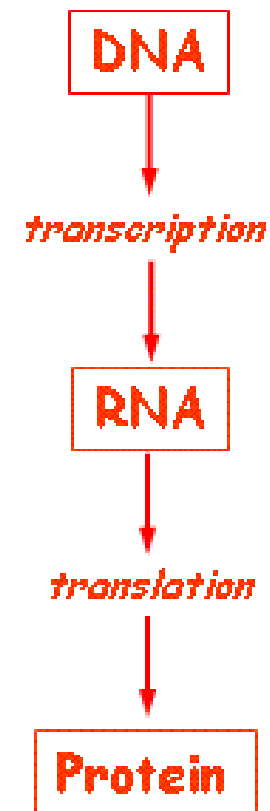
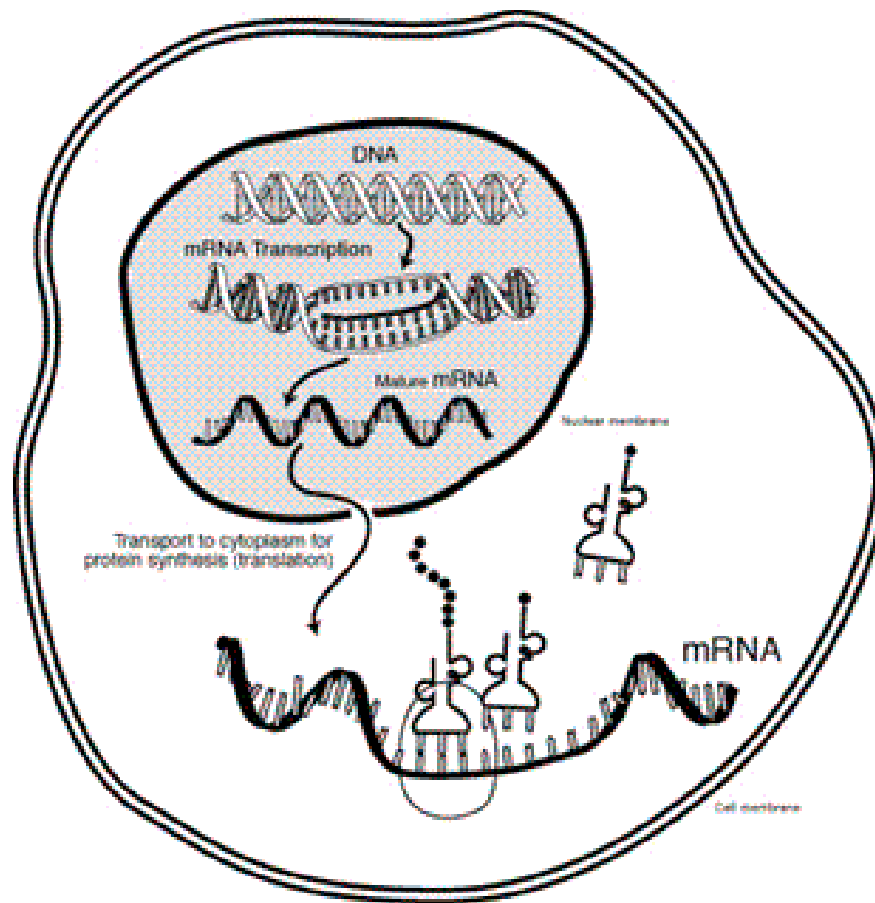


