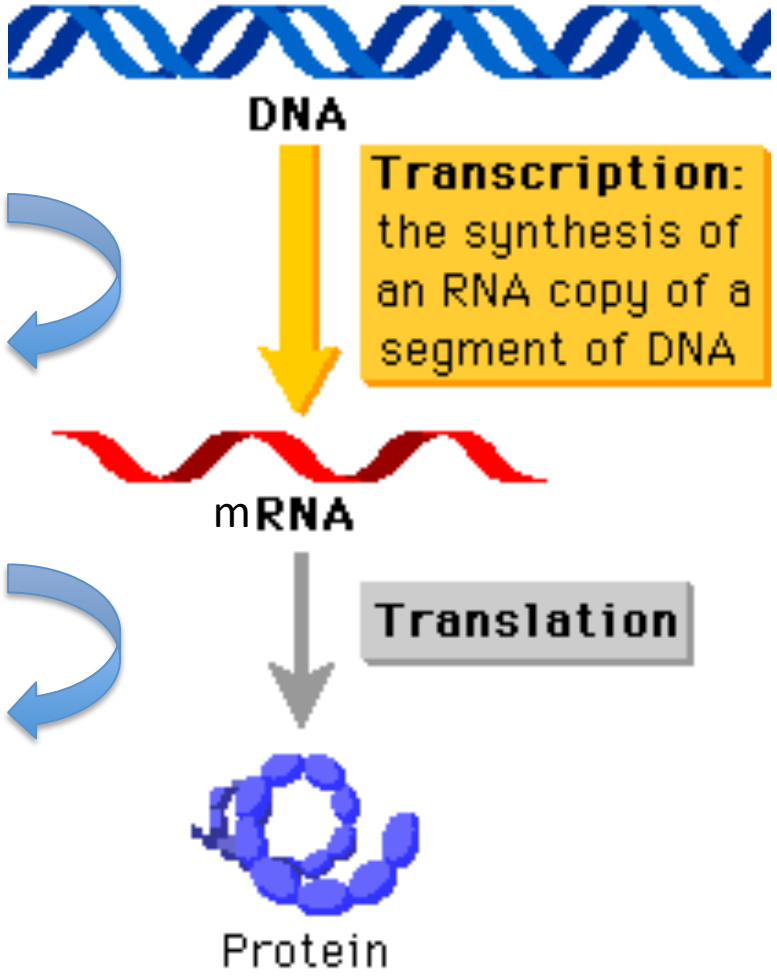
Module 7: The Central Dogma

CSE590: Molecular programming and neural computation. All slides by Eric Klavins.

The Central Dogma



RNA Polymerase

The Ribosome

- Note: We will look mainly at prokaryotic (e.g. e. coli) processes.
- Some of this is the same in eukaryotes, but there are important differences.

Important Molecules

DNA = Deoxyribonucleic acid
A sequence of A, T, C and G (deoxyribonucleotides)

RNA = Ribonucleic Acid (mRNA, tRNA, …)
A sequence of A, U, C and G (ribonucleotides)

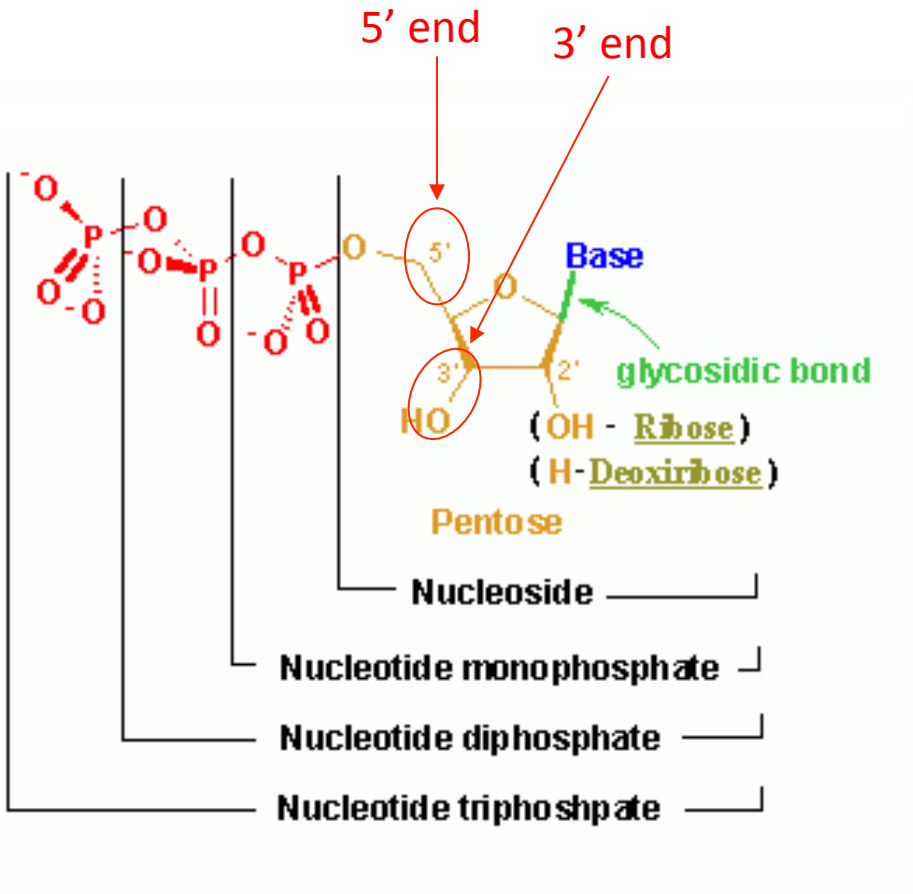
RNA Polymerase (RNAP)
Transcribes (copies) DNA segments into RNA

Amino Acids and Transfer RNA (tRNA)
Help build proteins

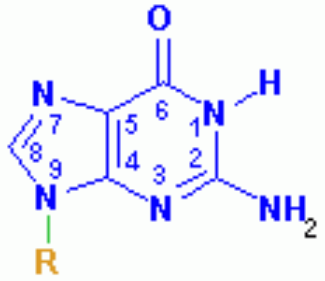
The Ribosome
Translates messenger RNA (mRNA) into protein

Protein
A sequence of amino acids

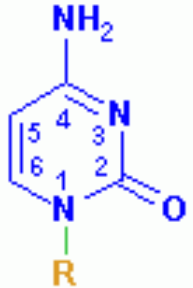
Nucleotides

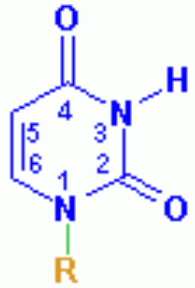


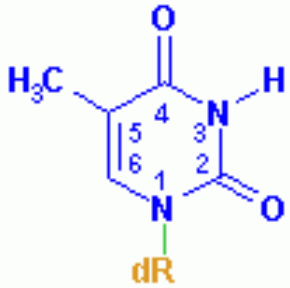
Purines



Pyrimidines



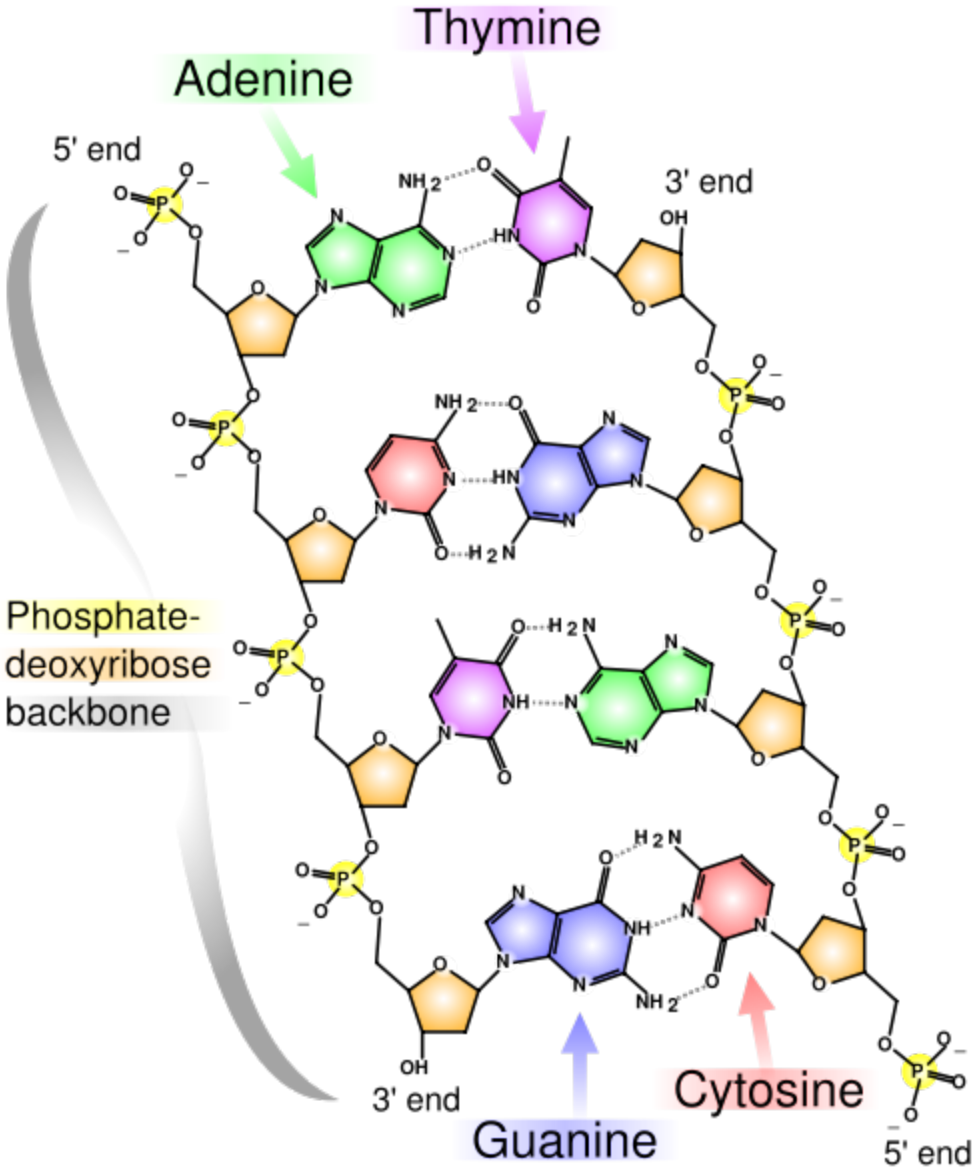




DNA

Nucleotides can make single stranded polymers called ssDNA (single stranded DNA).

Complementary ssDNAs can *hybridize* to form dsDNA (double stranded DNA).



Drawing DNA

You have to say which way the DNA goes, usually from 5’ to 3’.

