

Computational Biophysics: Algorithms to Applications (CS61060)

Instructor: Pralay Mitra
Email: pralay@cse.iitkgp.ac.in

Course Details

- TA
 - Nandita Sharma (nandita.sharma9494@kgpian.iitkgp.ac.in)
 - Shruti Agrawal (shrutiag23@kgpian.iitkgp.ac.in)
 - Suman Kumar Bera (skbera4@kgpian.iitkgp.ac.in)
- Lecture Hours
 - Mon (12:00 -12:55); Tue (10:00-11:55); Thu (08:00-08:55)
- Evaluation
 - TA Evaluation 30
 - Mid Semester 30
 - End Semester 40

Lecture notes and materials

- Will be uploaded at

<https://moodlecse.iitkgp.ac.in/moodle/>

Login to CSE Moodle. Join to the course

CS61060_S22-23 Computational Biophysics: Algorithms to Applications

as Student with enrolment key **spring22-23**

Course Coverage

Bioinformatics Topics:

1. Sequence alignment and comparison
2. Secondary structure prediction
3. Structure alignment
4. Protein folding
5. Protein docking
6. Protein Design

PCB Topics:

1. Protein and DNA structure

CS Topics:

1. Dynamic programming
2. Sequencing and Strings
3. Molecular surface
4. Simulation techniques
5. Relevant topics from ML and DL

Tutorials:

1. Databases and their usages
2. PDB file format
3. Rasmol, Pymol;
4. NW, SW, BLAST
5. PSIPred/PSSPred, DSSP, STRIDE
6. NACCESS
7. TM-align, TM-Score
8. Folding and Docking software

References

- **Text Book:**

1. Neil C. Jones and Pavel A. Pevzner. An Introduction to Bioinformatics Algorithms.
2. Gary D. Stormo. Introduction to Protein-DNA Interactions: Structure, Thermodynamics, and Bioinformatics.
3. Bruce R. Donald. Algorithms in Structural Molecular Biology (Computational Molecular Biology).
4. Dan Gusfield. Algorithms on Strings, Trees and Sequences

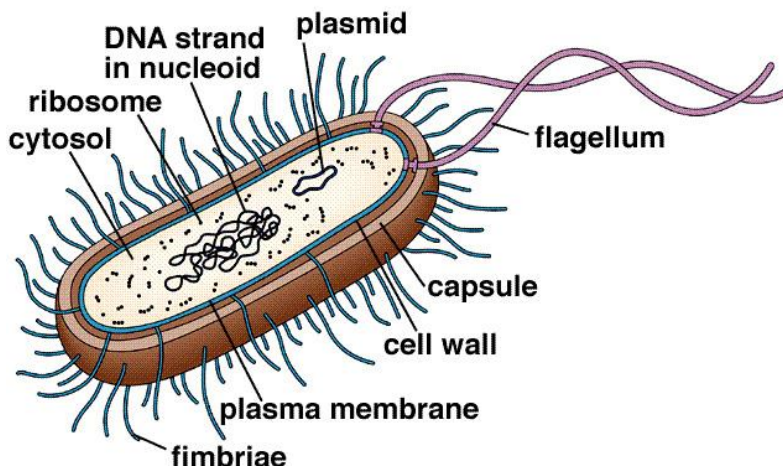
- **Literature:**

1. Christopher M. Dobson, Andrej Sali, Martin Karplus. (1998) Protein Folding: A Perspective from Theory and Experiment. *Angewandte Chemie International Edition* 37(7):868-893.
2. Inbal Halperin, Buyong Ma, Haim Wolfson, Ruth Nussinov. (2002) Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins: Structure, Function, and Bioinformatics* 47(4):409-443.
3. Joan-Emma Shea and Charles L Brooks III. (2001). From Folding Theories To Folding Proteins: A Review and Assessment of Simulation Studies of Protein Folding and Unfolding. *Annual Review of Physical Chemistry* 52:499-535.
4. Literature will be provided from time to time.

Prokaryotic Cell

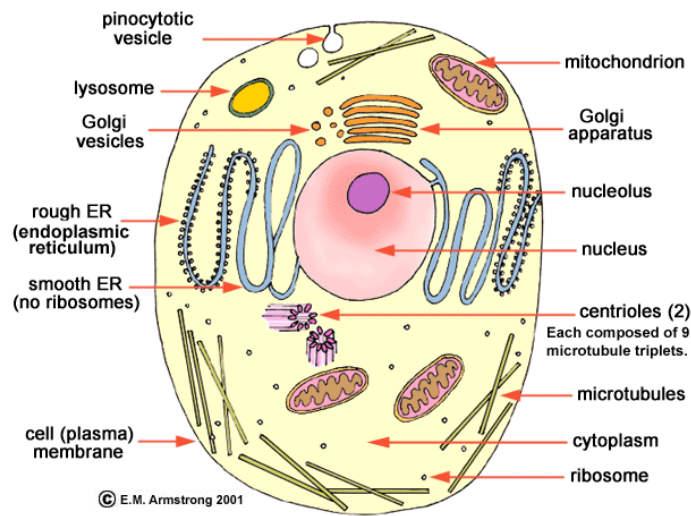
Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

Generalized structure of a prokaryote



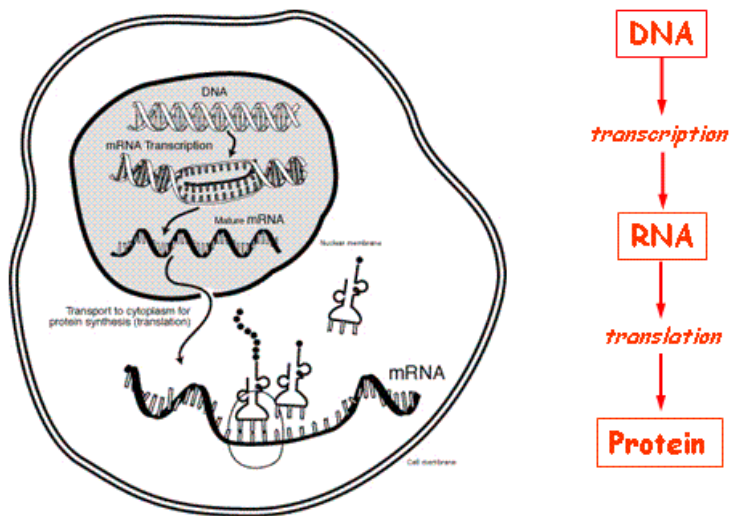
Source: <http://classes.midlandstech.edu/carterp/Courses/bio101/labquiz2/prokaryote.jpg>