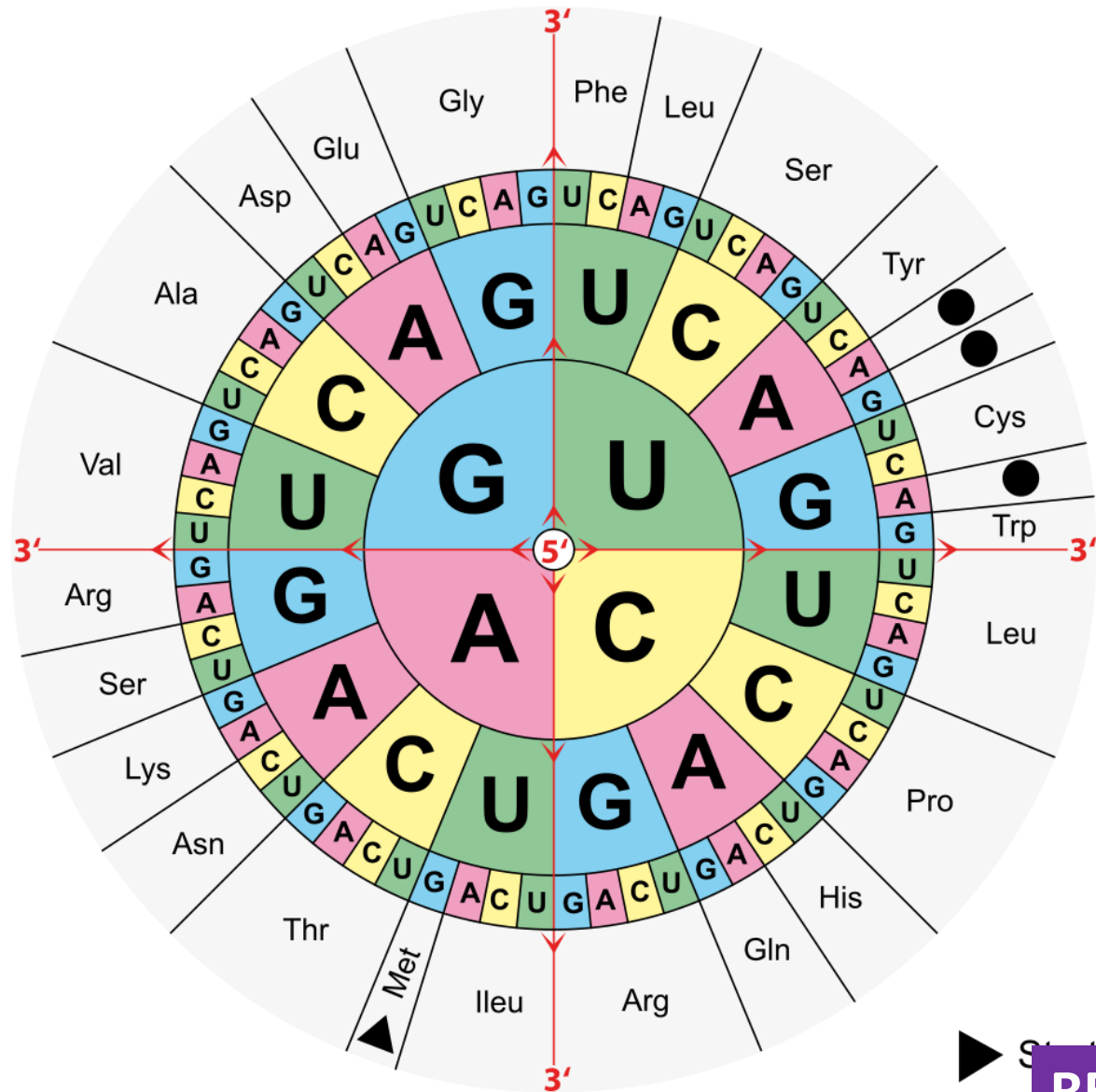
# Lecture 05

# Central Dogma of Molecular Biology



RECAP

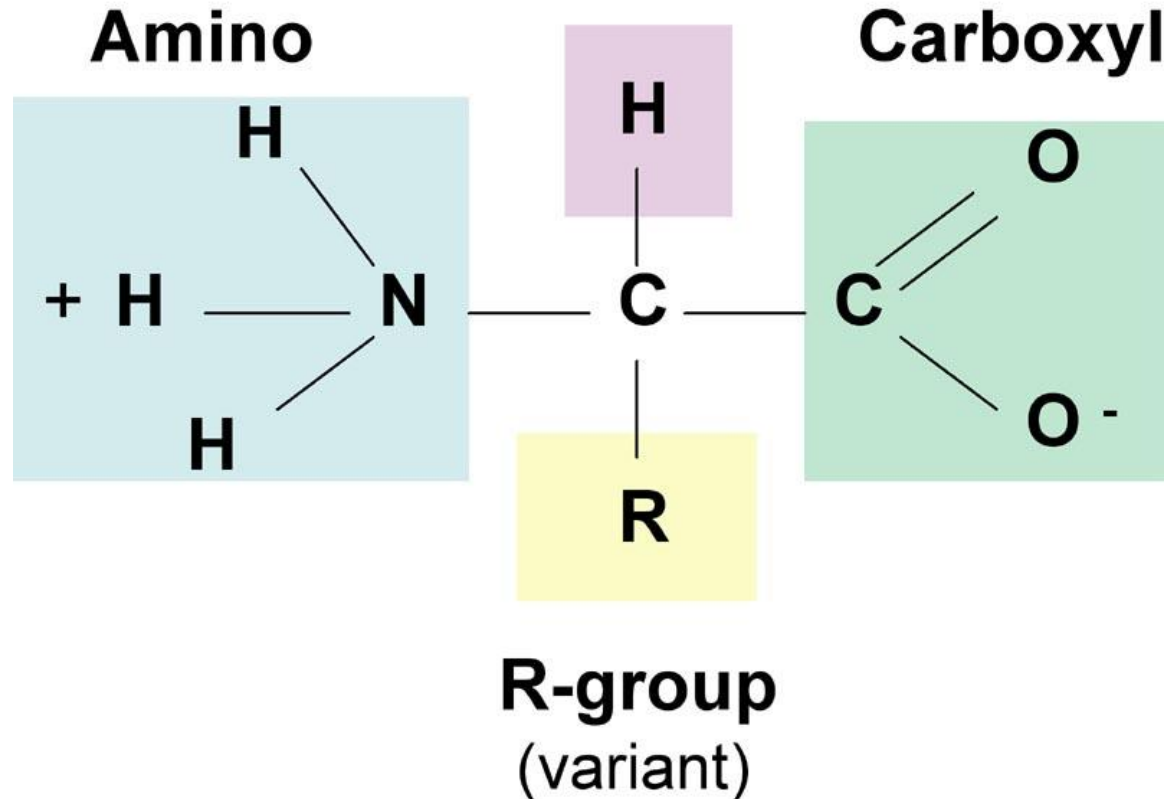
# CODON WHEEL



## RECAP

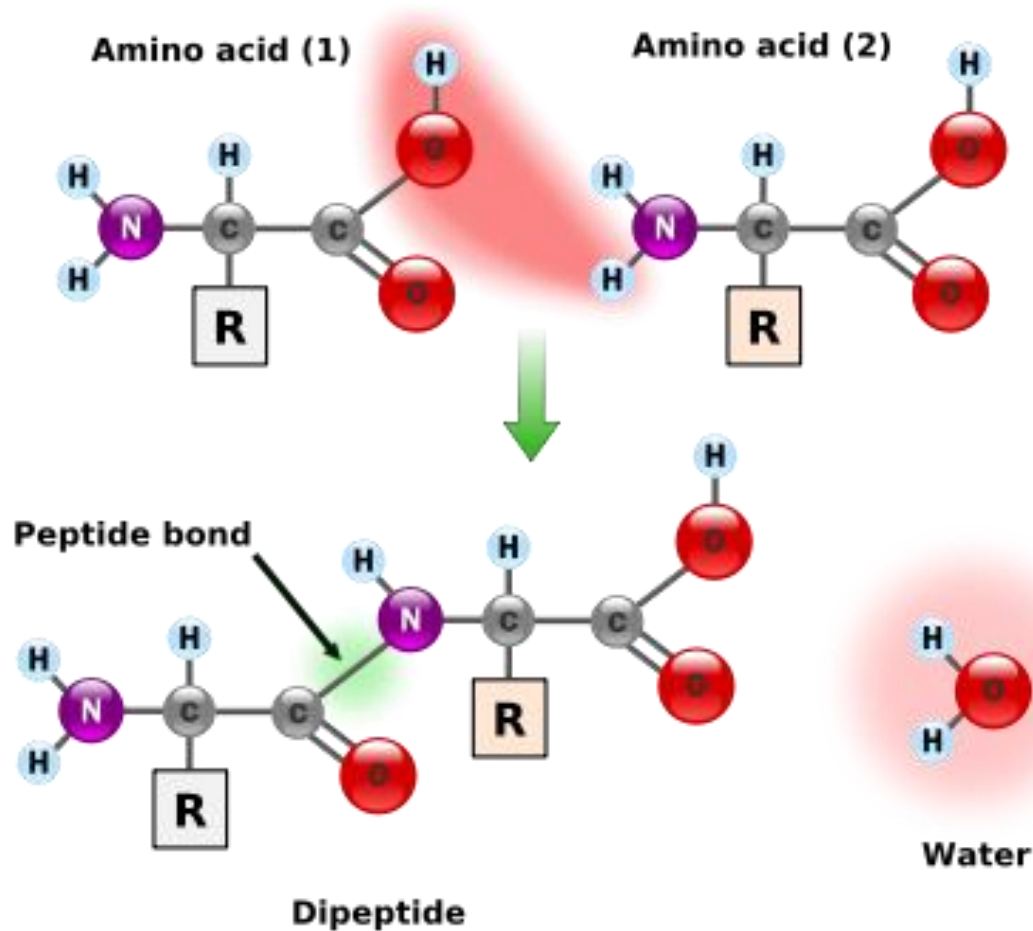
# Amino Acid Structure

Hydrogen



RECAP

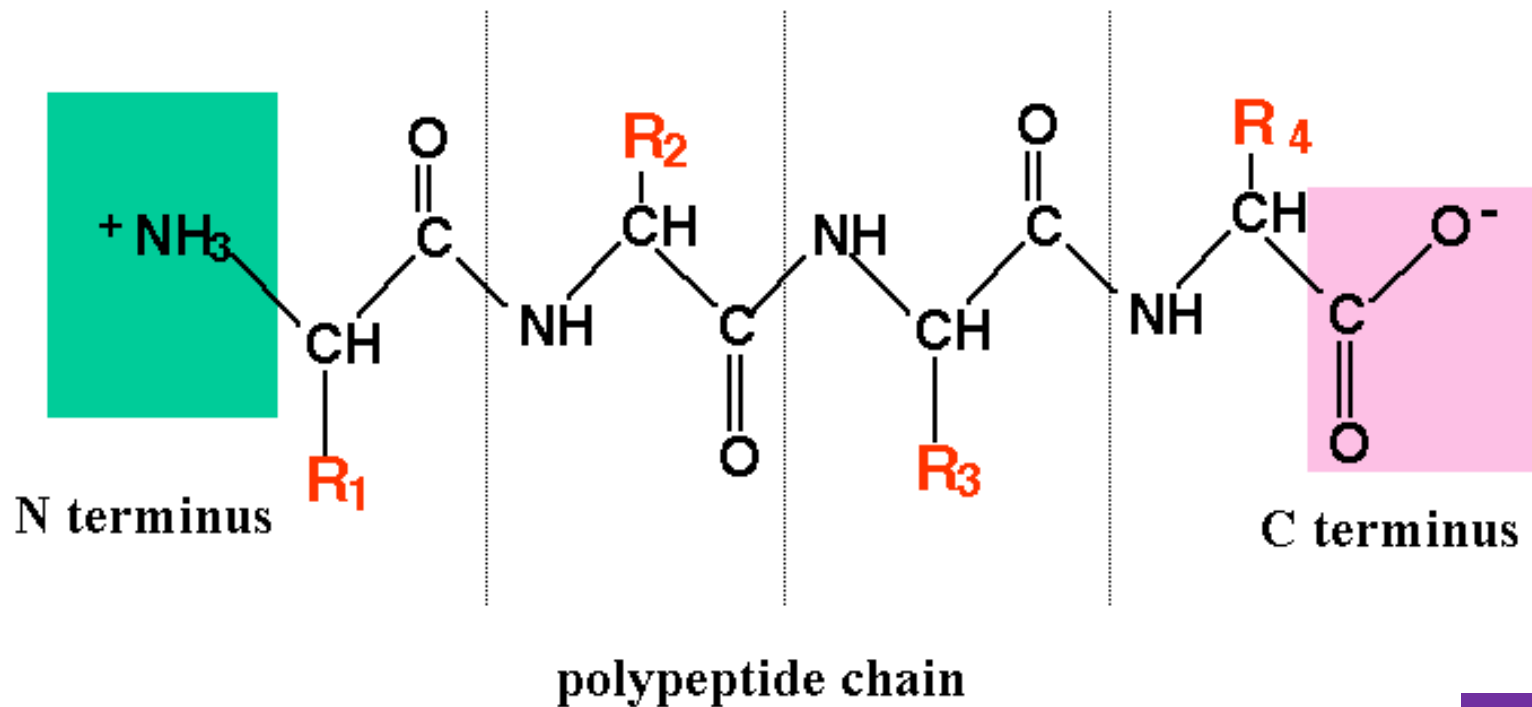
# Peptide Bond Formation



RECAP

# Peptide

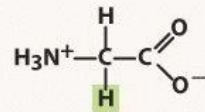
Peptide = chain of amino acids



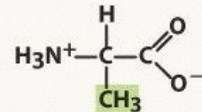
RECAP

# Essential Amino Acids

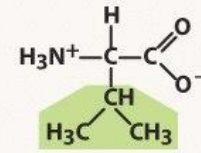
## Nonpolar side chains



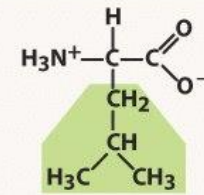
Glycine (G)  
Gly



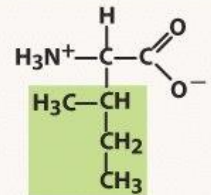
Alanine (A)  
Ala



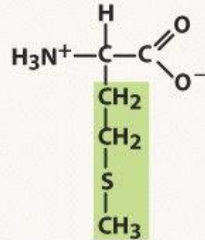
Valine (V)  
Val



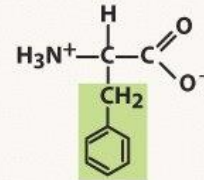