A single stranded DNA:

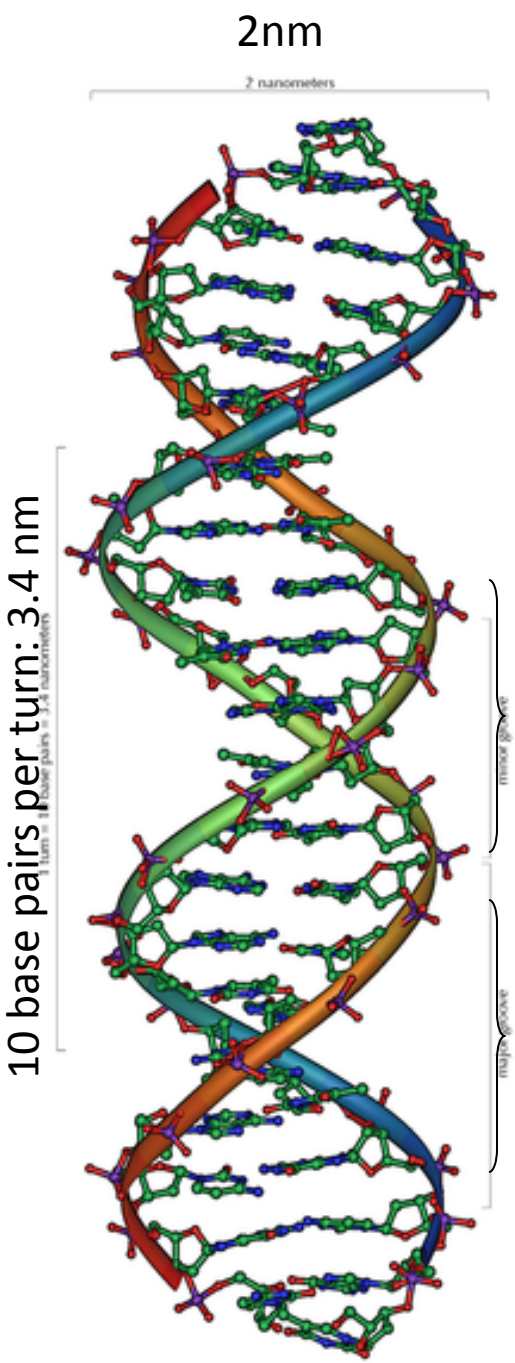
```
5’-ATCCTGTAATGC-3’
```

A double stranded DNA:

```
5’-ATCCTGTAATGC-3’
   ||||||||||||
3’-TAGGACATTACG-5’
```

Sometimes, we write dsDNA by writing only one strand, since the other is implied, as in genebank data.

The double helix

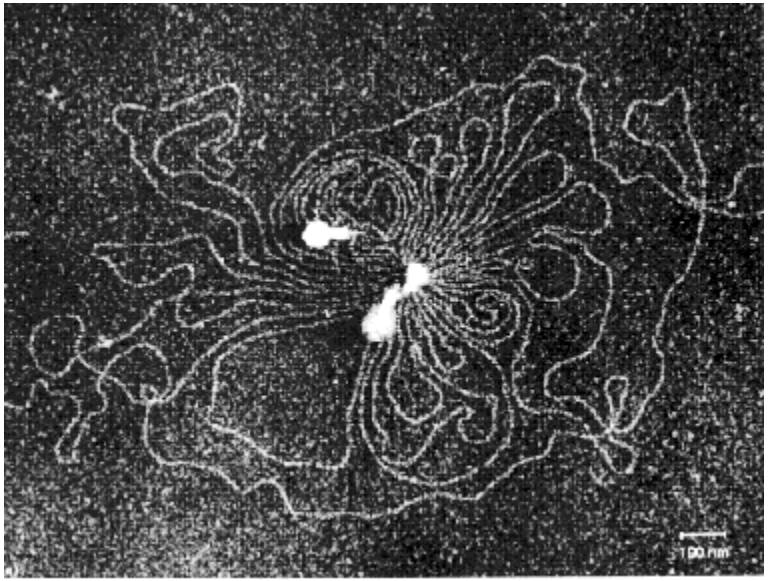


The information of DNA is encoded in the sequence of bases.

The sequence can also effect the detailed structure of the major and minor grooves.

Repressor and promotor proteins can bind with high specificity to the “outside” of the DNA helix (more on this next week).

DNA is long!



DNA is stored under pressure in virus capsids. This 160,000 bp piece of DNA exploded from within a T4 bacteriophage.

E. Coli genome = 4.6 Mbp (one molecule!)

$$L = 4,600,000bp \times \frac{3.4nm}{10bp} \approx 1.6mm$$

human genome = 2m long!

Important Molecules

DNA = Deoxyribonucleic acid
A sequence of A, T, C and G (deoxyribonucleotides)

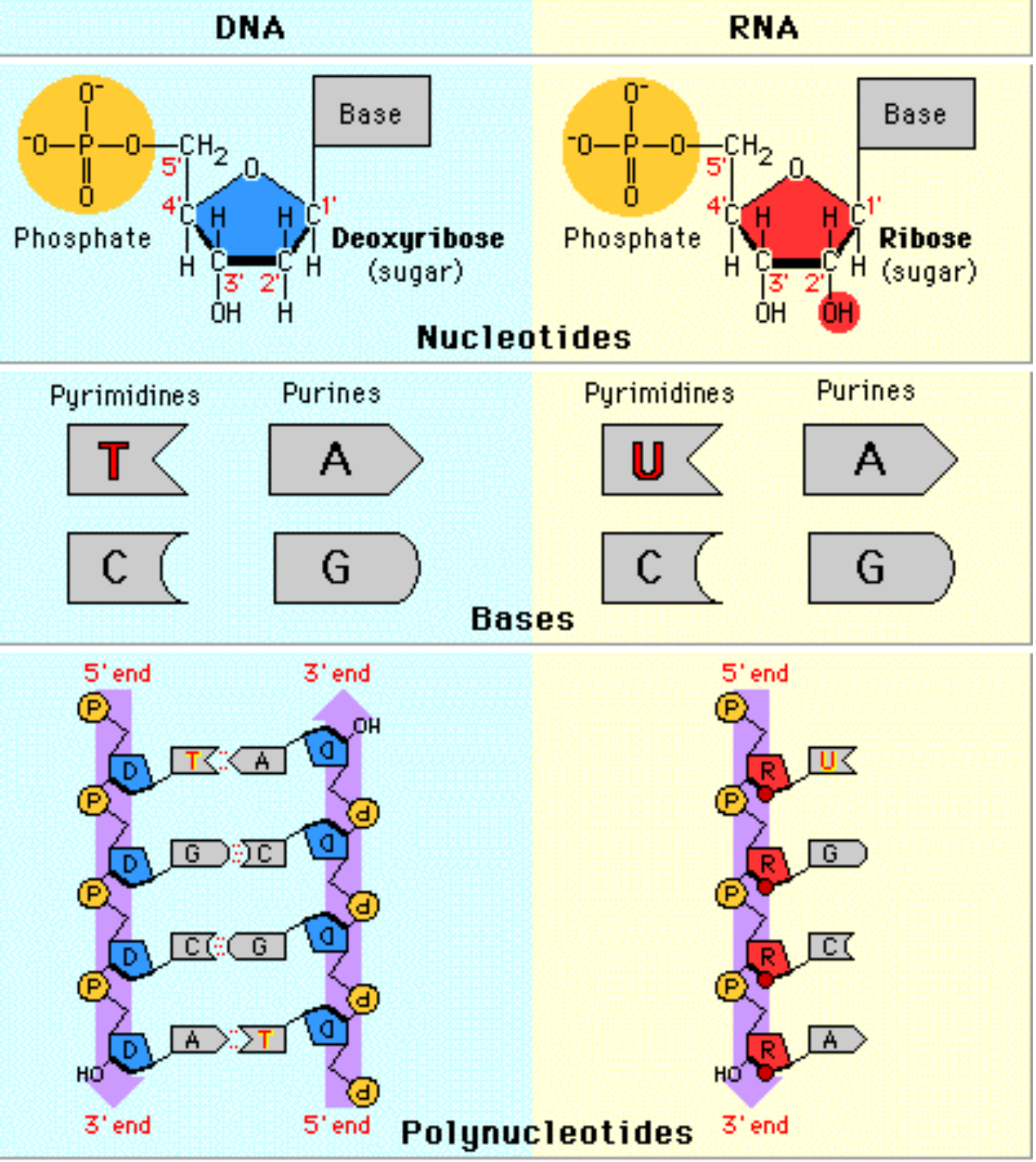
RNA = Ribonucleic Acid (mRNA, tRNA, …)
A sequence of A, U, C and G (ribonucleotides)

RNA Polymerase (RNAP)
Transcribes (copies) DNA segments into RNA

Amino Acids and Transfer RNA (tRNA)
Help build proteins

The Ribosome
Translates messenger RNA (mRNA) into protein

Protein
A sequence of amino acids



DNA is long.

RNA is short.

DNA is stable.

RNAs appear and disappear.

DNA is double stranded.

RNA is single stranded and sometimes folds up on itself in funny ways.

DNA stores information.

RNA does <u>many</u> things!

Important Molecules

DNA = Deoxyribonucleic acid
A sequence of A, T, C and G (deoxyribonucleotides)

RNA = Ribonucleic Acid (mRNA, tRNA, …)
A sequence of A, U, C and G (ribonucleotides)

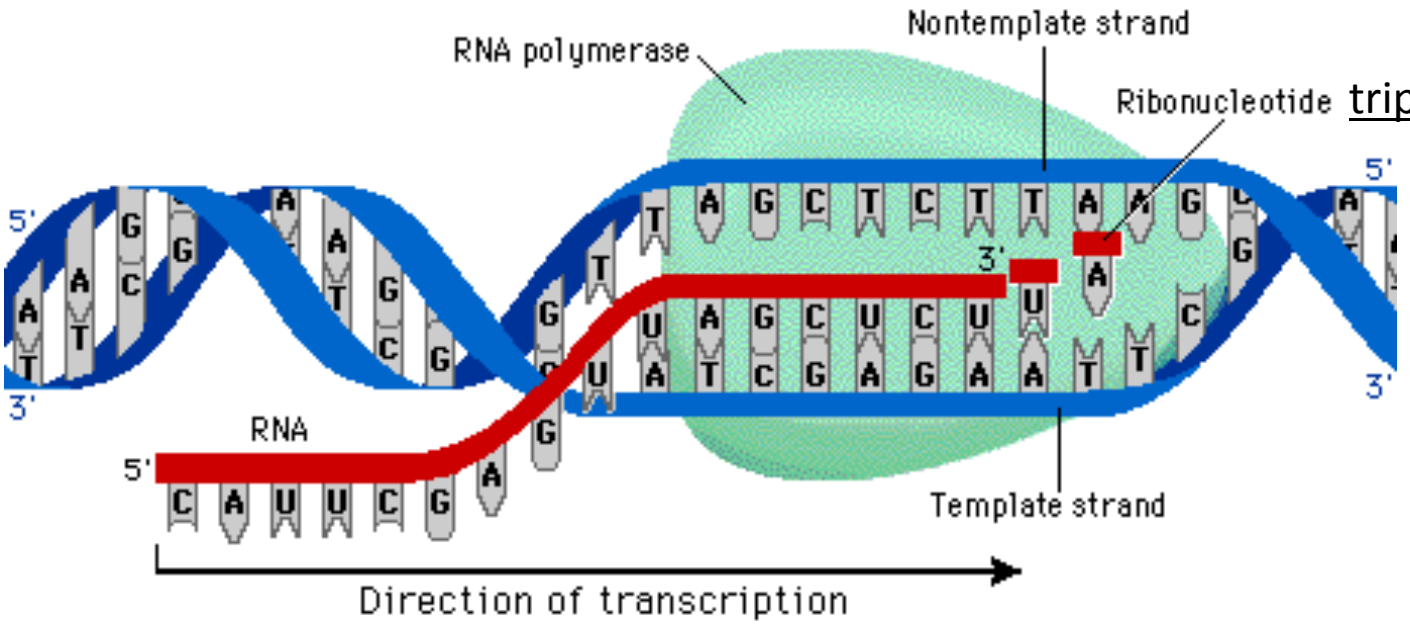
RNA Polymerase (RNAP)
Transcribes (copies) DNA segments into RNA