Eukaryotic Cell



Source: <http://waynesword.palomar.edu/images/animal4.gif>

Central Dogma of Molecular Biology



Source: http://www.ncbi.nlm.nih.gov/Class/MLACourse/Modules/MolBioReview/central_dogma.html

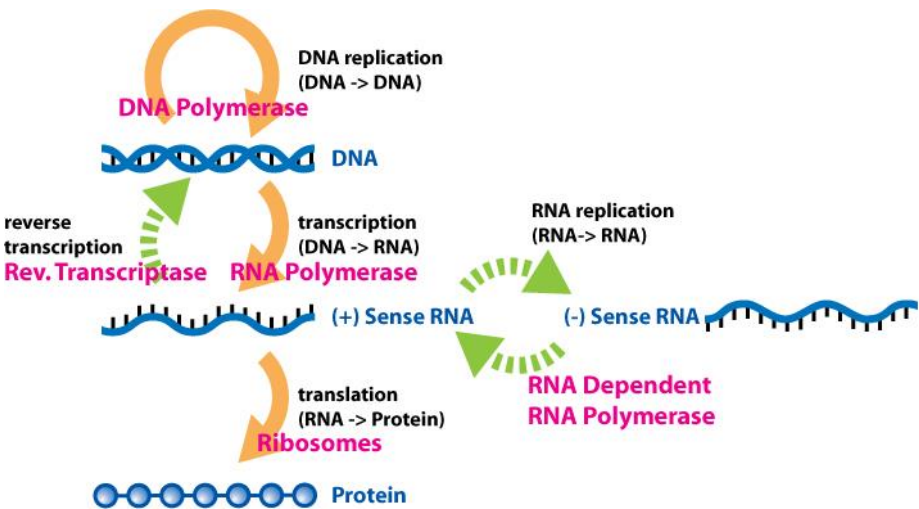
General transfers of biological sequential information

General	Special	Unknown
DNA → DNA	RNA → DNA	protein → DNA
DNA → RNA	RNA → RNA	protein → RNA
RNA → protein	DNA → protein	protein → protein

General transfers of biological sequential information

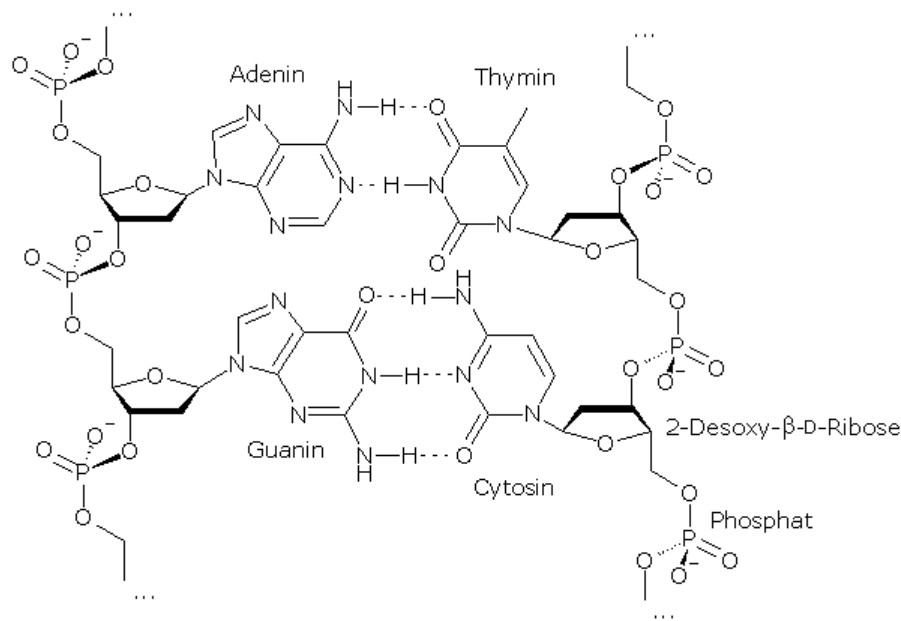
General	Special	Unknown
DNA → DNA (DNA replication)	RNA → DNA (reverse transcription)	protein → DNA
DNA → RNA (transcription)	RNA → RNA (RNA replication)	protein → RNA
RNA → protein (translation)	DNA → protein (direct translation)	protein → protein

Central Dogma with Enzymes

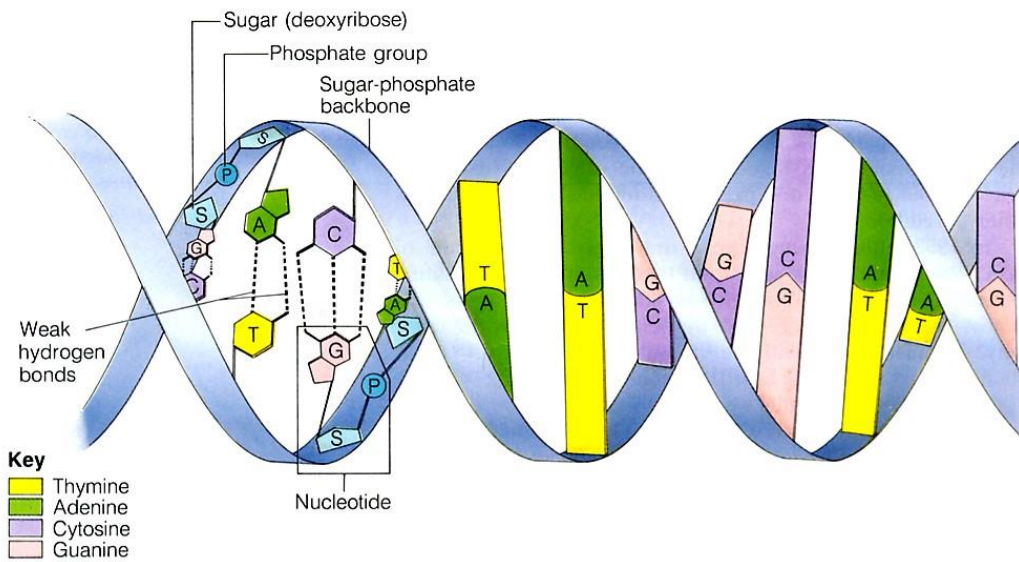


Source: http://en.wikipedia.org/wiki/File:Extended_Central_Dogma_with_Enzymes.jpg