Leucine (L)  
Leu



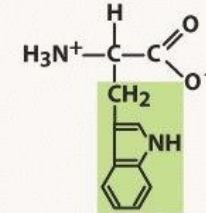
Isoleucine (I)  
Ile



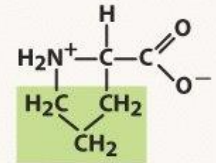
Methionine (M)  
Met



Phenylalanine (F)  
Phe

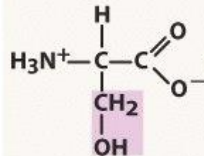


Tryptophan (W)  
Trp

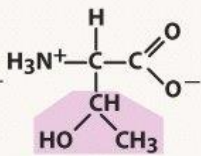


Proline (P)  
Pro

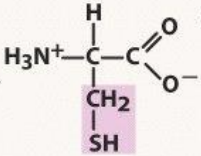
## Polar side chains



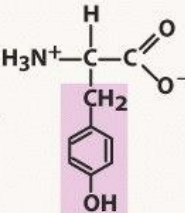
Serine (S)  
Ser



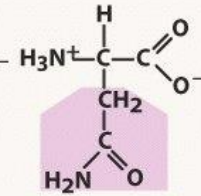
Threonine (T)  
Thr



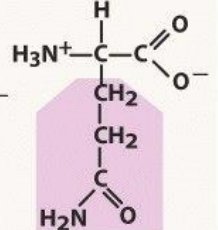
Cysteine (C)  
Cys



Tyrosine (Y)  
Tyr



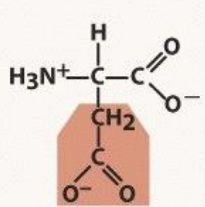
Asparagine (N)  
Asn



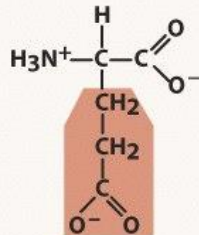
Glutamine (Q)  
Gln