Amino Acids and Transfer RNA (tRNA)
Help build proteins

The Ribosome
Translates messenger RNA (mRNA) into protein

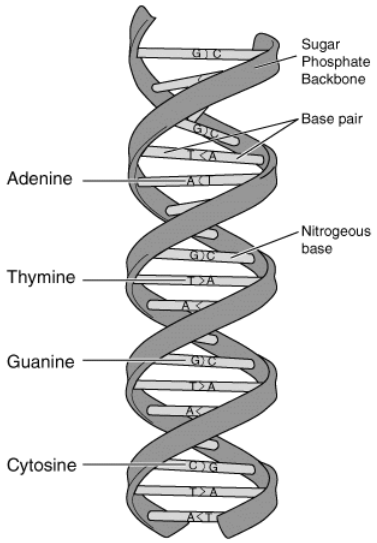
Protein
A sequence of amino acids

RNA is transcribed from DNA templates



- The region that expresses an RNA is called a *gene.*
- Most mRNAs are 3000 bp or, usually, shorter.

All types of RNA are generated this way



Transcription

mRNA ← Sometimes several mRNAs are copied in a row to produce a string of protein coding mRNAs.

tRNA (e. coli has ~60 tRNA genes)

siRNA

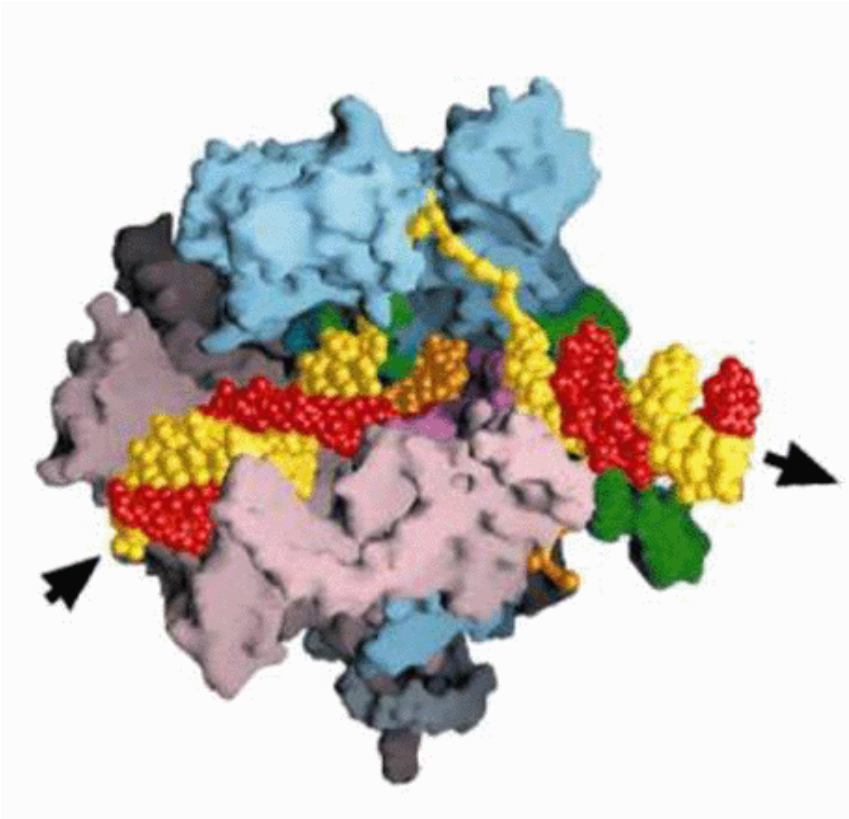
microRNA

Eukaryotes

Ribosomal RNA (e. coli has 7 copies of its rRNA genes so that it can produce a lot of ribosomes)

RNA is totally the coolest molecule.

RNA Polymerase (RNAP)



RNAP is a *holoenzyme*, consisting of 4 proteins subunits and a detachable cofactor enzyme (σ-cofactor).

The workhorse of transcription is RNAP.

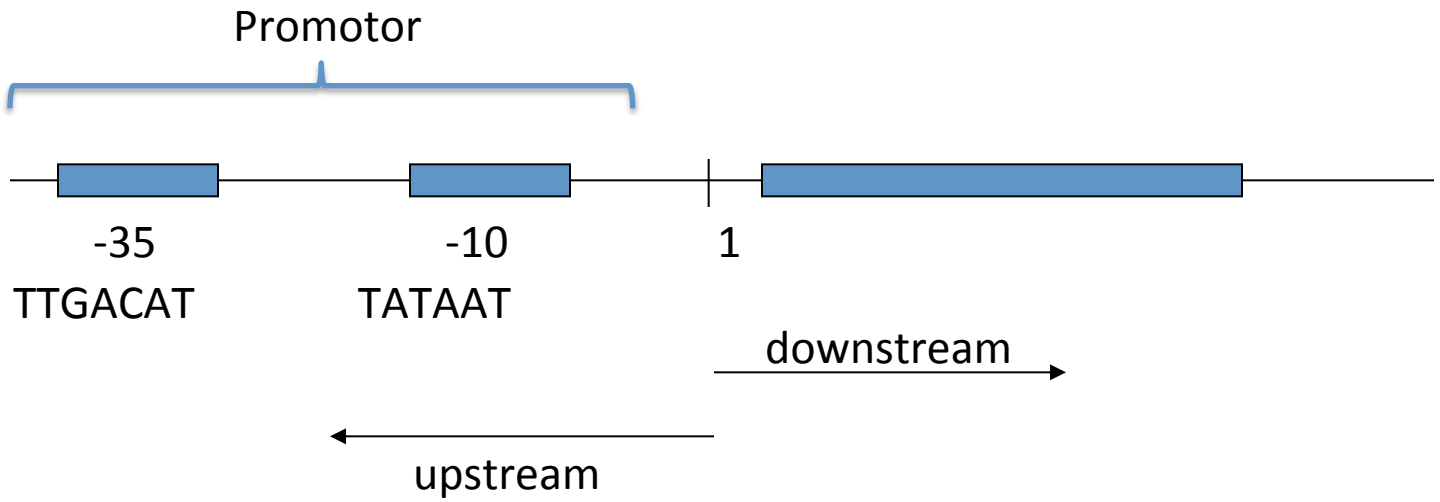
RNAP catalyzes the formation of the phosphodiester bond linking the nucleotides together in RNA.

RNAP moves at about 20 nucleotides per second!

Some genes are transcribed faster than others.

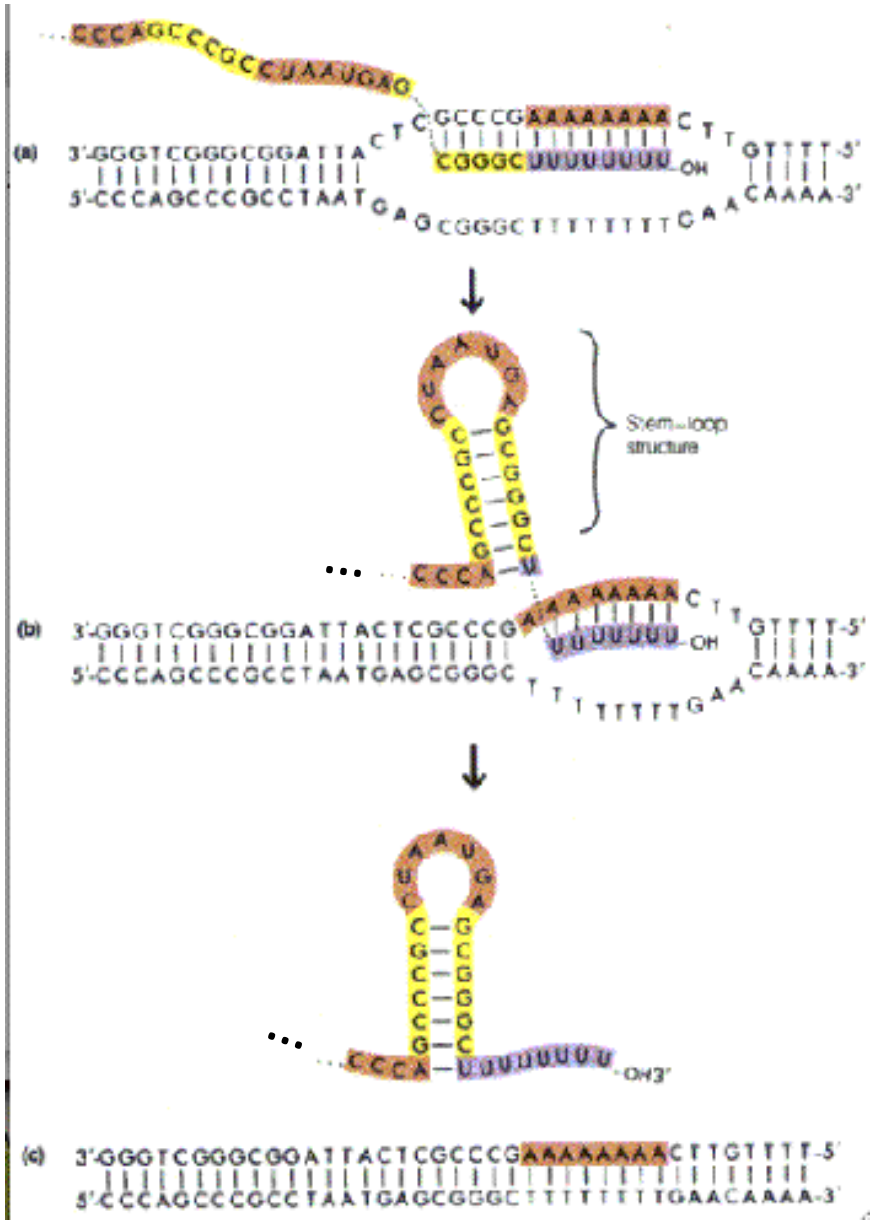
If the wrong nucleotide is added, RNAP backs up and fixes its mistakes!