DNA Structure

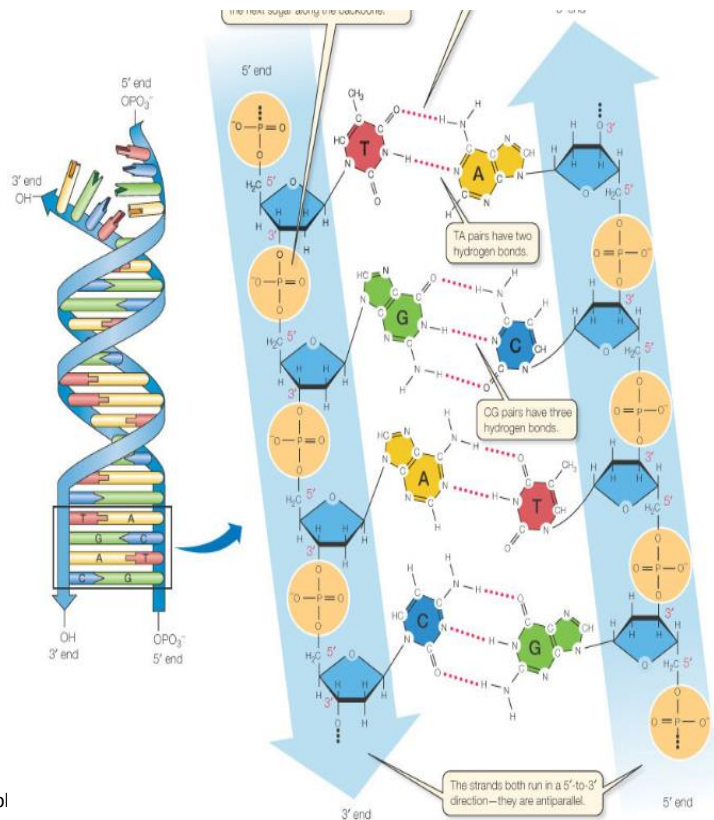


DNA Structure



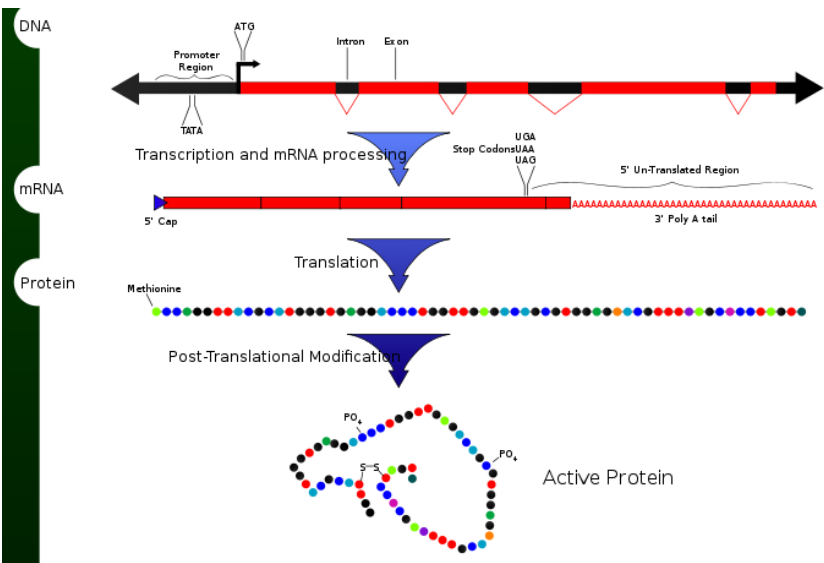
Source: <http://karimedalla.files.wordpress.com/2012/11/dna-structure.jpg>