## Electrically charged side chains

### Acidic

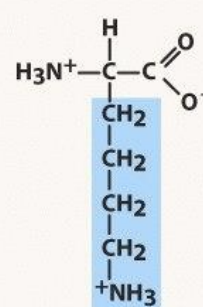


Aspartate (D)  
Asp

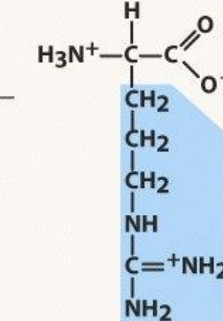


Glutamate (E)  
Glu

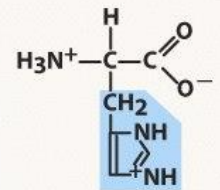
### Basic



Lysine (K)  
Lys



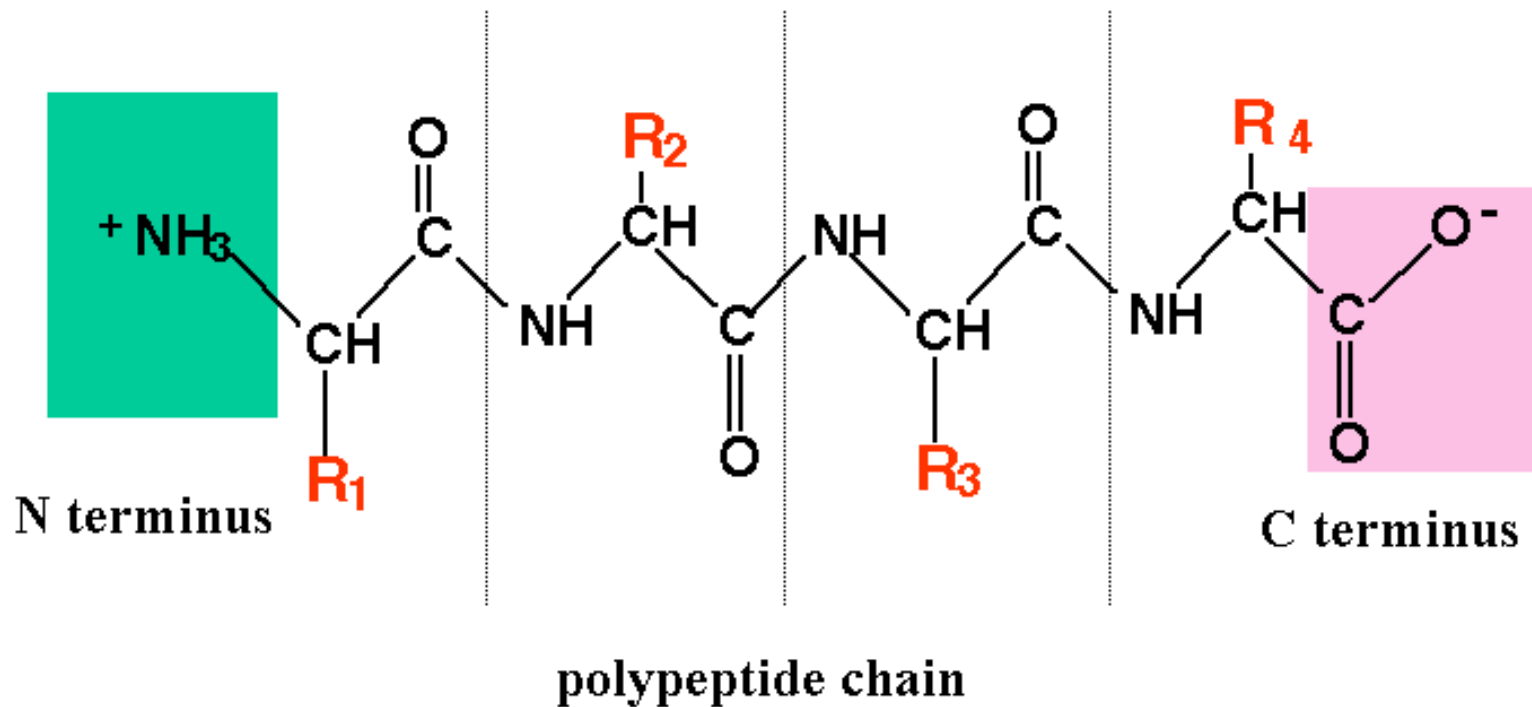
Arginine (R)  
Arg



Histidine (H)  
His

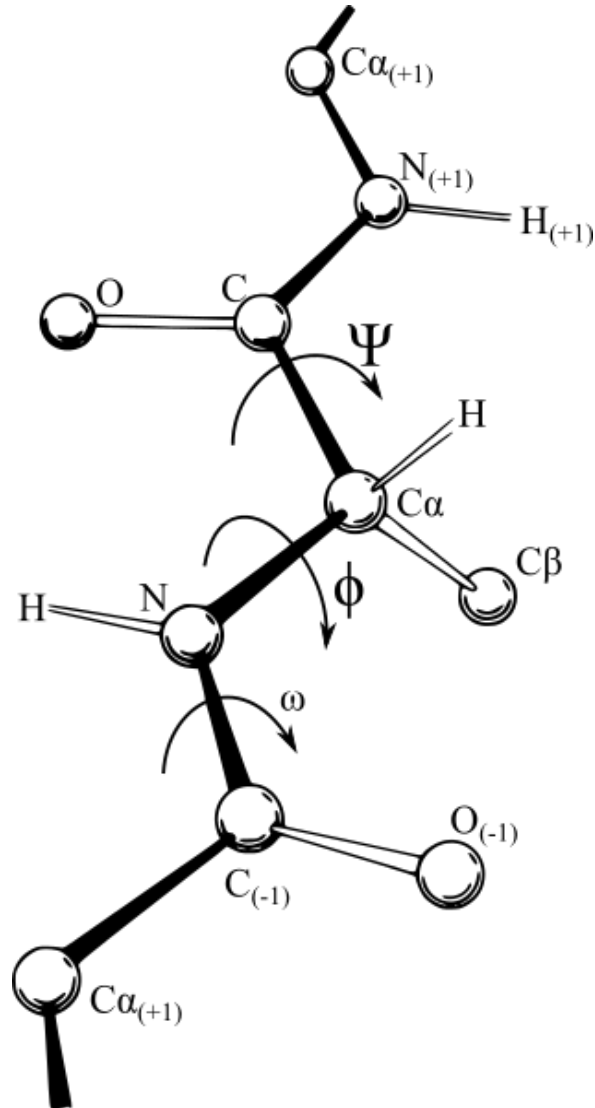
# Peptide

Peptide = chain of amino acids

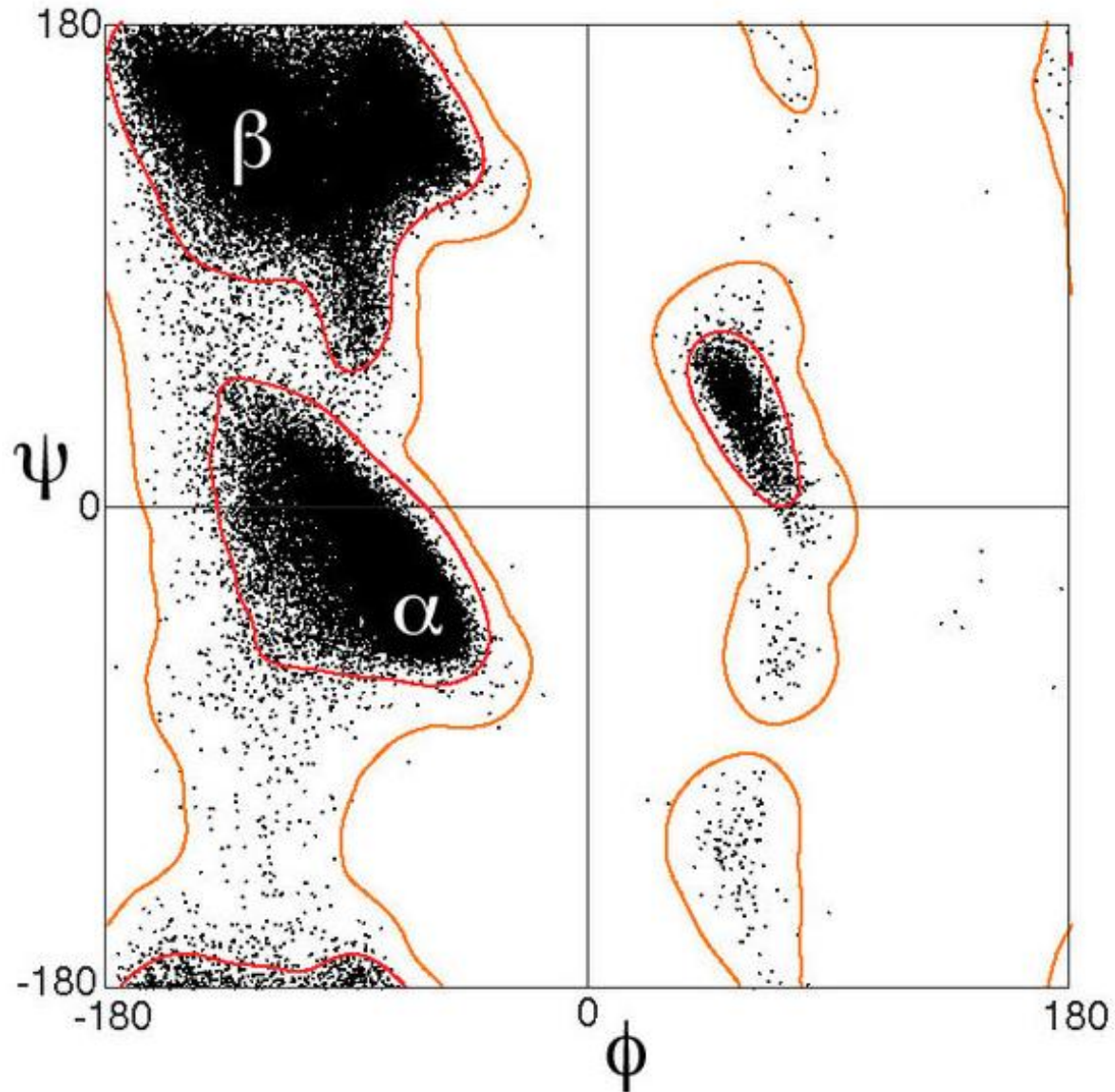




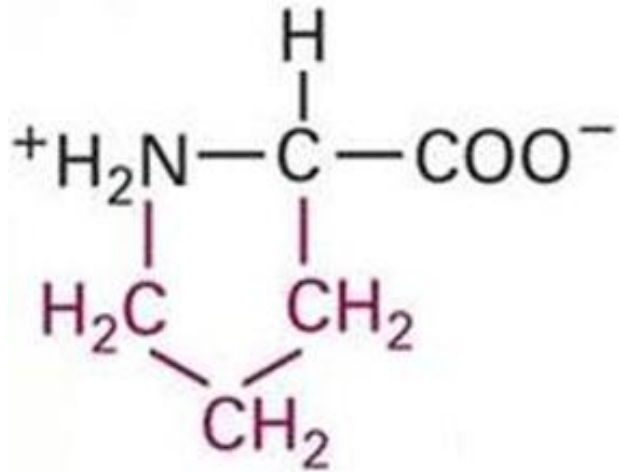
# Protein Backbone



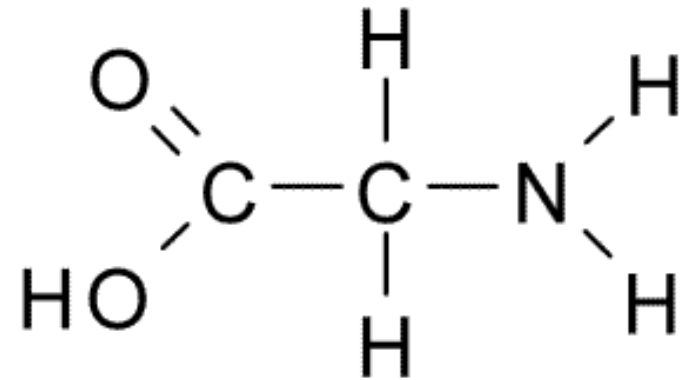
# Ramachandran Plot



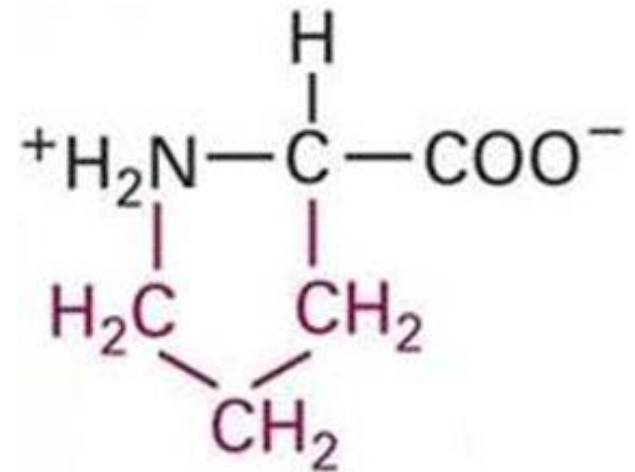
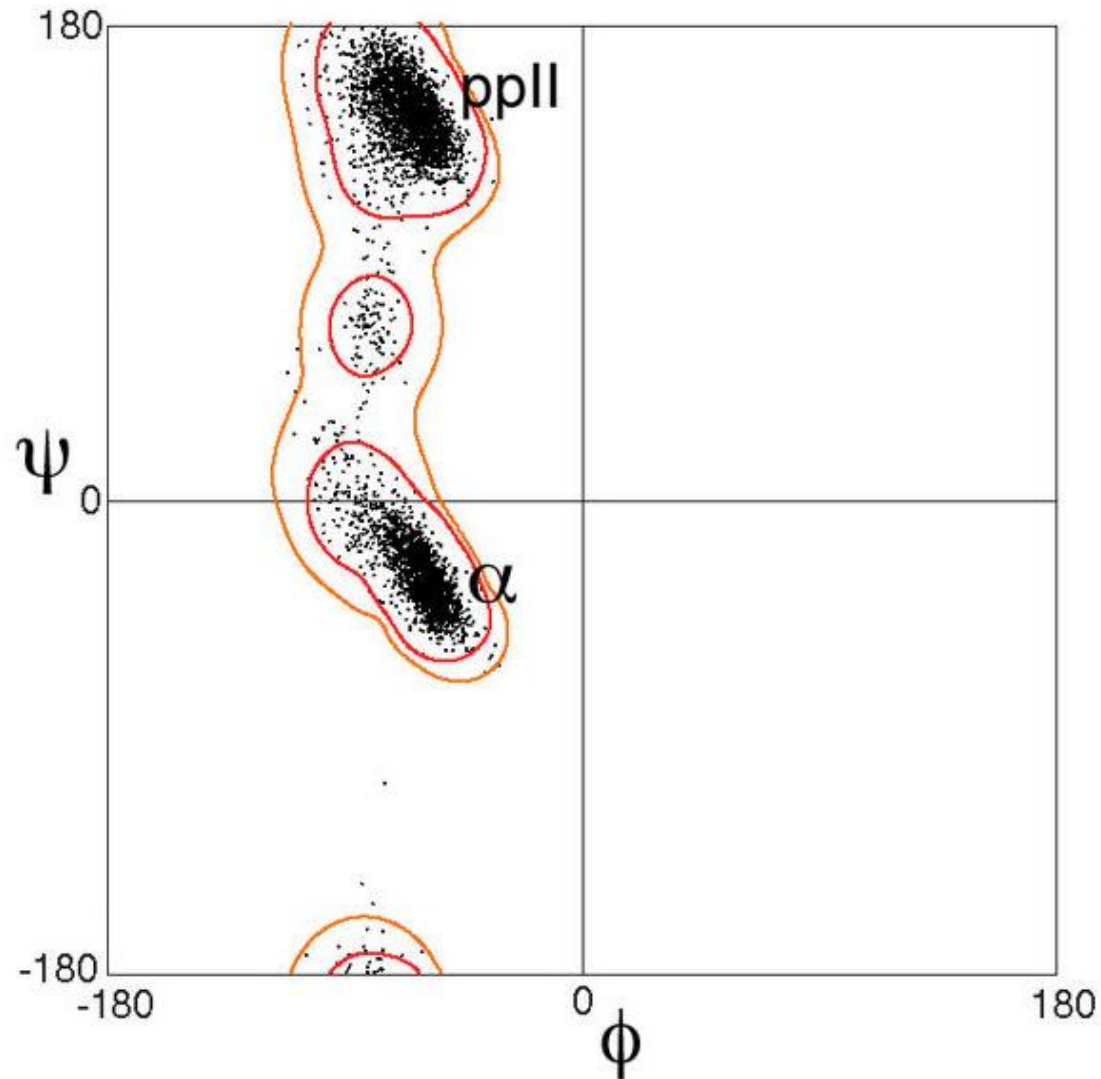
## Proline structure



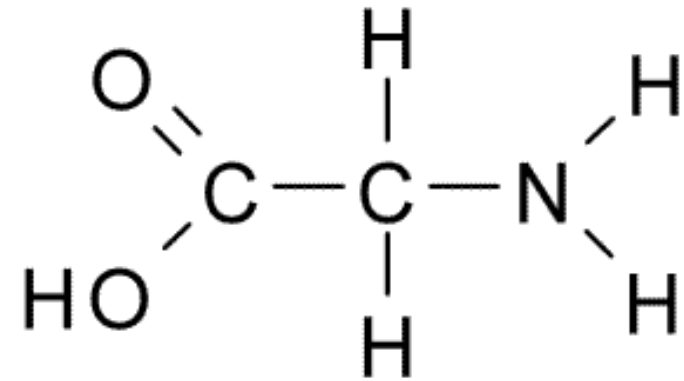
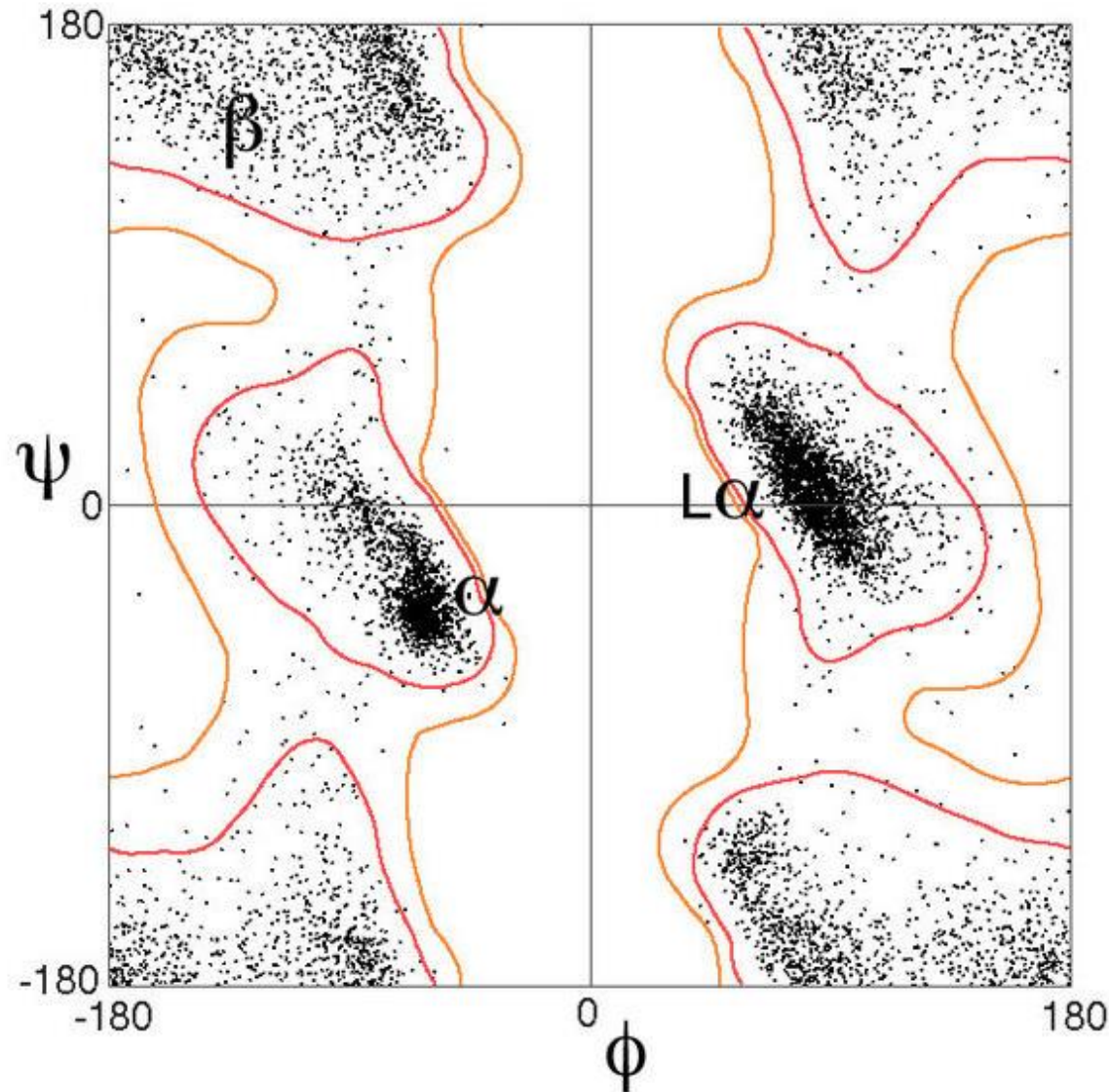
## Glycine structure