Start signals are coded in DNA



1. RNAP slides along DNA without transcribing.
2. σ−factor binds to promotor
3. RNAP binds with σ-factor until about 10 bases are transcribed.
4. When you write DNA programs, these are the things you specify!

Stop signals are encoded in DNA



- Terminators are AAAAAA sequences preceded by a short palindrome.
- The palindrome forms a *hairpin* in the growing RNA.
- The shape of the hairpined RNA pops RNAP off the gene and transcription stops!

Important Molecules

DNA = Deoxyribonucleic acid
A sequence of A, T, C and G (deoxyribonucleotides)

RNA = Ribonucleic Acid (mRNA, tRNA, …)
A sequence of A, U, C and G (ribonucleotides)

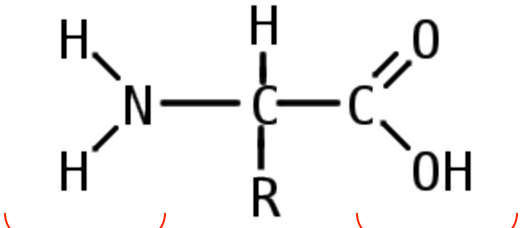
RNA Polymerase (RNAP)
Transcribes (copies) DNA segments into RNA

Amino Acids and Transfer RNA (tRNA)
Help build proteins

The Ribosome
Translates messenger RNA (mRNA) into protein

Protein
A sequence of amino acids

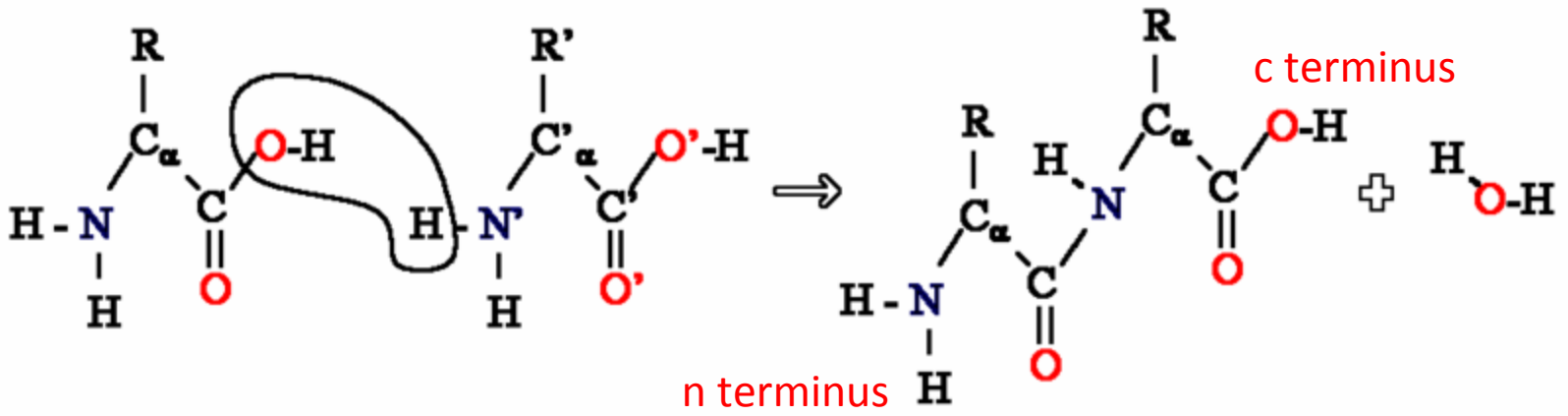
Amino Acids



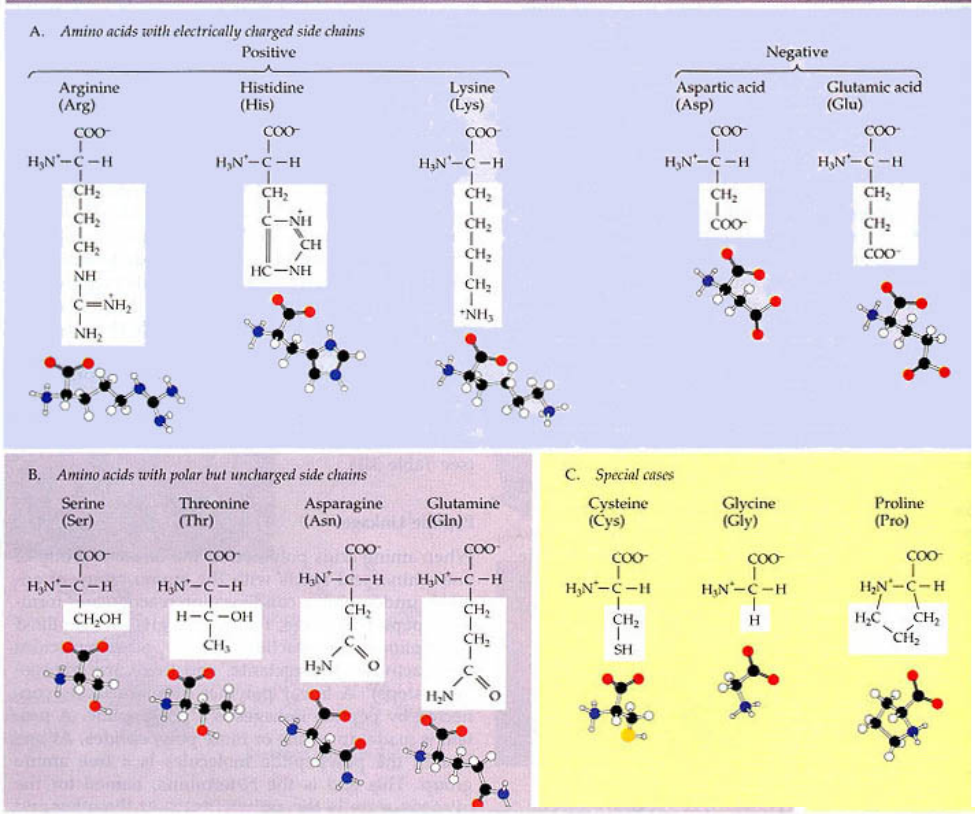
Amino

Carboxyl Group

One of 20 *residues*



Amino acids *condense* to form chains called *polypeptides.*



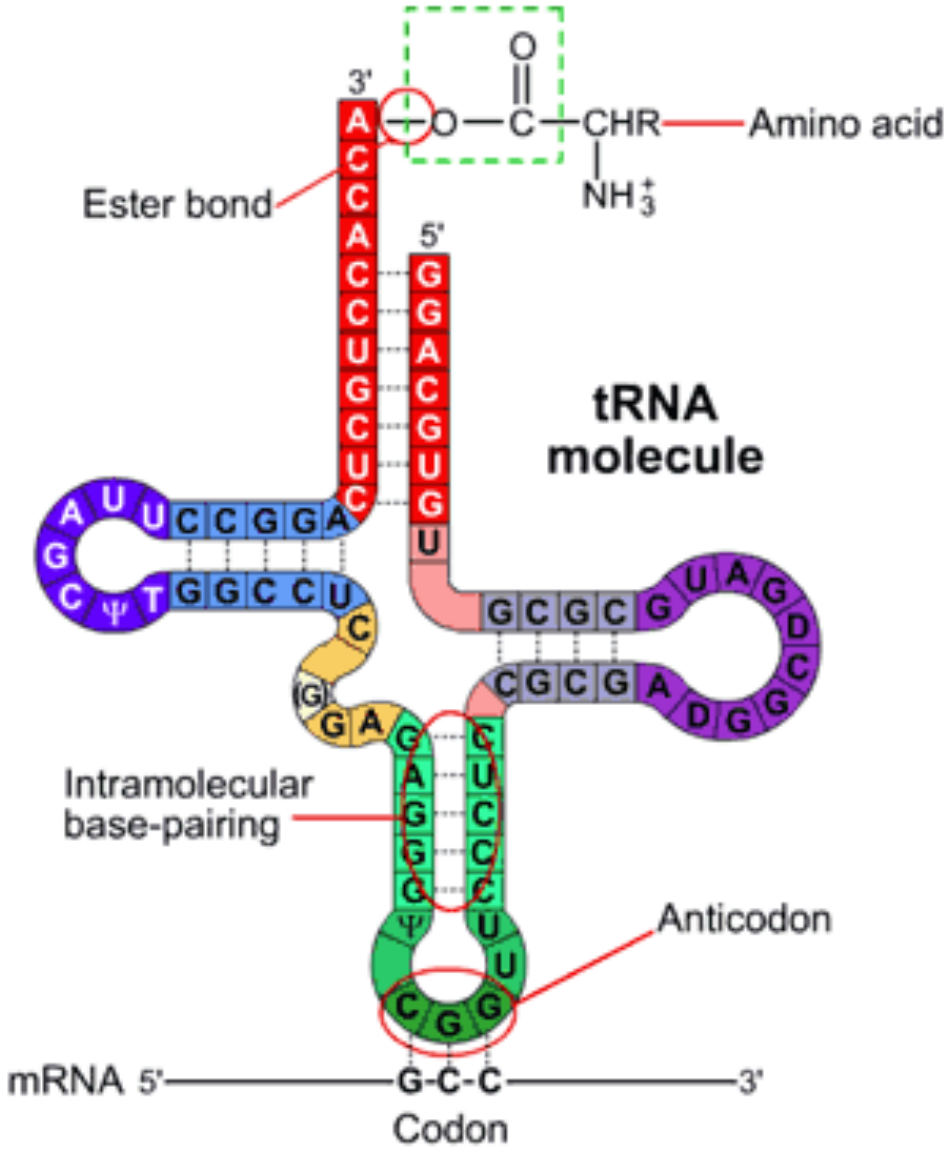
The 20 Amino Acids

Each one has a different shape, charge, and hydrophobcity.

By linking up some into a long chain, you can make a little machine.

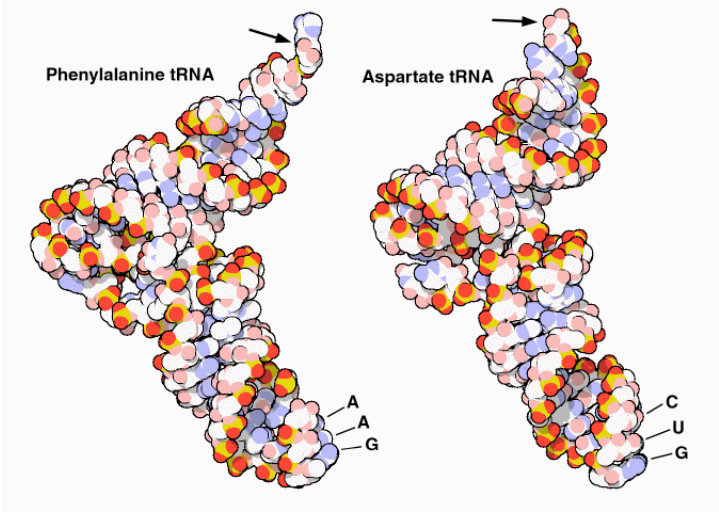
We know how to make some kinds of machines this way, but for the most part we don't know much about how to do this in a principled manner.