DNA Structure



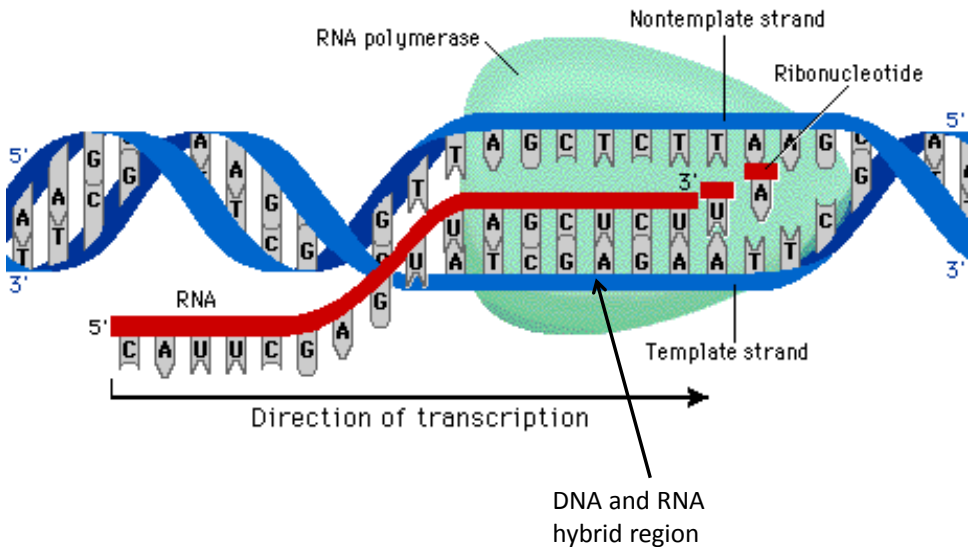
Source: <http://www.nature.com/scitabl>

Central Dogma of Molecular Biology: Eukaryotic Model



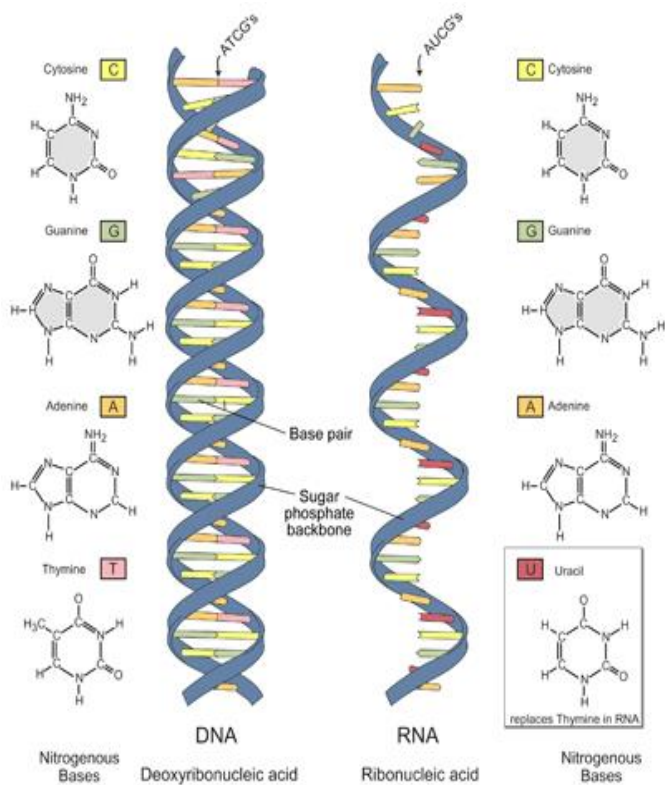
Source: <http://en.wikipedia.org/wiki/File:Cdmb.svg>

Transcription

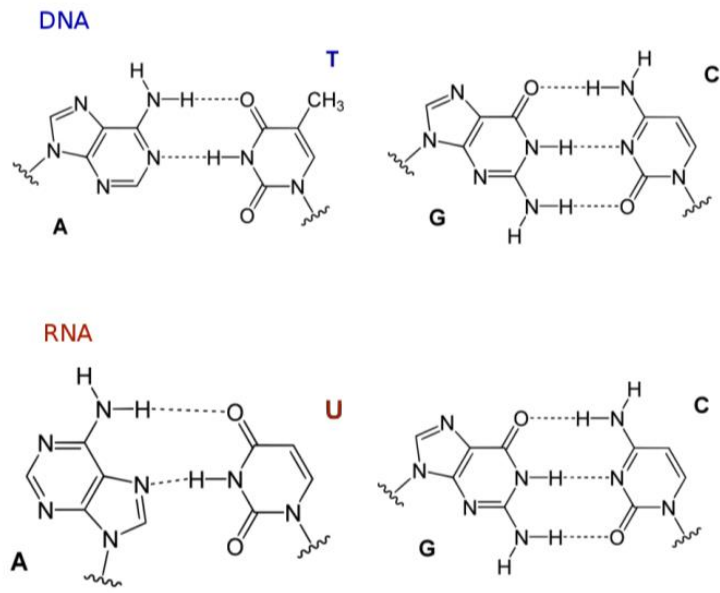


Source: http://www.phschool.com/science/biology_place/biocoach/images/transcription/startrans.gif