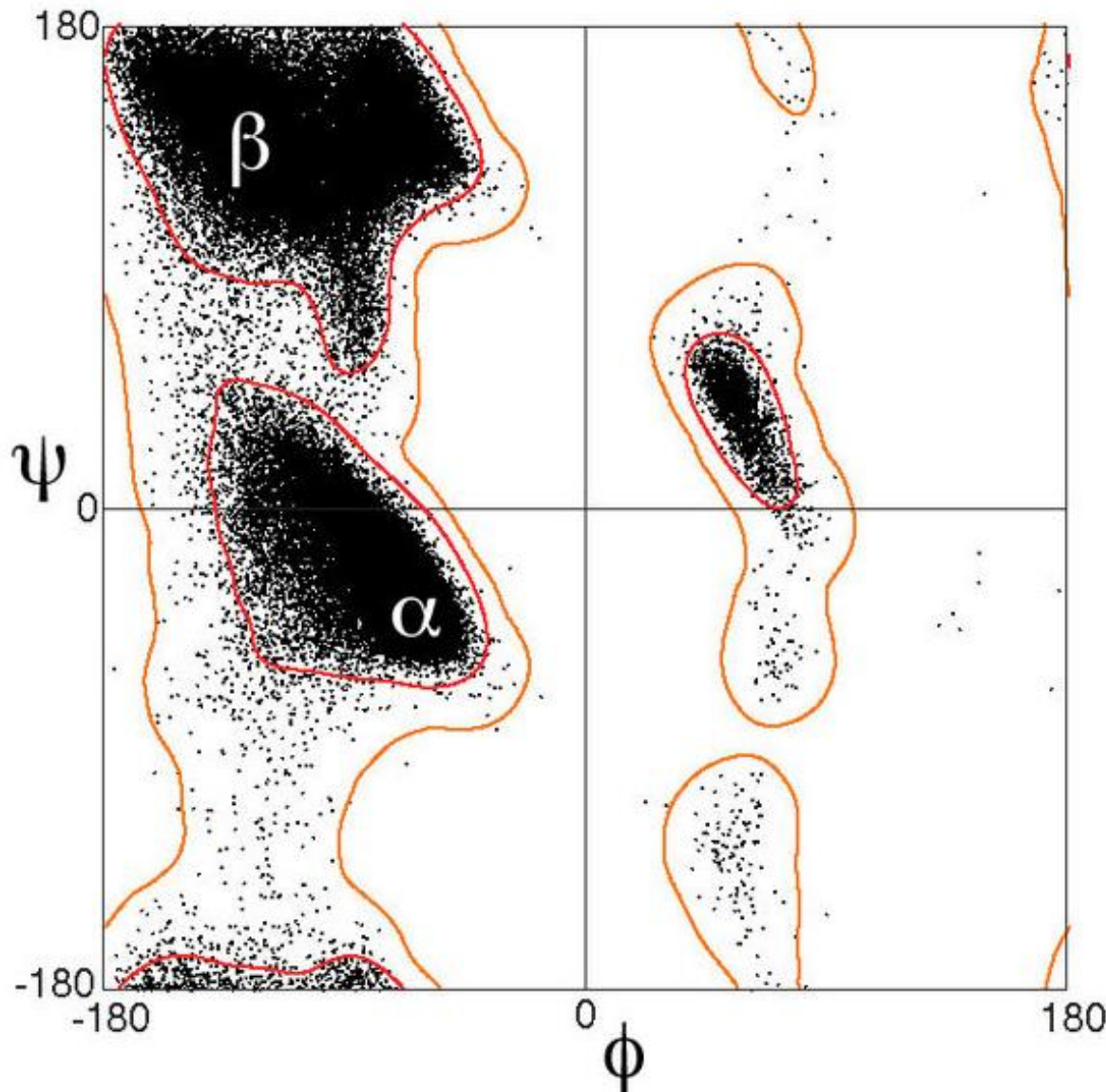
# Ramachandran Plot for Proline



# Ramachandran Plot for Glycine



# How to implement Ramachandran Plot?

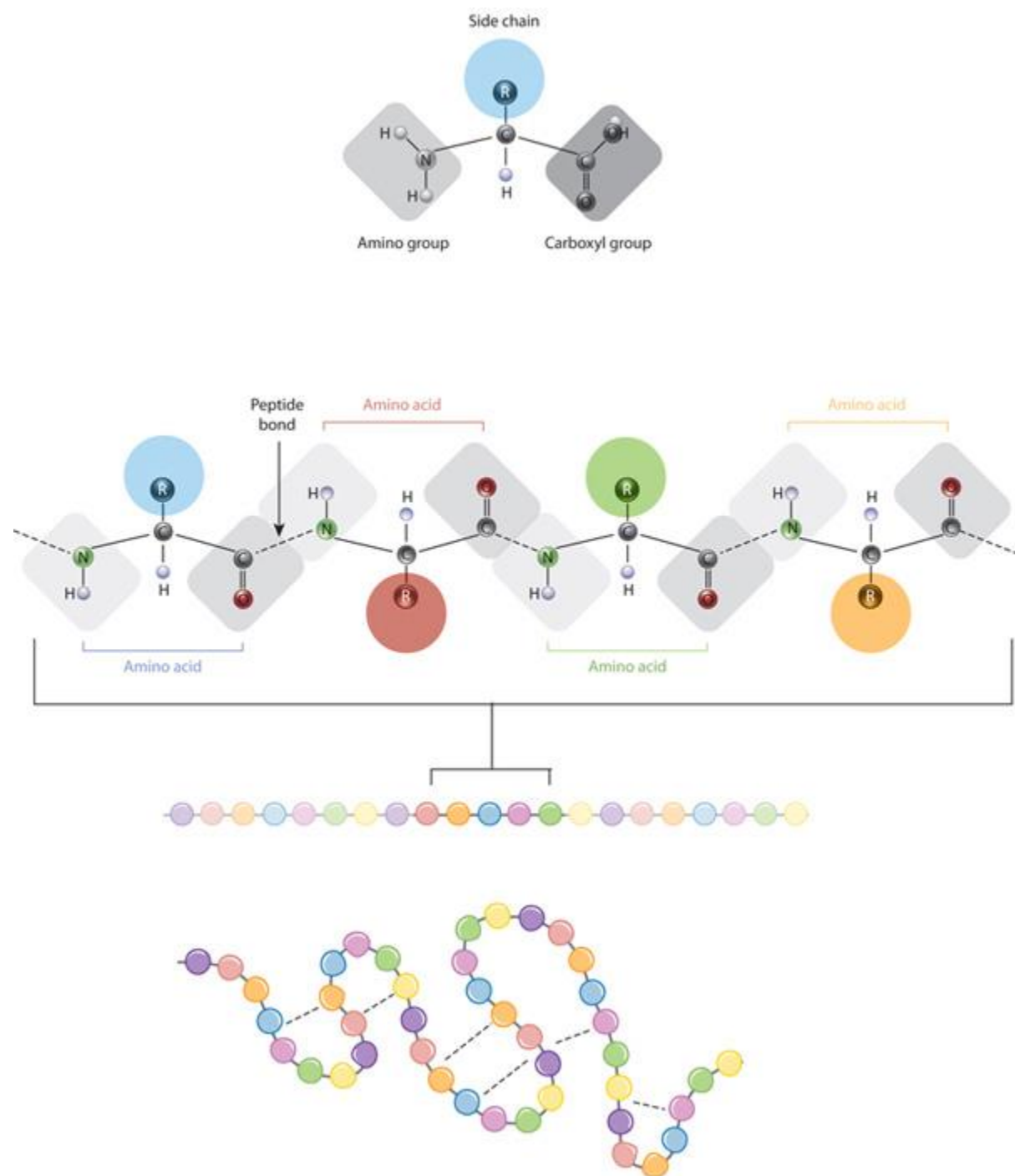


**Food for your  
thought**

Come up with an  
efficient data structure  
to implement for  
frequent use of  
Ramachandran Plot.