Transfer RNA (tRNA)

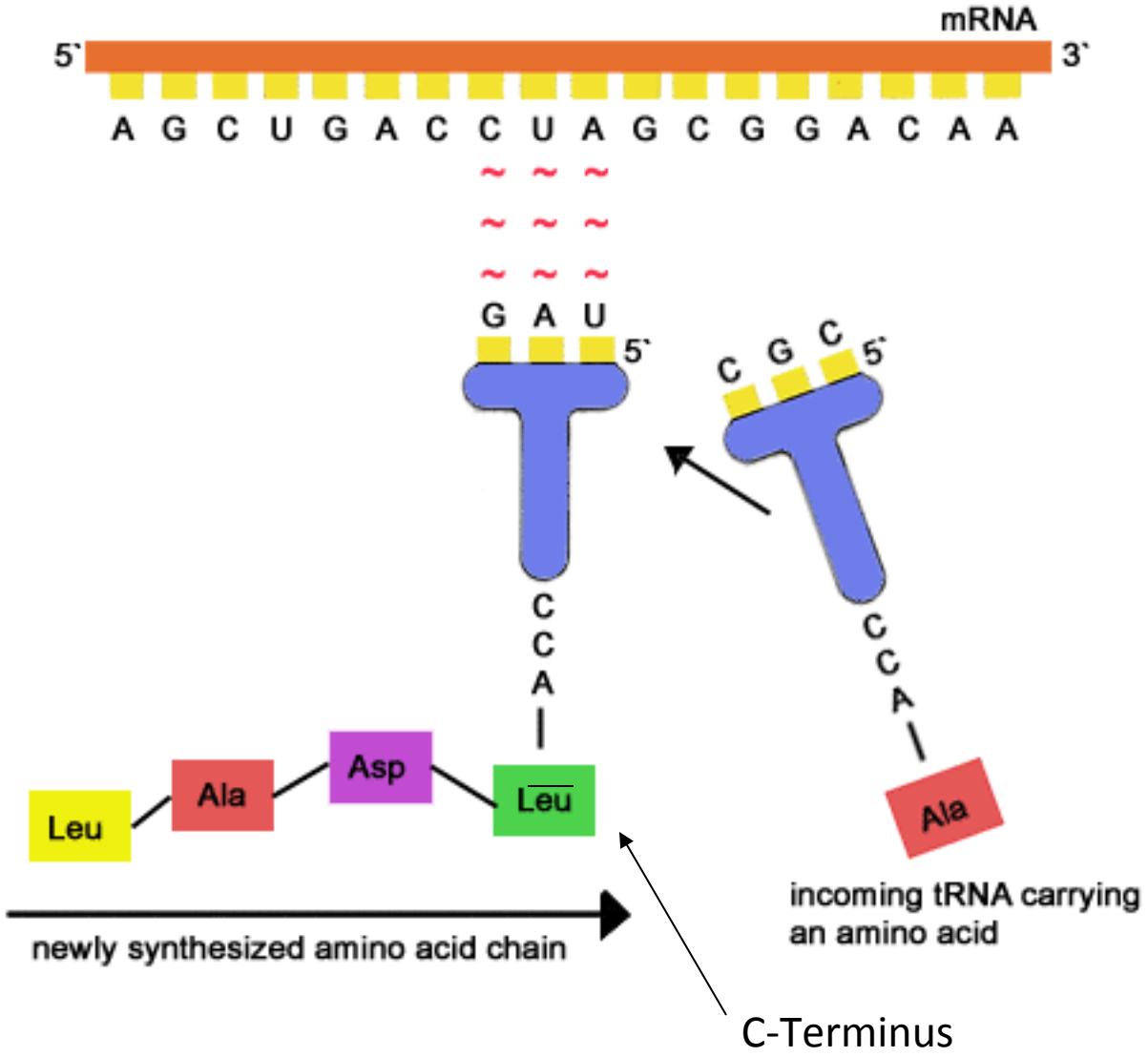


- Each amino acid can have several tRNAs, one for each codon variation.

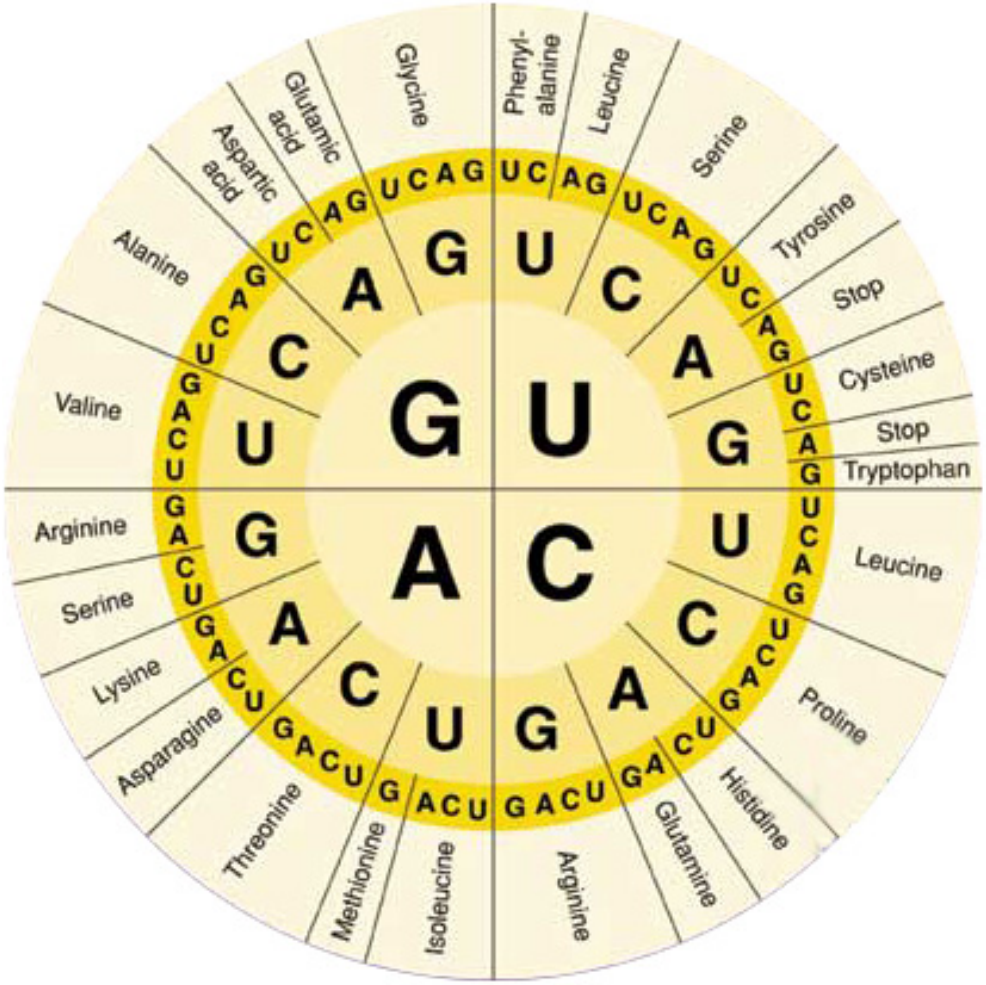
- Various tRNA synthetases and other enzymes provide a post-translational modification that adds the amino acid.



Proteins grow at the carboxyl group (C-terminus)



RNA is translated to protein



- Each 3 nucleotide sequence, called a *codon,* in mRNA (from 5’ to 3’) codes for an amino acid.

- Translation is the process of building the corresponding protein from this code.

5’-CUC AGC GUU ACC-3’

Leu Ser Val Thr

The start and stop codons

- Translation starts with the codon AUG.
  - So all proteins start with *methionine*.

- Translation stops with UAA, UAG or UGA.
  - These do not code for amino acids.

Important Molecules

DNA = Deoxyribonucleic acid
A sequence of A, T, C and G (deoxyribonucleotides)

RNA = Ribonucleic Acid (mRNA, tRNA, …)
A sequence of A, U, C and G (ribonucleotides)

RNA Polymerase (RNAP)
Transcribes (copies) DNA segments into RNA

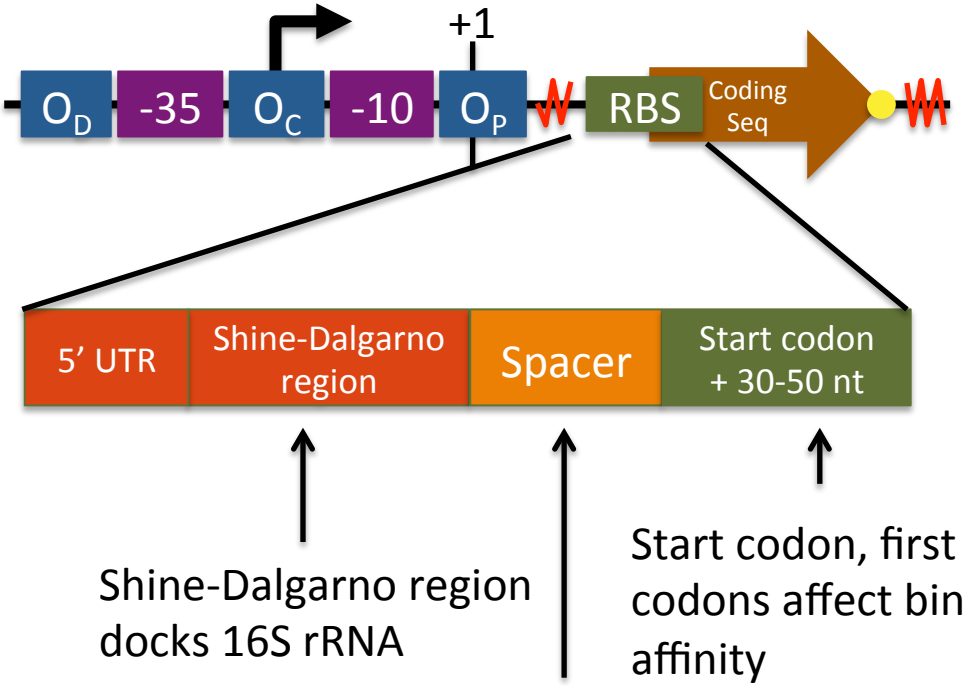
Amino Acids and Transfer RNA (tRNA)
Help build proteins

The Ribosome
Translates messenger RNA (mRNA) into protein

Protein
A sequence of amino acids

The Ribosome Binding Site

Ribosomes bind to specific sequences on mRNAs to initiate translation.