DNA vs RNA

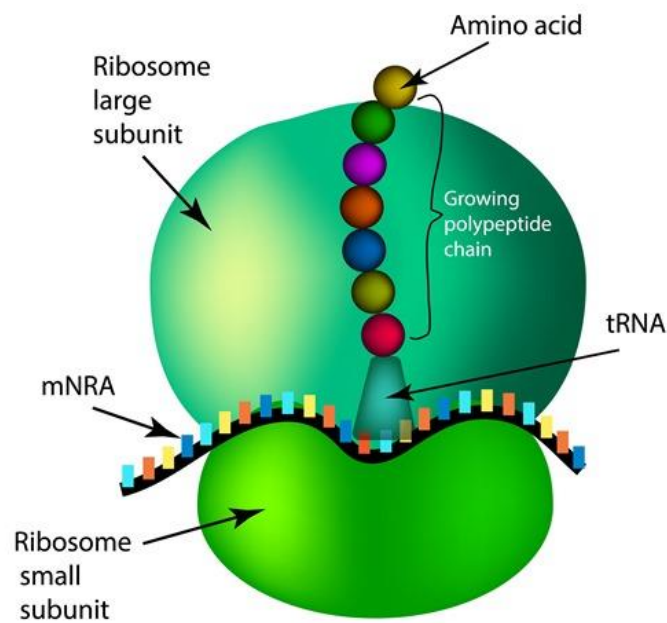


Base Pairs

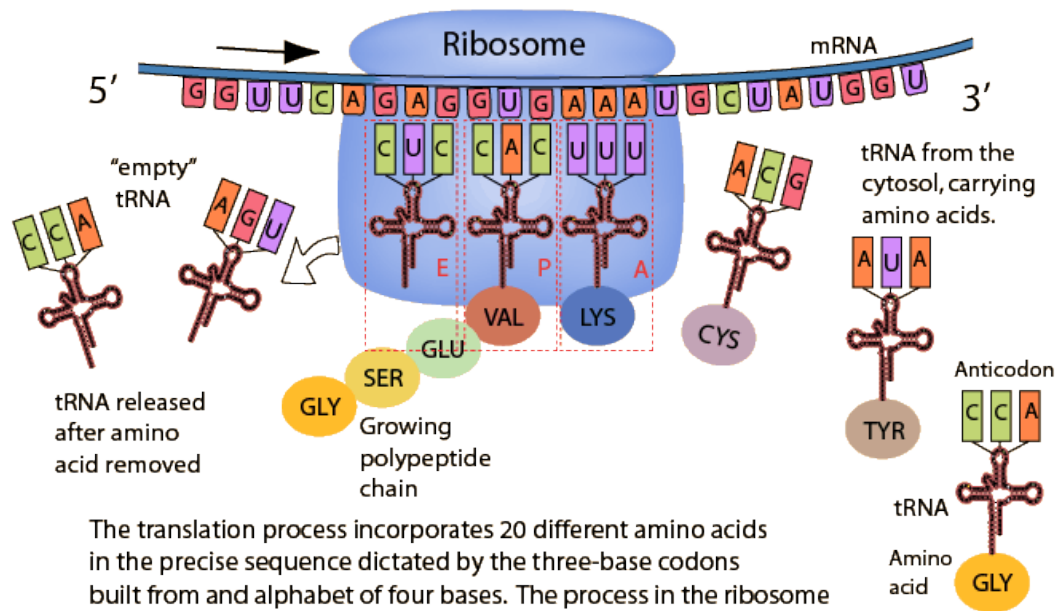


Source: http://andromeda.rutgers.edu/~huskey/images/base_pairs1_w.png

Ribosome



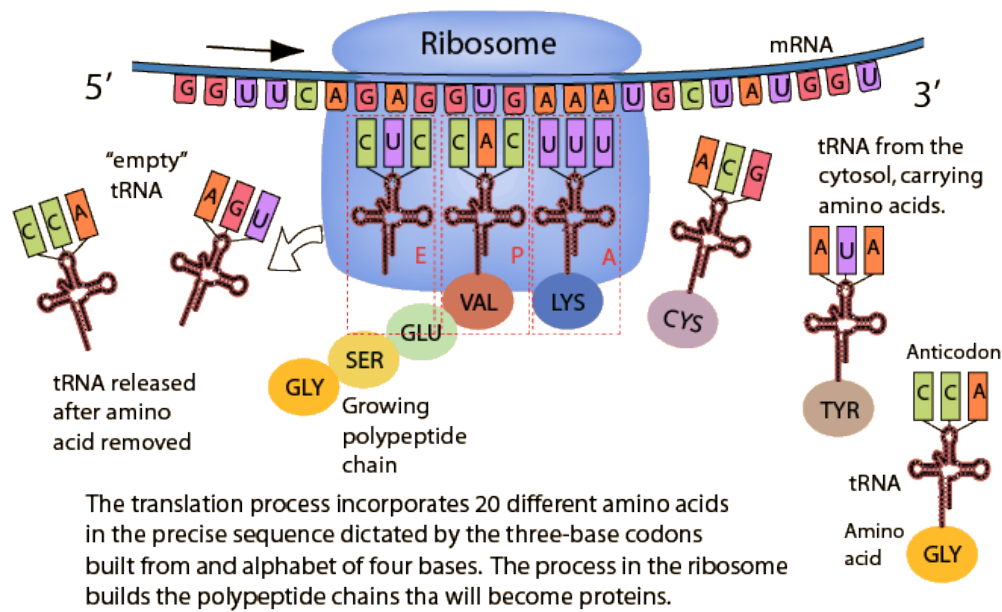
Translation



The translation process incorporates 20 different amino acids in the precise sequence dictated by the three-base codons built from an alphabet of four bases. The process in the ribosome builds the polypeptide chains that will become proteins.

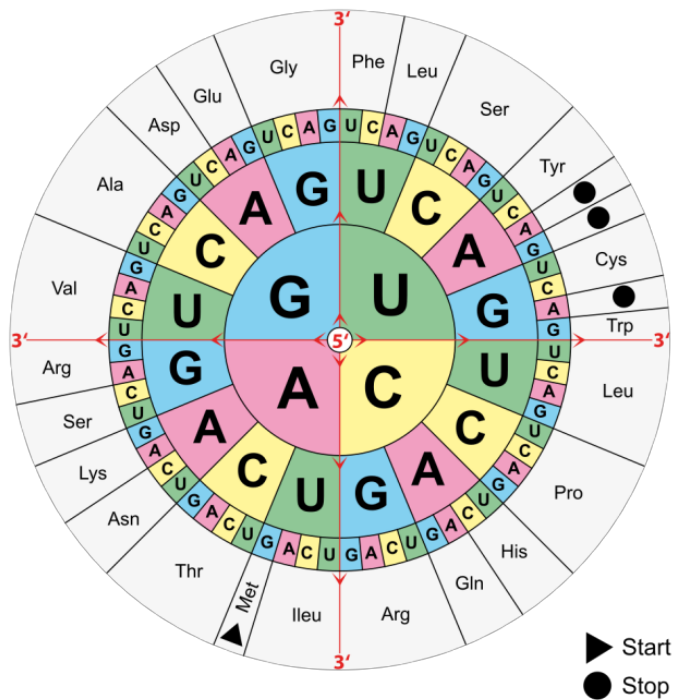
Source: <http://hyperphysics.phy-astr.gsu.edu/hbase/organic/imgorg/translation2.gif>

Translation



Source: <http://hyperphysics.phy-astr.gsu.edu/hbase/organic/imgorg/translation2.gif>