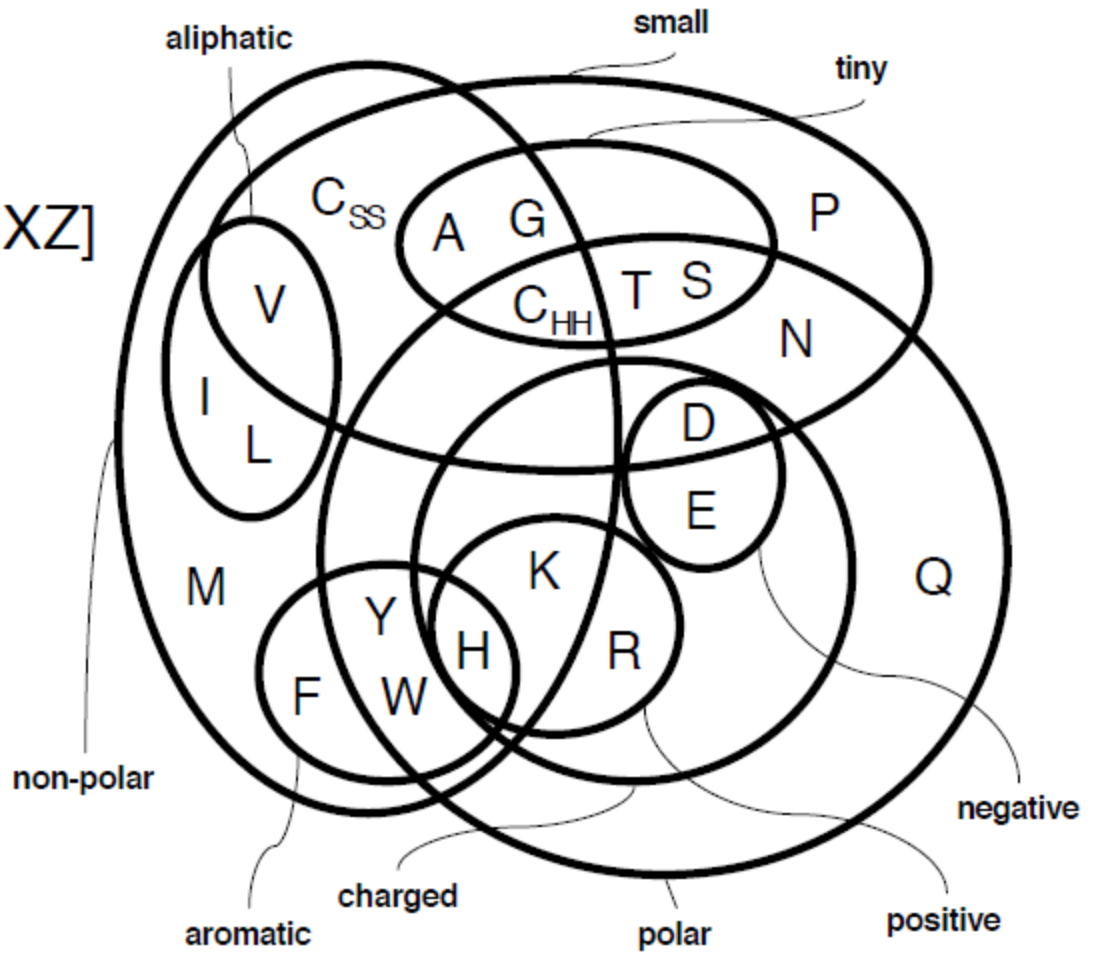


# Peptide

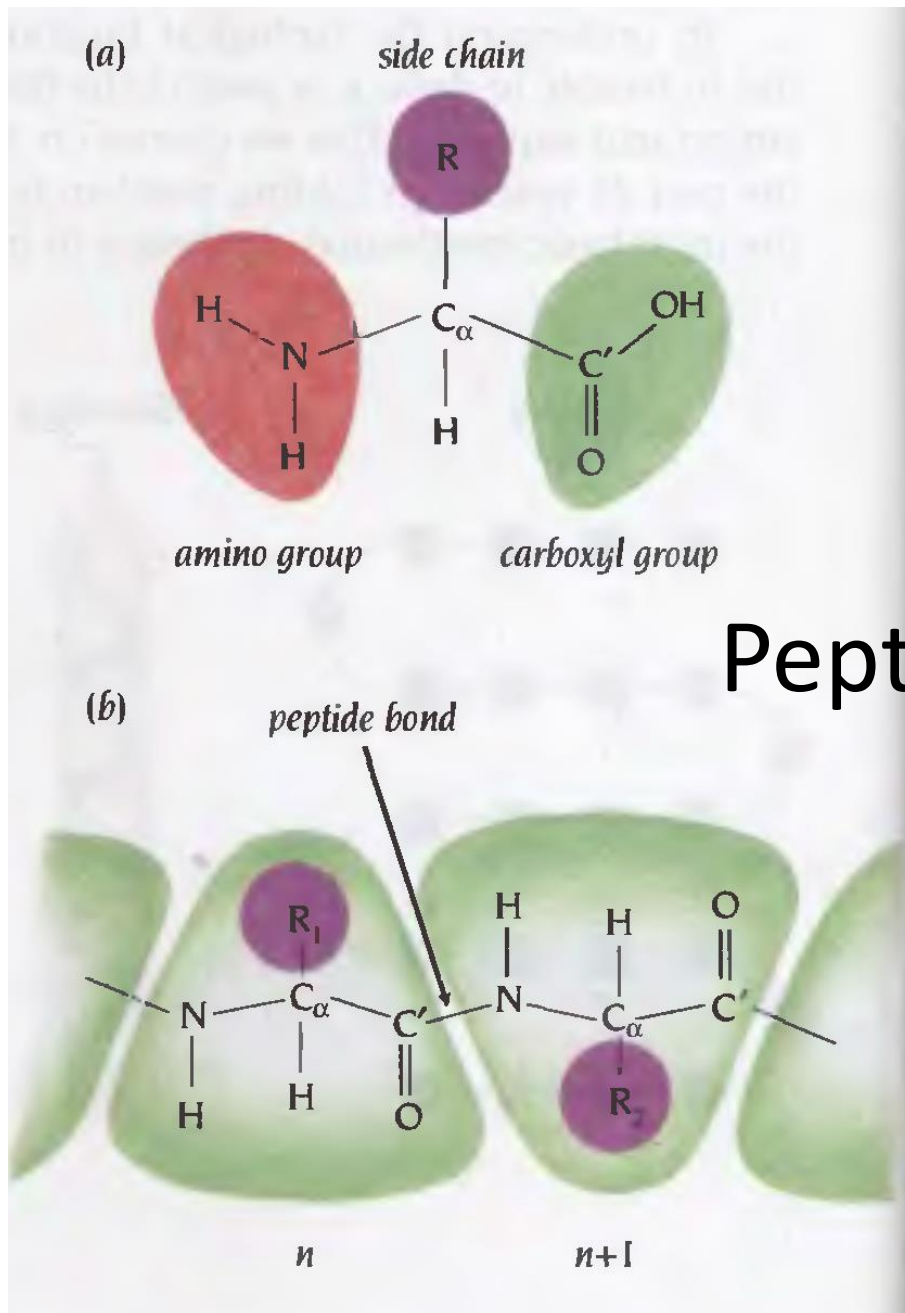


# Proteins

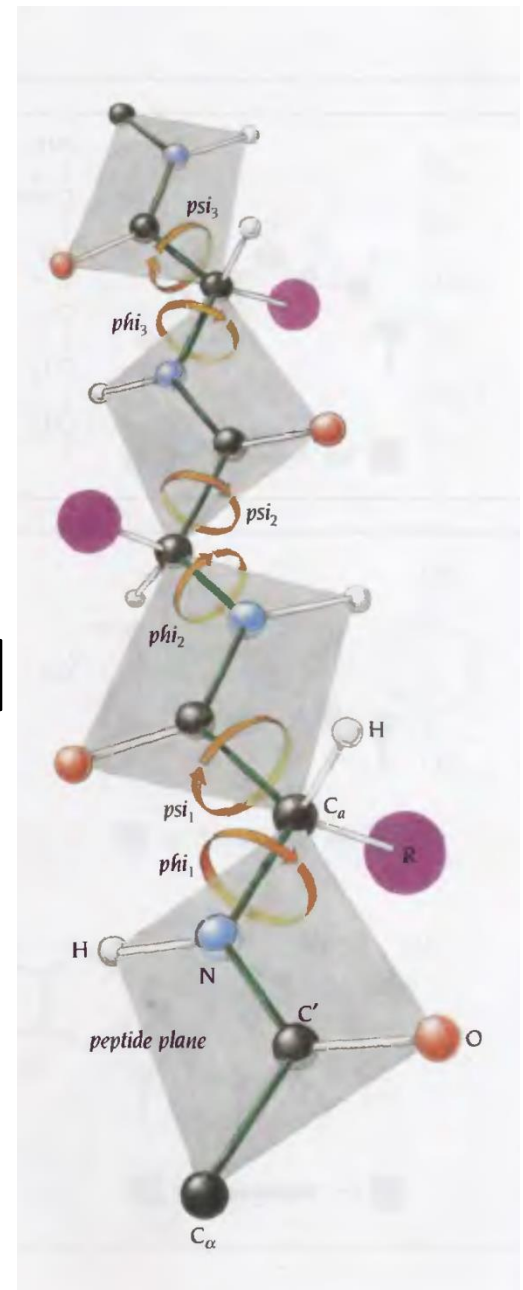
- ❑ 20 amino acids:  
[A-Z] but not [BJOUXZ]
- ❑ ~300 aa/protein  
(bacteria), ~200  
aa/domain
- ❑ ~200 k known  
sequences