Consensus RBS Sequences

Prokaryotic (Shine-Dalgarno sequence)

+1

5' – AGGAGGACAGCUAUG – 3'

RBS spacer initiator

Eukaryotic (Kozak sequence)

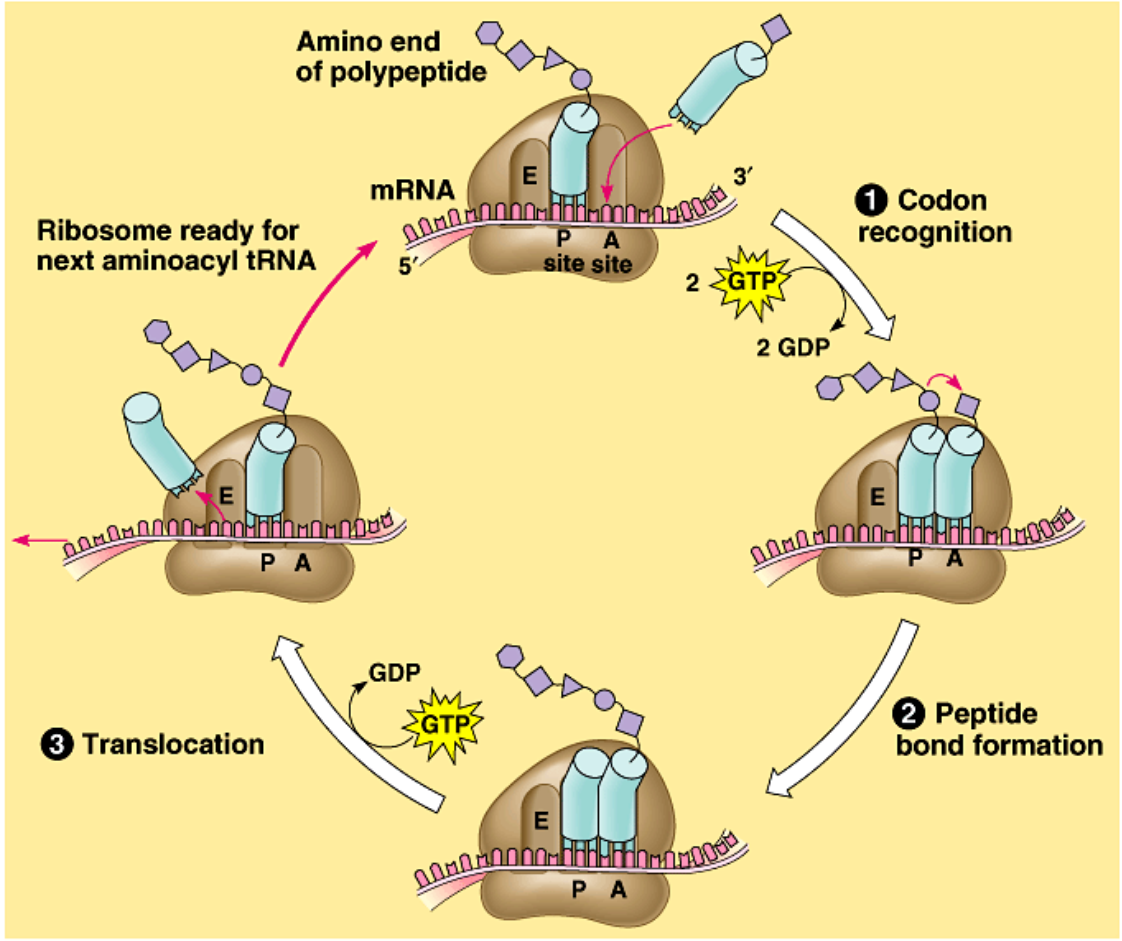
+1

5' – A/GCCACCAUGG – 3'

RBS

initiator

Ribosomes translate mRNA to protein

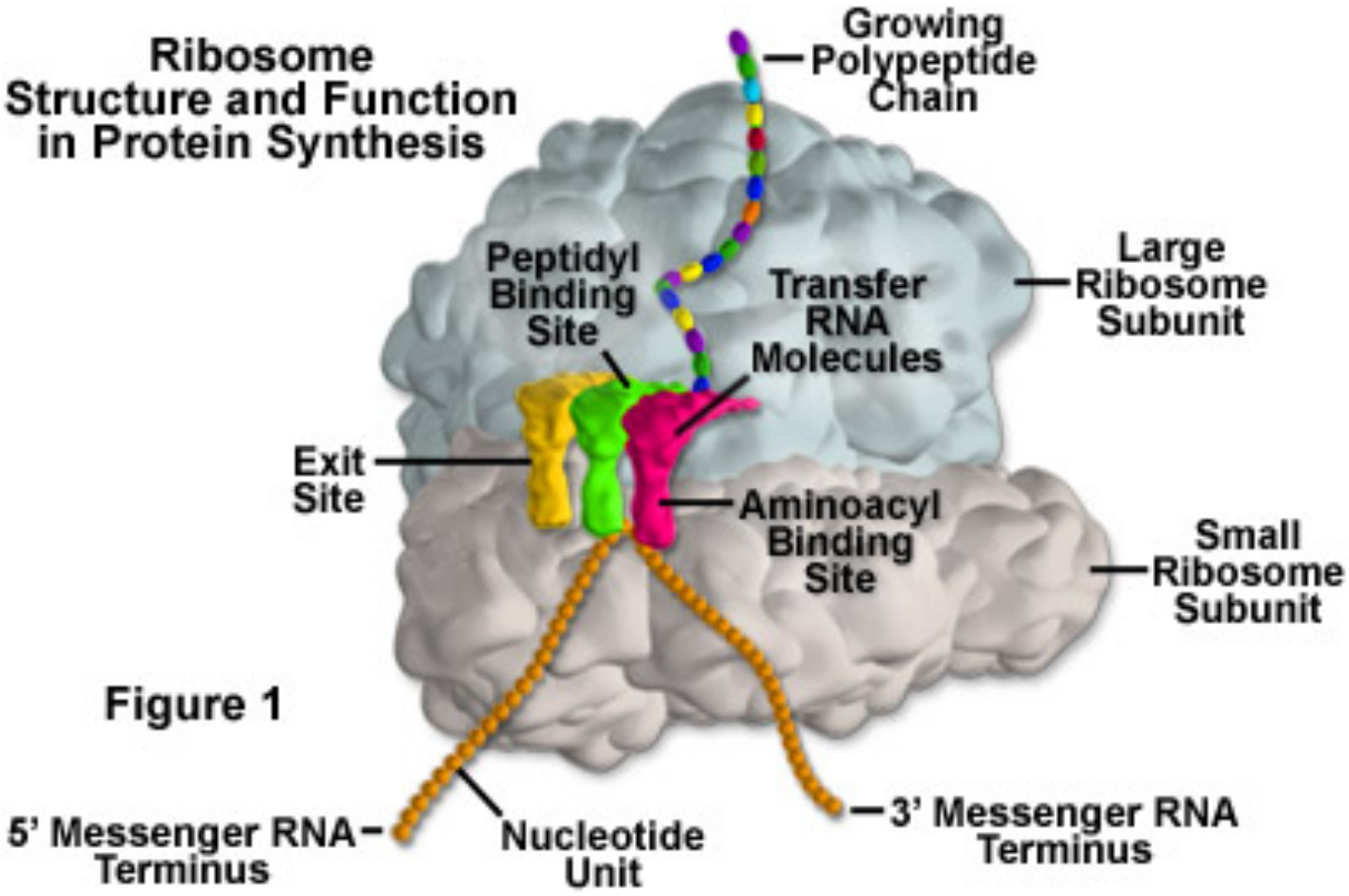


The two subunits of the ribosome are separate until translation starts.

In bacteria, a ribosome processes 20 amino acids per second.

The ribosome makes <u>one</u> mistake every 10000 amino acids

The ribosome (2/3 RNA, 1/3 Protein)



The RNA World

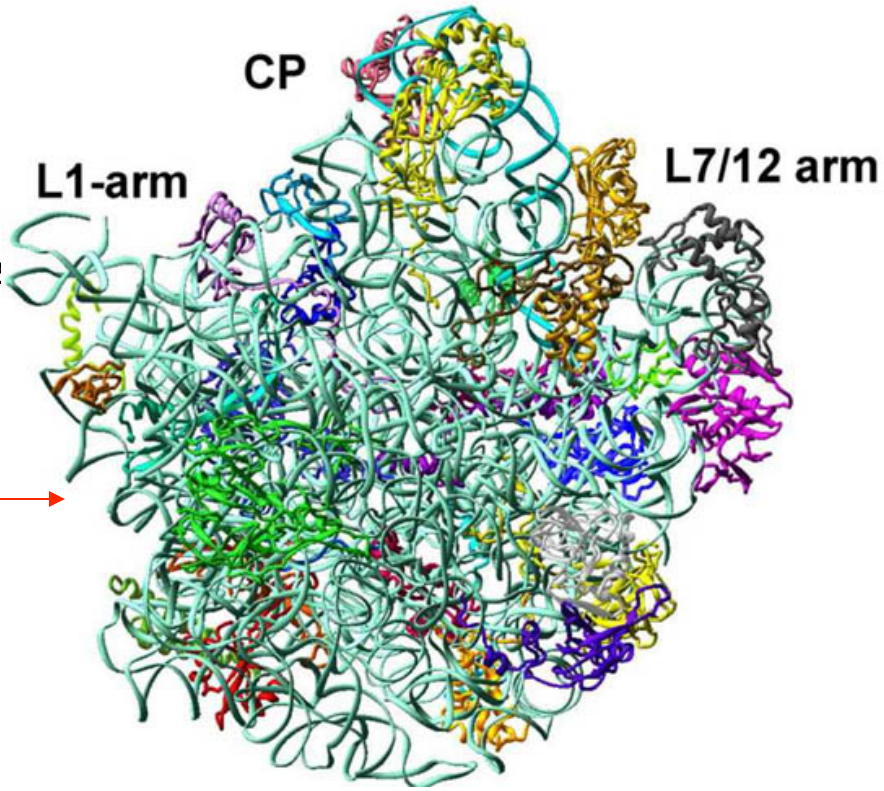
There are those who believe that RNA was th first autocatalytic molecular system.

RNA can cleave and ligate itself.

RNAs have been designed that can transcribe RNA.

The ribosome is made almost entirely out of RNA.

Protein and DNA may have come along later. Having separate molecules for information and structure may have been evolutionarily advantageous.