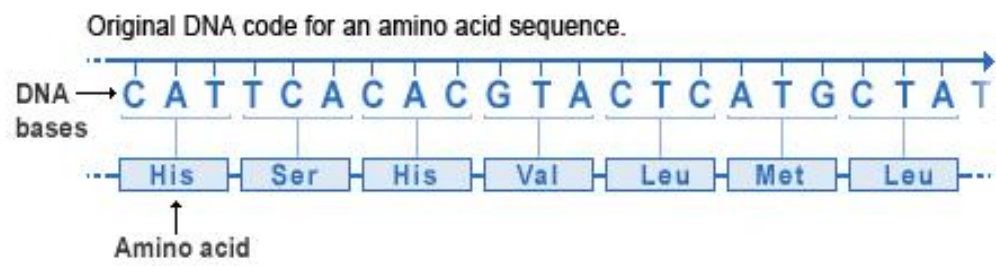
CODON WHEEL



Amino Acids

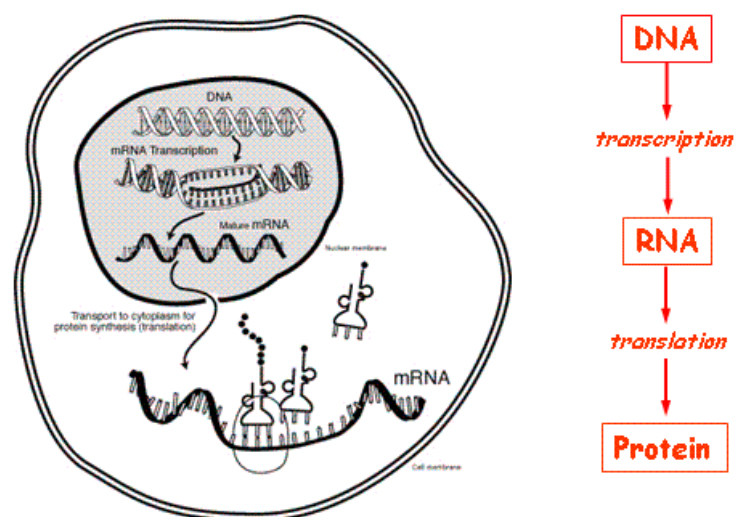
Amino Acid	3-Letter Code	1-Letter Code
Alanine	Ala	A
Cysteine	Cys	C
Aspartic acid or aspartate	Asp	D
Glutamic acid or glutamate	Glu	E
Phenylalanine	Phe	F
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Lysine	Lys	K
Leucine	Leu	L
Methionine	Met	M
Asparagine	Asn	N
Proline	Pro	P
Glutamine	Gln	Q
Arginine	Arg	R
Serine	Ser	S
Threonine	Thr	T
Valine	Val	V
Tryptophan	Trp	W
Tyrosine	Tyr	Y

Protein Sequence



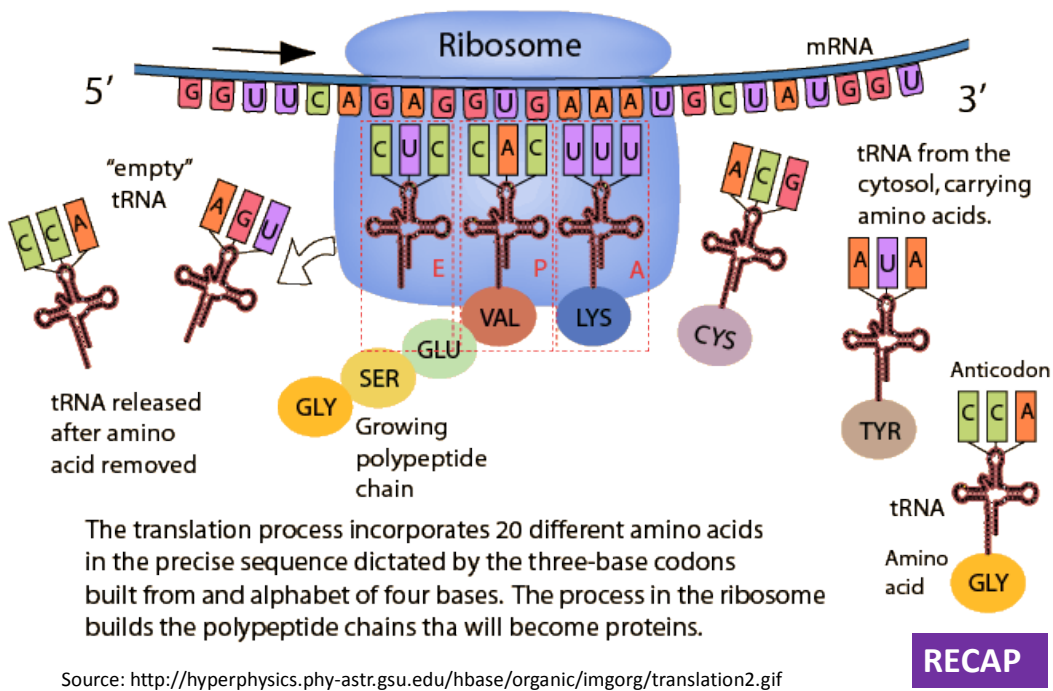
Lecture 02

Central Dogma of Molecular Biology

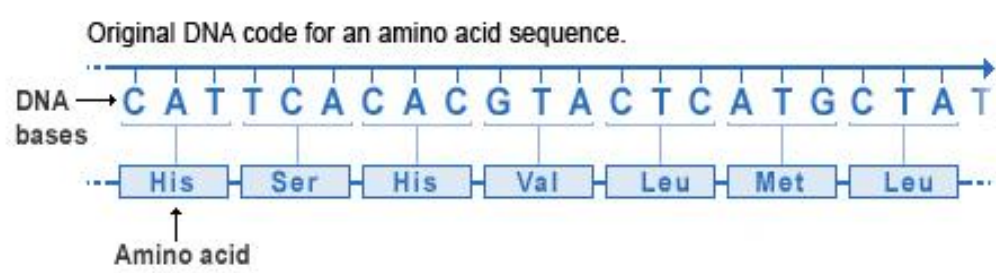
**RECAP**

Source: http://www.ncbi.nlm.nih.gov/Class/MCACourse/Modules/MolBioReview/central_dogma.html