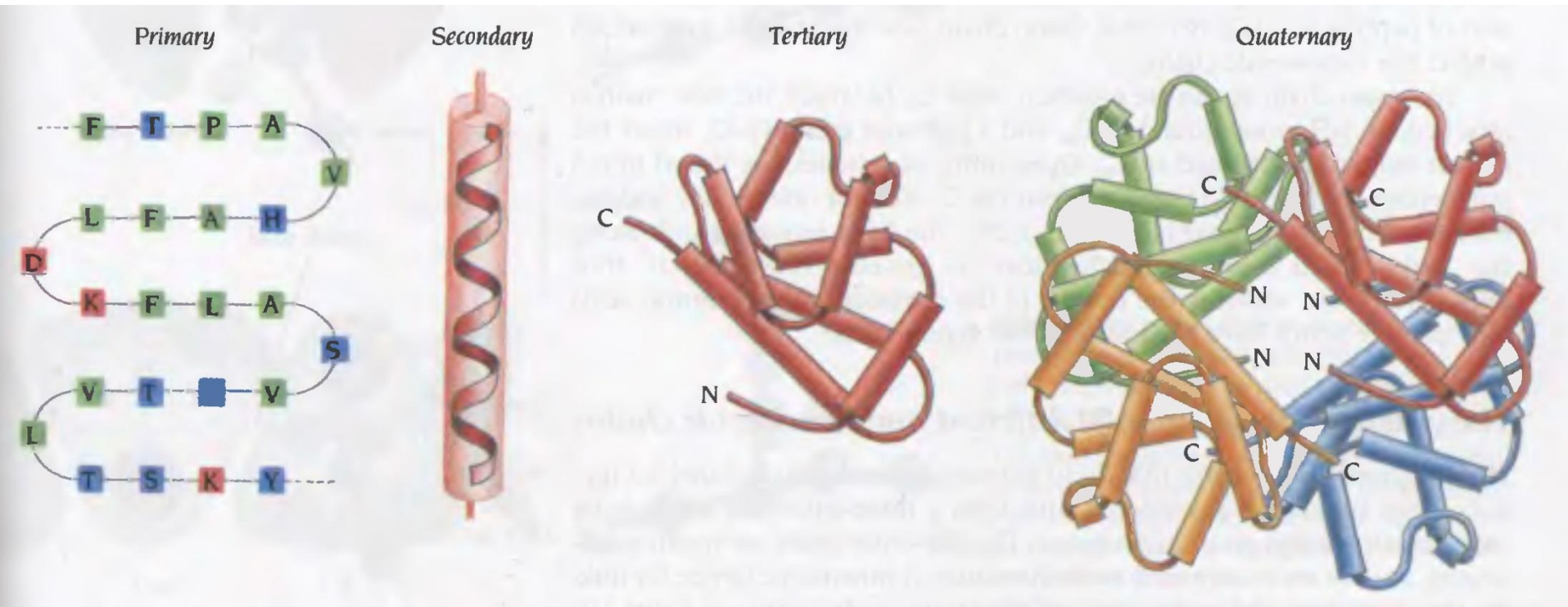




# Peptide bond



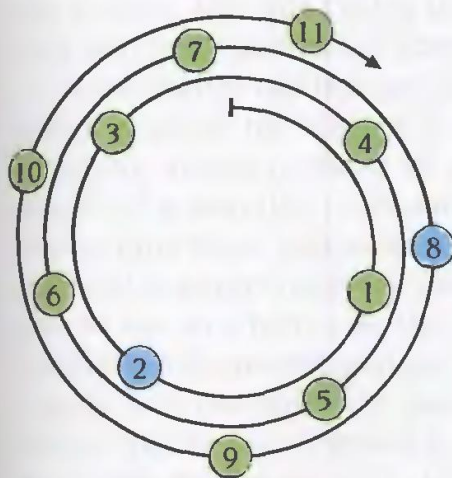
# Proteins



# Preference of AA in helix

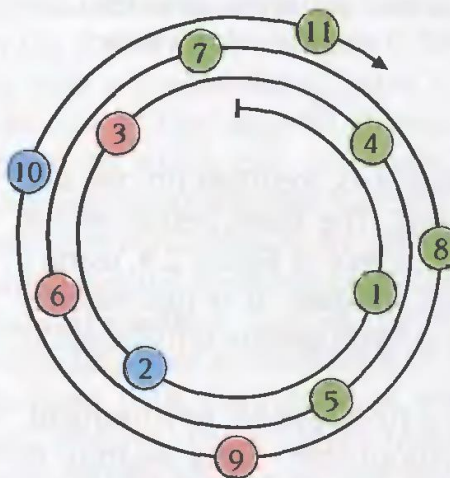
**Table 2.1** Amino acid sequences of three  $\alpha$  helices

1. - Leu - Ser - Phe - Ala - Ala - Ala - Met - Asn - Gly - Leu - Ala -
2. - Ile - Asn - Glu - Gly - Phe - Asp - Leu - Leu - Arg - Ser - Gly -
3. - Lys - Glu - Asp - Ala - Lys - Gly - Lys - Ser - Glu - Glu - Glu -