Important Molecules

- **DNA = Deoxyribonucleic acid**
  A sequence of A, T, C and G (deoxyribonucleotides)

- **RNA = Ribonucleic Acid (mRNA, tRNA, …)**
  A sequence of A, U, C and G (ribonucleotides)

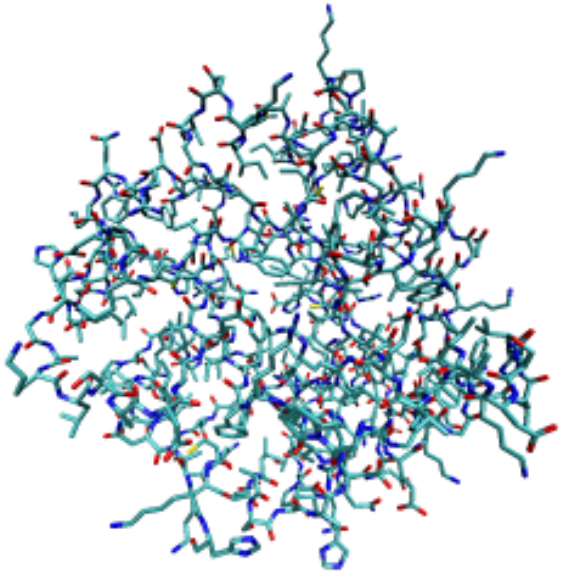
- **RNA Polymerase (RNAP)**
  Transcribes (copies) DNA segments into RNA

- **Amino Acids and Transfer RNA (tRNA)**
  Help build proteins

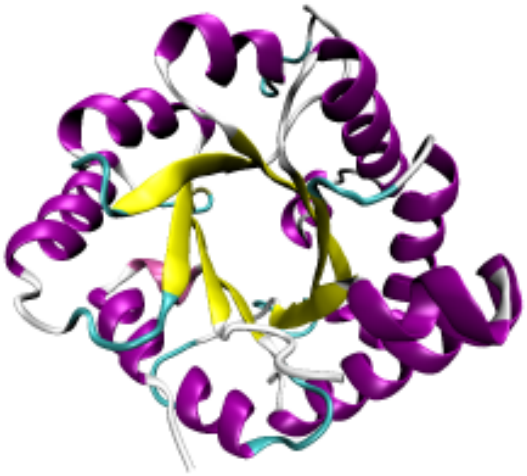
- **The Ribosome**
  Translates messenger RNA (mRNA) into protein

- **Protein**
  A sequence of amino acids

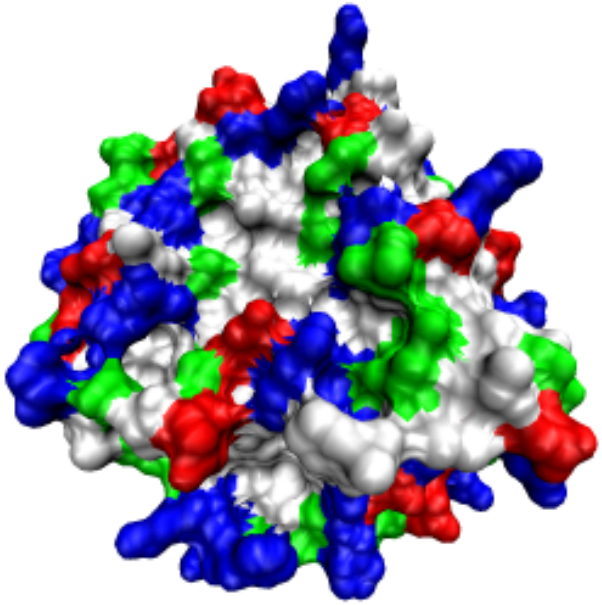
Proteins



Ball and stick



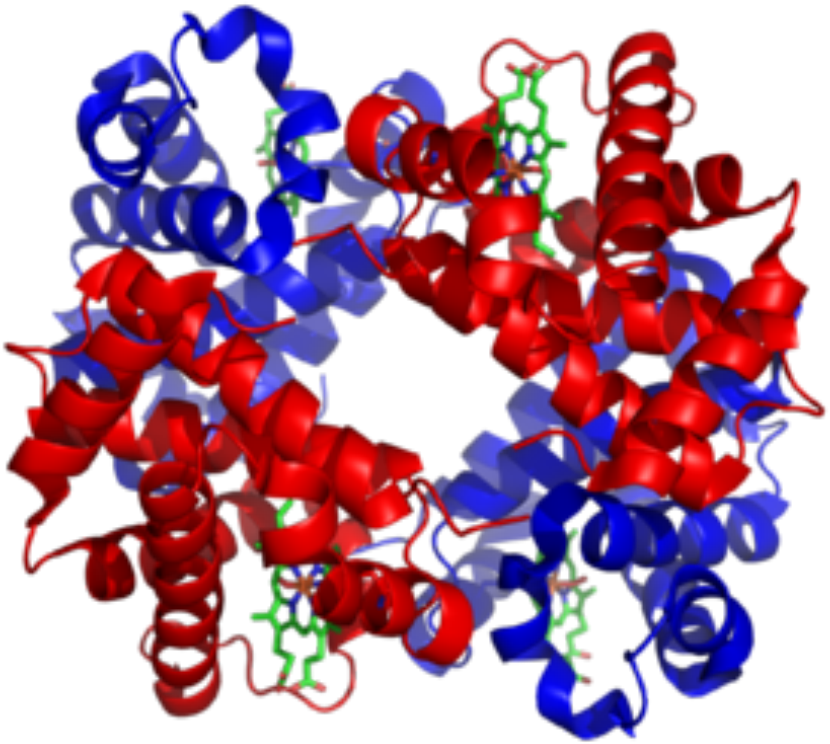
Structure cartoon



Space filling

- Primary structure: The sequence of amino acids
- Secondary structure: The local shape (helix, coil or sheet)
- Tertiary structure: The global 3D shape
- Quaternary structure: How proteins form groups

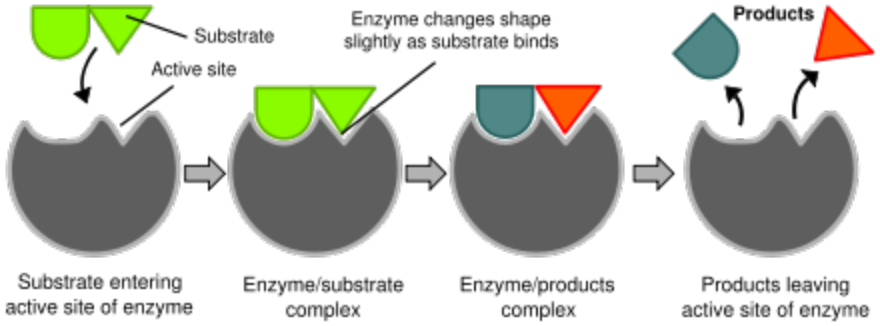
Proteins form groups

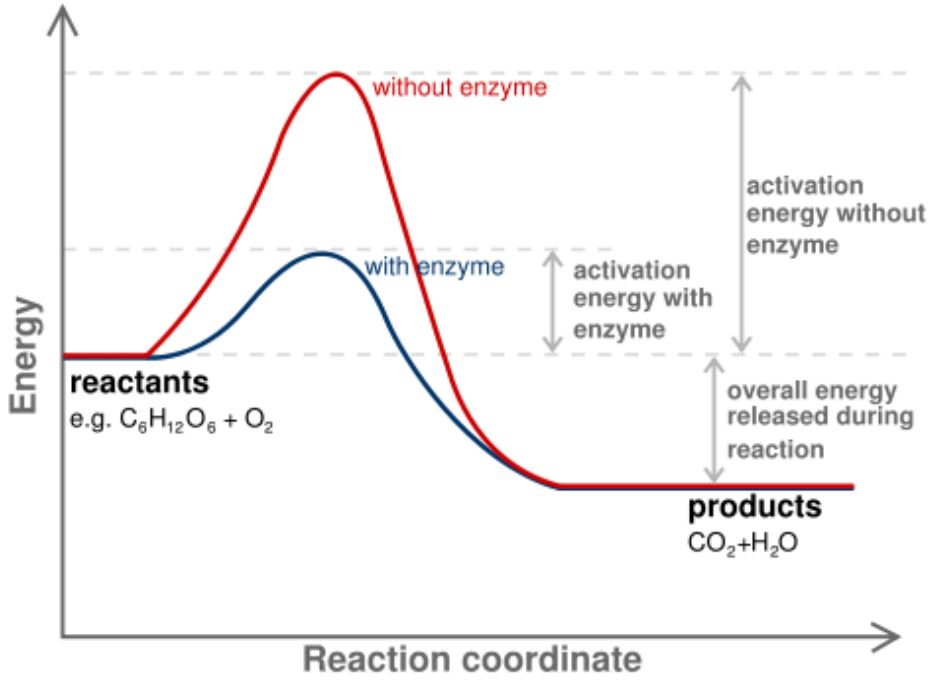


Hemoglobin consists of four protein subunits and four non-protein iron containing heme units. It is self-assembled inside the cell once the components are present.

Proteins can be Enzymes

An enzyme is a proteins that accelerates a chemical reaction, usually very specifically.

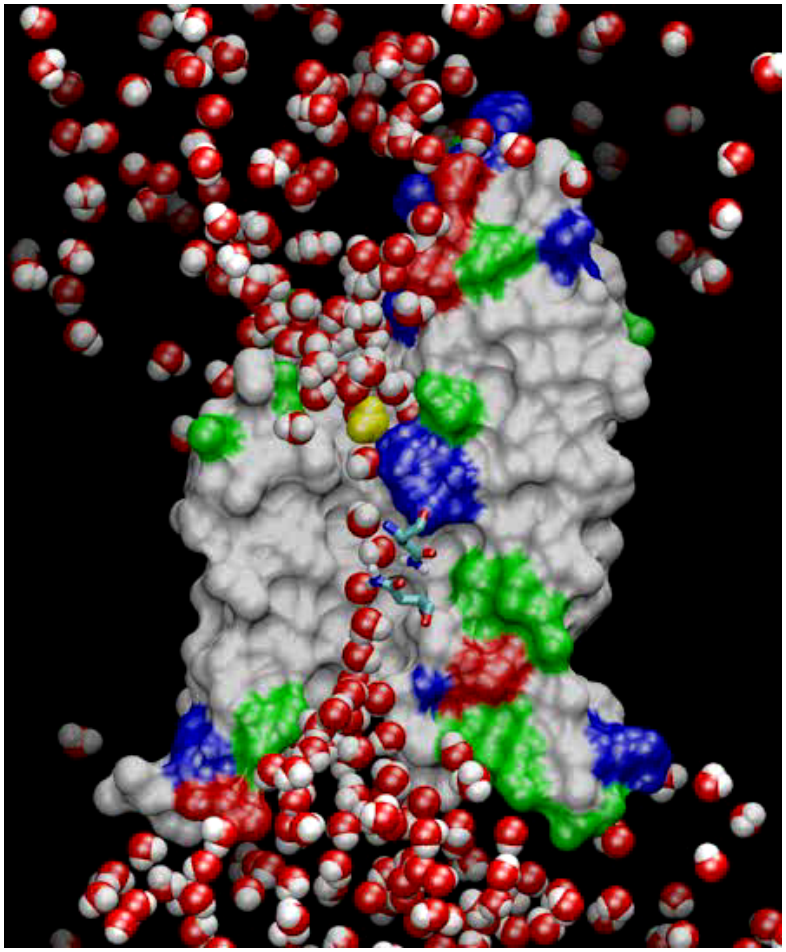




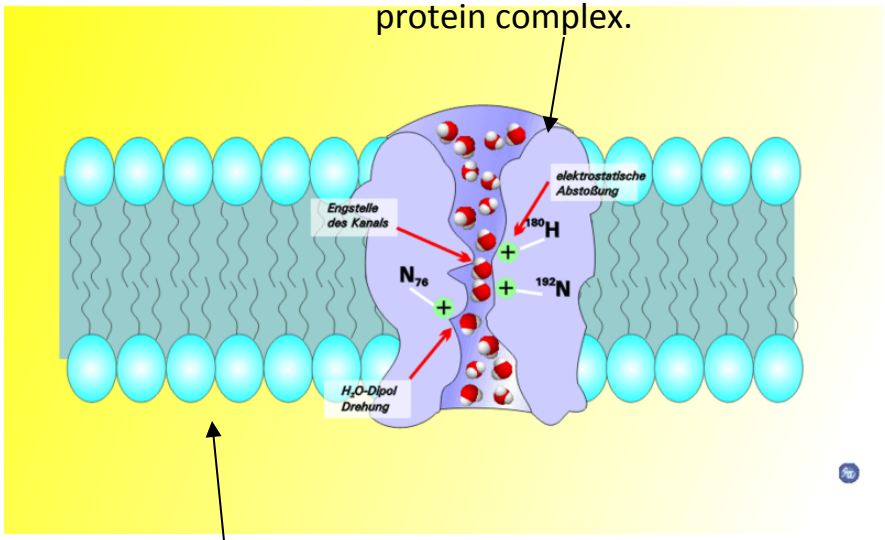
Proteins can pump small molecules

(2003 Nobel Prize to Peter Agre)

A 1 nanosecond simulation of the 60,000 atom [aquaporin-1 water channel] with full electrostatics and constant pressure in a single week (Schulten Group, UIUC).