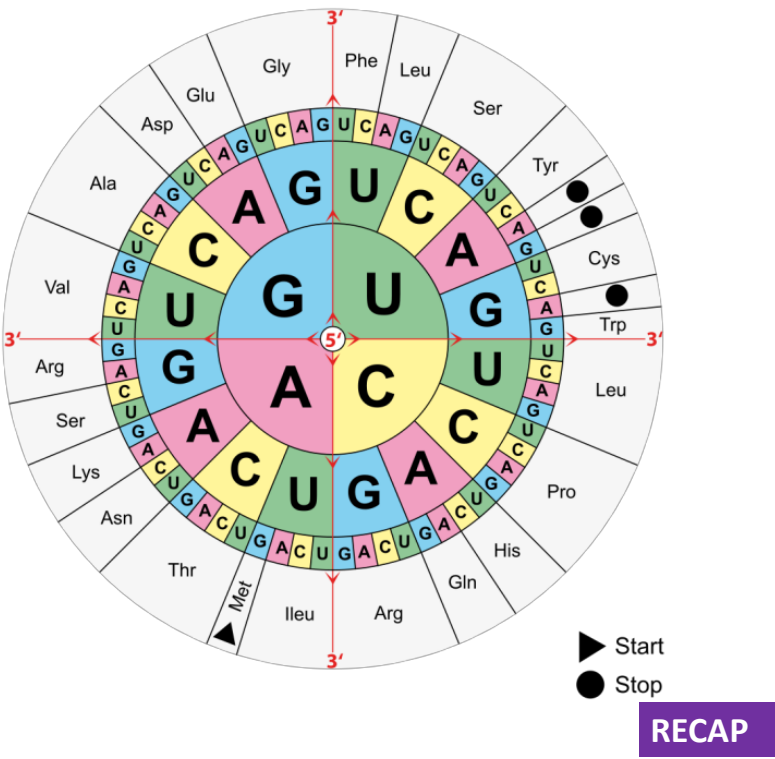
Translation



Protein Sequence



CODON
WHEEL

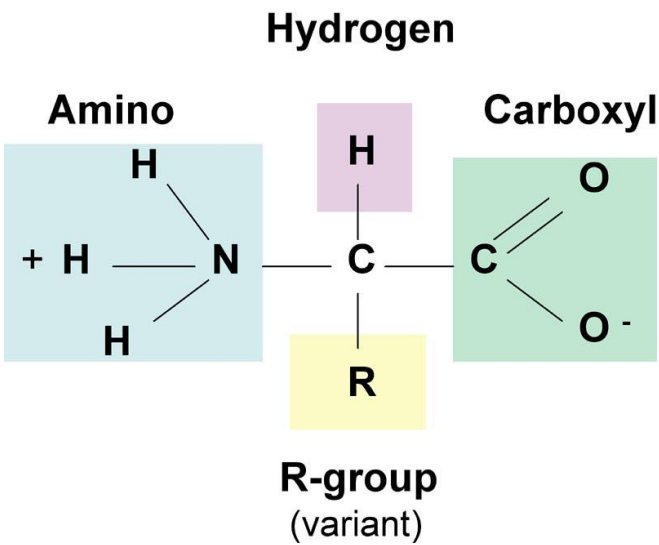


Amino Acids

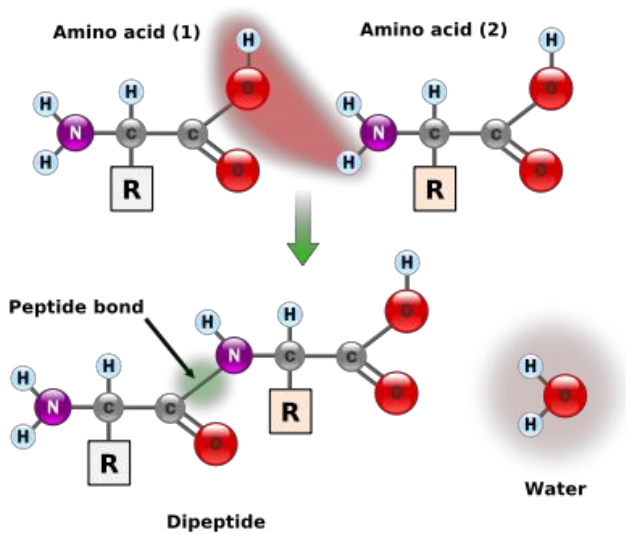
Amino Acid	3-Letter Code	1-Letter Code
Alanine	Ala	A
Cysteine	Cys	C
Aspartic acid or aspartate	Asp	D
Glutamic acid or glutamate	Glu	E
Phenylalanine	Phe	F
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Lysine	Lys	K
Leucine	Leu	L
Methionine	Met	M
Asparagine	Asn	N
Proline	Pro	P
Glutamine	Gln	Q
Arginine	Arg	R
Serine	Ser	S
Threonine	Thr	T
Valine	Val	V
Tryptophan	Trp	W
Tyrosine	Tyr	Y