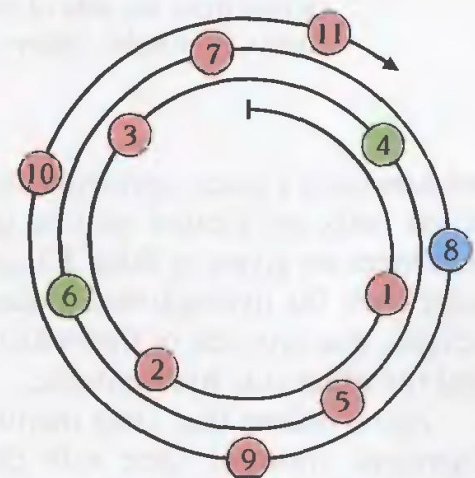
*citrate synthase*

1	2	3	4	5	6	7	8	9	10	11
L	S	F	A	A	A	M	N	G	L	A



*alcohol dehydrogenase*

1	2	3	4	5	6	7	8	9	10	11
I	N	E	G	F	D	L	L	R	S	G

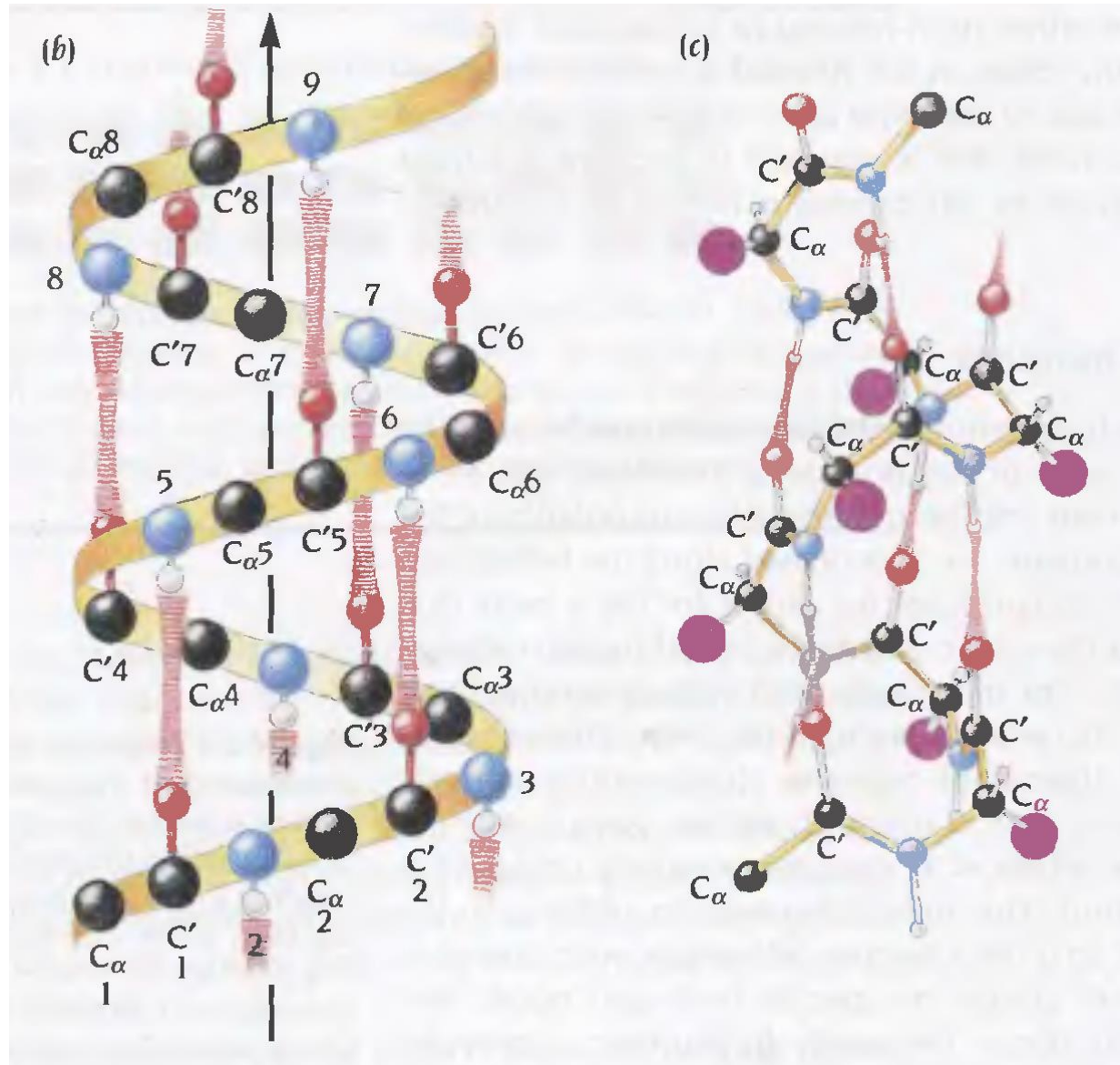


*troponin-C*

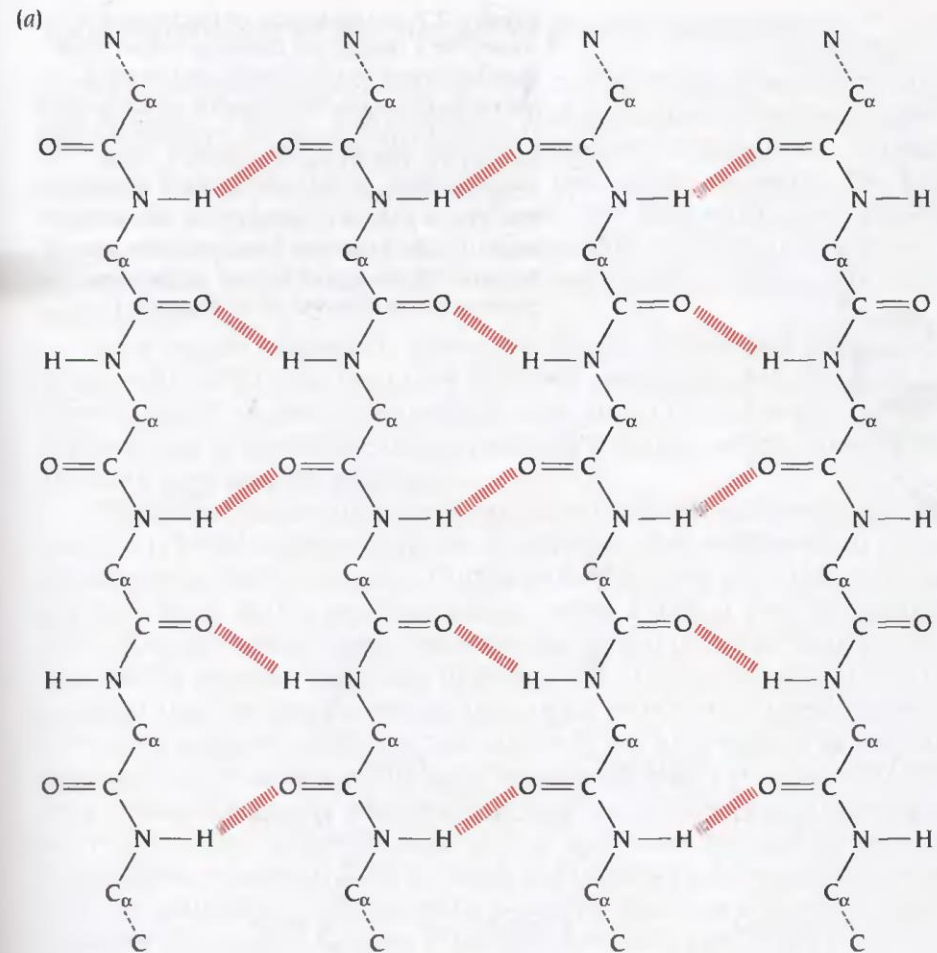
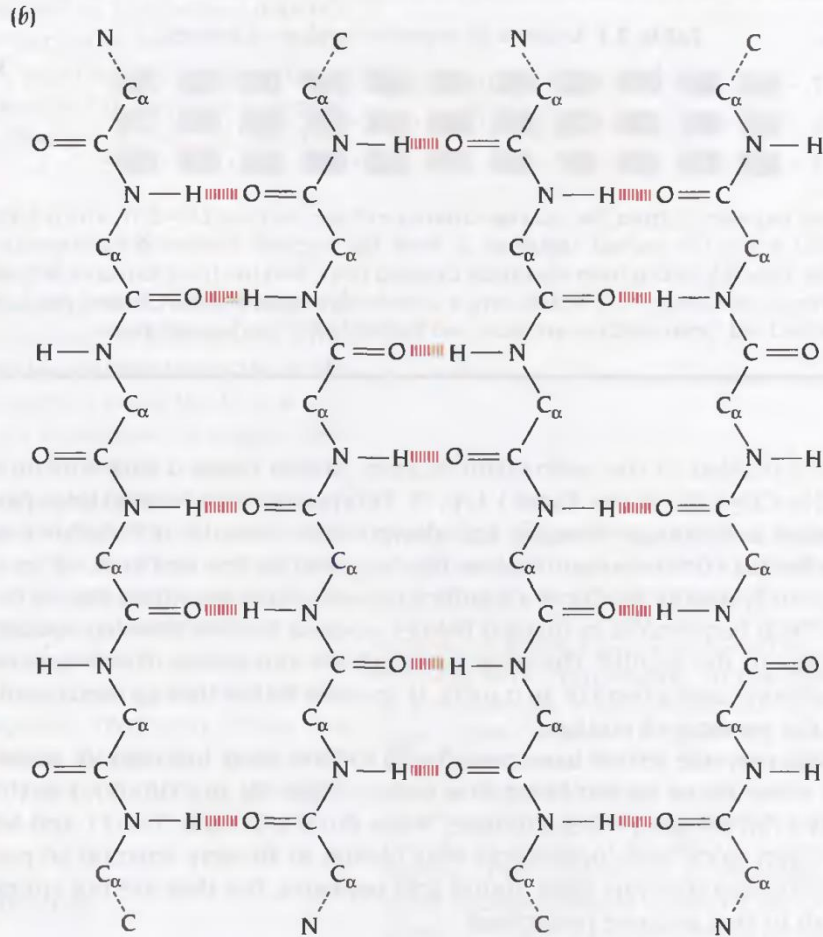
1	2	3	4	5	6	7	8	9	10	11
K	E	D	A	K	G	K	S	E	E	E



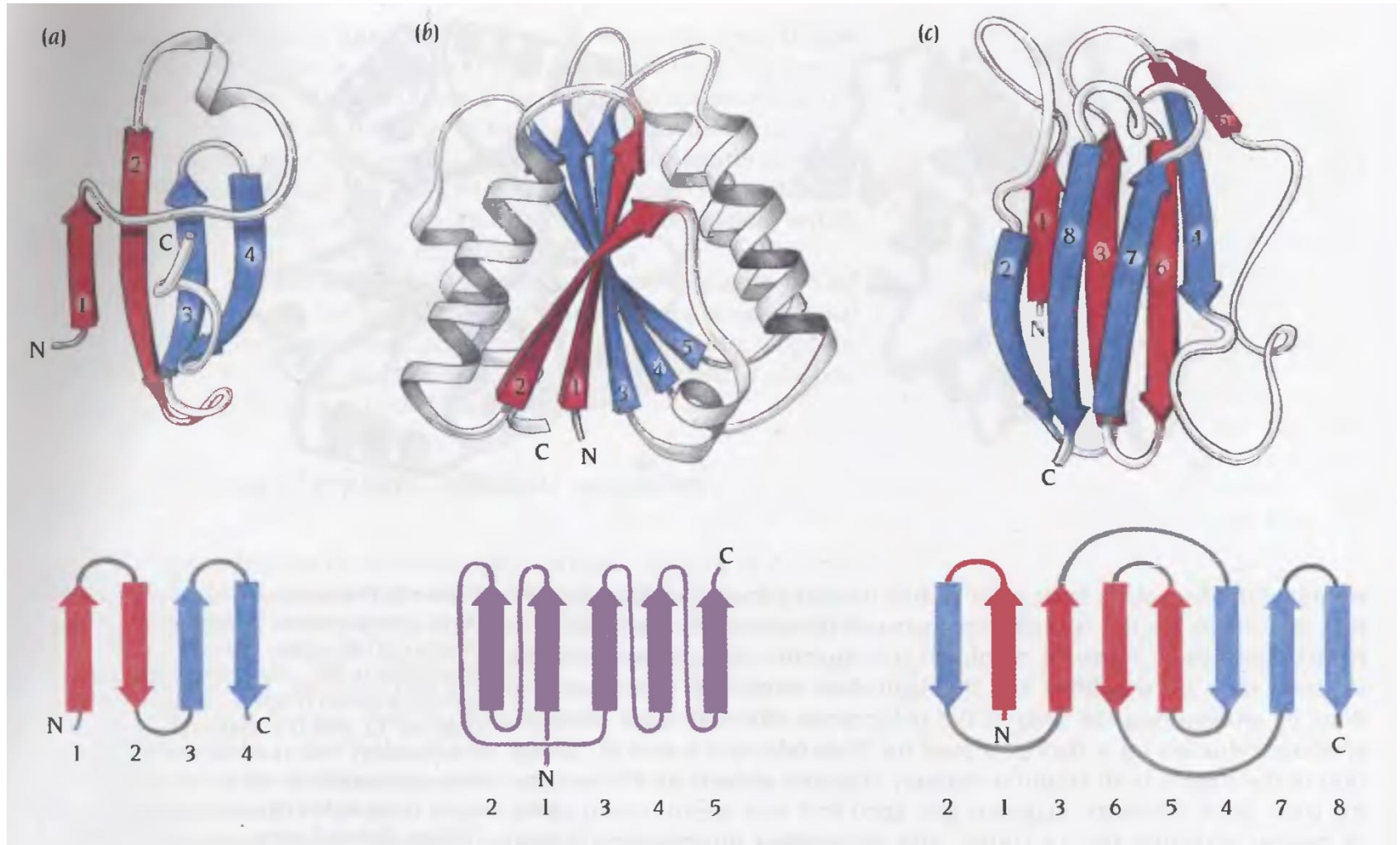
# Formation of helix



# Formation of sheet



# Topology in Structure



# Systems

- Deterministic System
  - In a deterministic system no randomness is involved in the development of future states of the system. A deterministic model will thus always produce the same output from a given starting condition or initial state.
- Dynamical System
  - A dynamical system is a concept in mathematics where a fixed rule describes the time dependence of a point in a geometrical space. Examples include the mathematical models that describe the swinging of a clock pendulum, the flow of water in a pipe etc.

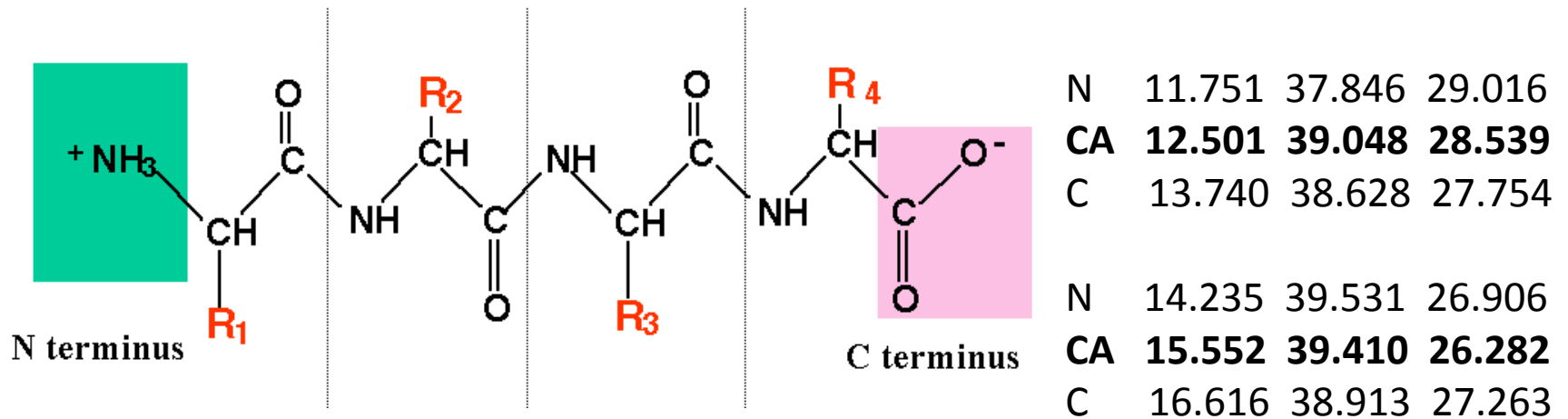
# Biological Parameter Space

- N-Dimensional
  - Temperature
  - Pressure
  - pH
  - Ionic Concentration
  - Solvent
  - Mode of interactions
  - Entropy
  - .....

Value of N varies from problem to problem.



# Bonds, Angles and Dihedral Angles



## Homework

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

N	16.789	39.630	28.369
<b>CA</b>	<b>17.791</b>	<b>39.281</b>	<b>29.375</b>
C	17.598	37.844	29.863
N	16.368	37.519	30.261
<b>CA</b>	<b>16.004</b>	<b>36.186</b>	<b>30.742</b>
C	16.371	35.097	29.741

# References

- <http://kinemage.biochem.duke.edu/teaching/anatax/>
- **Introduction to Protein Structure**  
Authors: Carl Branden, John Tooze