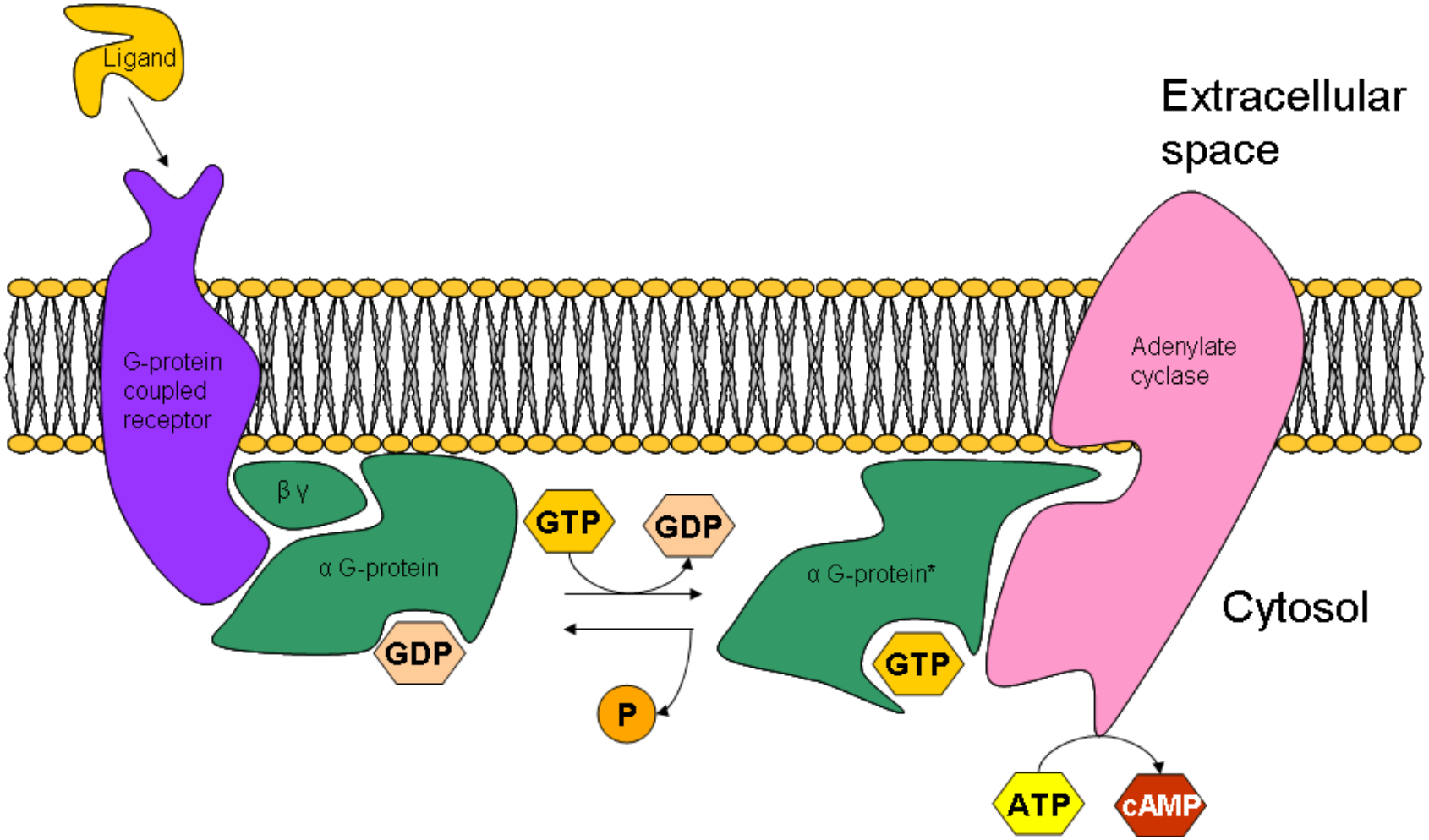


Aquaporin transmembrane four-protein complex.



Phospholipid membrane

Proteins are involved with signaling



RNA and Protein Are Degraded

- RNA is degraded by Ribonucleases (Rnase)
- Protein is degraded by Proteases
- Some RNAs and Proteins are more stable than others.
- Synthetic Biologists can tune degradation rates.

Important Molecules

- **DNA = Deoxyribonucleic acid**
  A sequence of A, T, C and G (deoxyribonucleotides)

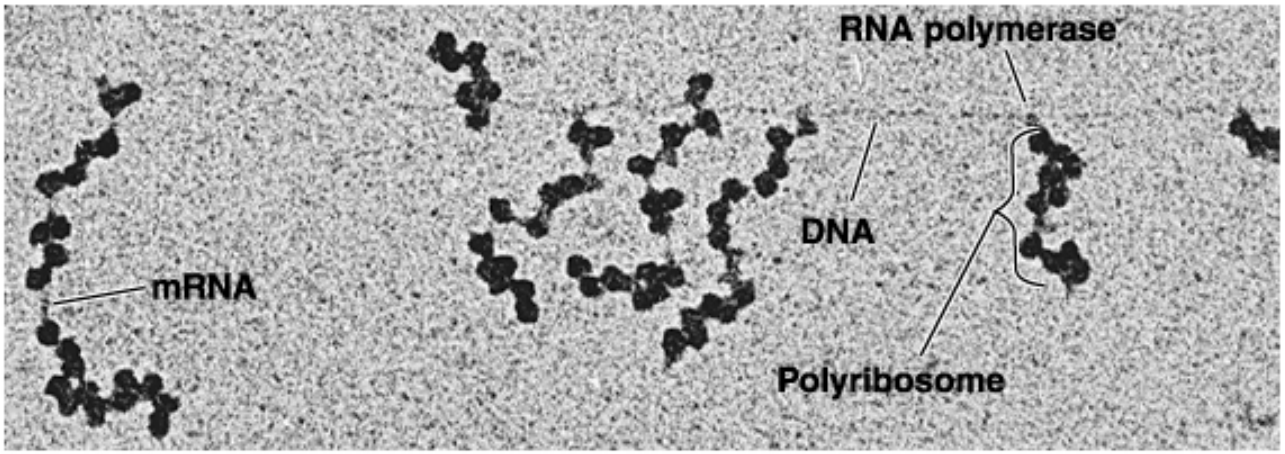
- **RNA = Ribonucleic Acid (mRNA, tRNA, …)**
  A sequence of A, U, C and G (ribonucleotides)

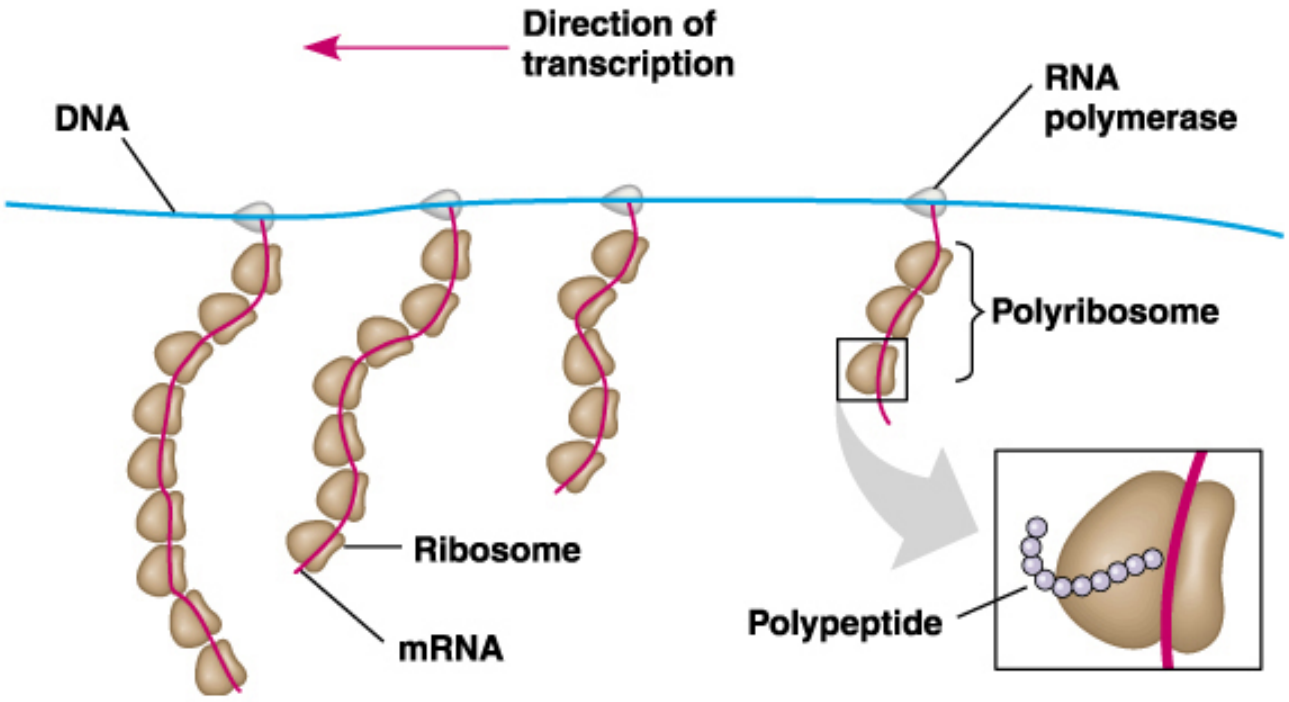
- **RNA Polymerase (RNAP)**
  Transcribes (copies) DNA segments into RNA

- **Amino Acids and Transfer RNA (tRNA)**
  Help build proteins

- **The Ribosome**
  Translates messenger RNA (mRNA) into protein

- **Protein**
  A sequence of amino acids





In prokaryotes, everything happens at once!