RECAP

Amino Acid Structure

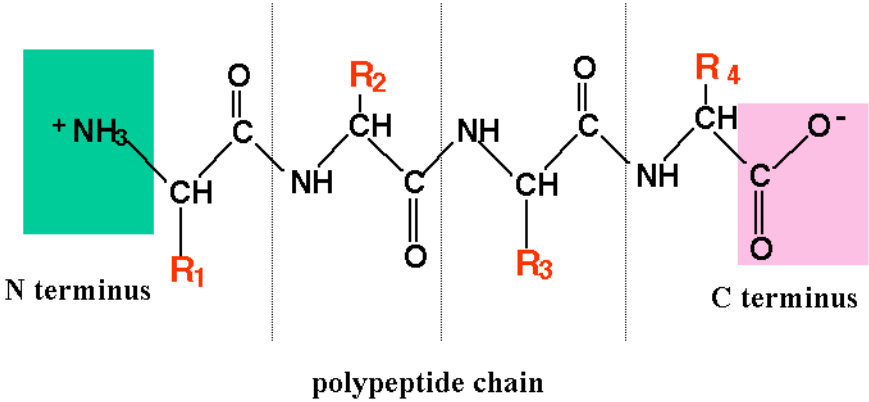


Peptide Bond Formation



Peptide

Peptide = chain of amino acids



Essential Amino Acids

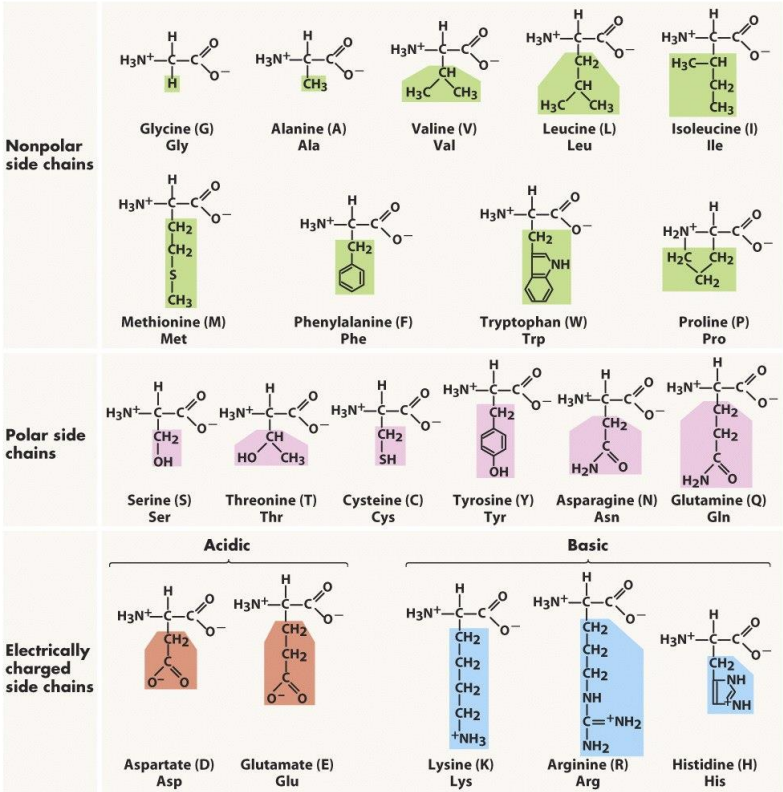
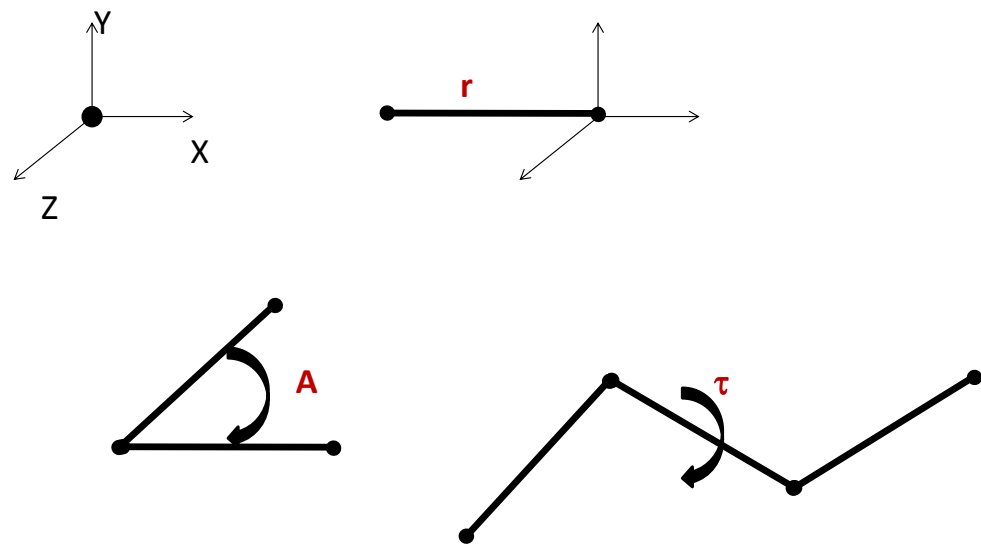


Figure 3-5 Biological Science, 2/e © 2005 Pearson Prentice Hall, Inc.

Point, Line, Angle, Torsional Angle



Bonds, Angles and Dihedral Angles

N terminus

C terminus

N	11.751	37.846	29.016
CA	12.501	39.048	28.539
C	13.740	38.628	27.754
N	14.235	39.531	26.906
CA	15.552	39.410	26.282
C	16.616	38.913	27.263
N	16.789	39.630	28.369
CA	17.791	39.281	29.375
C	17.598	37.844	29.863
N	16.368	37.519	30.261
CA	16.004	36.186	30.742
C	16.371	35.097	29.741

Homework

Write down a pseudo-code to parse the coordinate file and output the torsional angles.

Essential Amino Acids

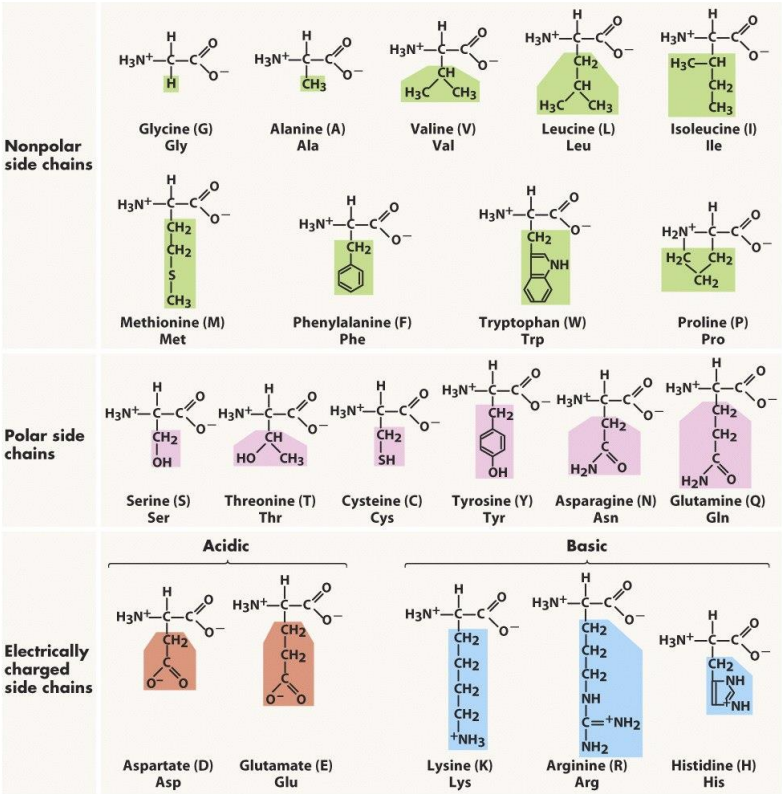


Figure 3-5 Biological Science, 2/e © 2005 Pearson Prentice Hall, Inc.