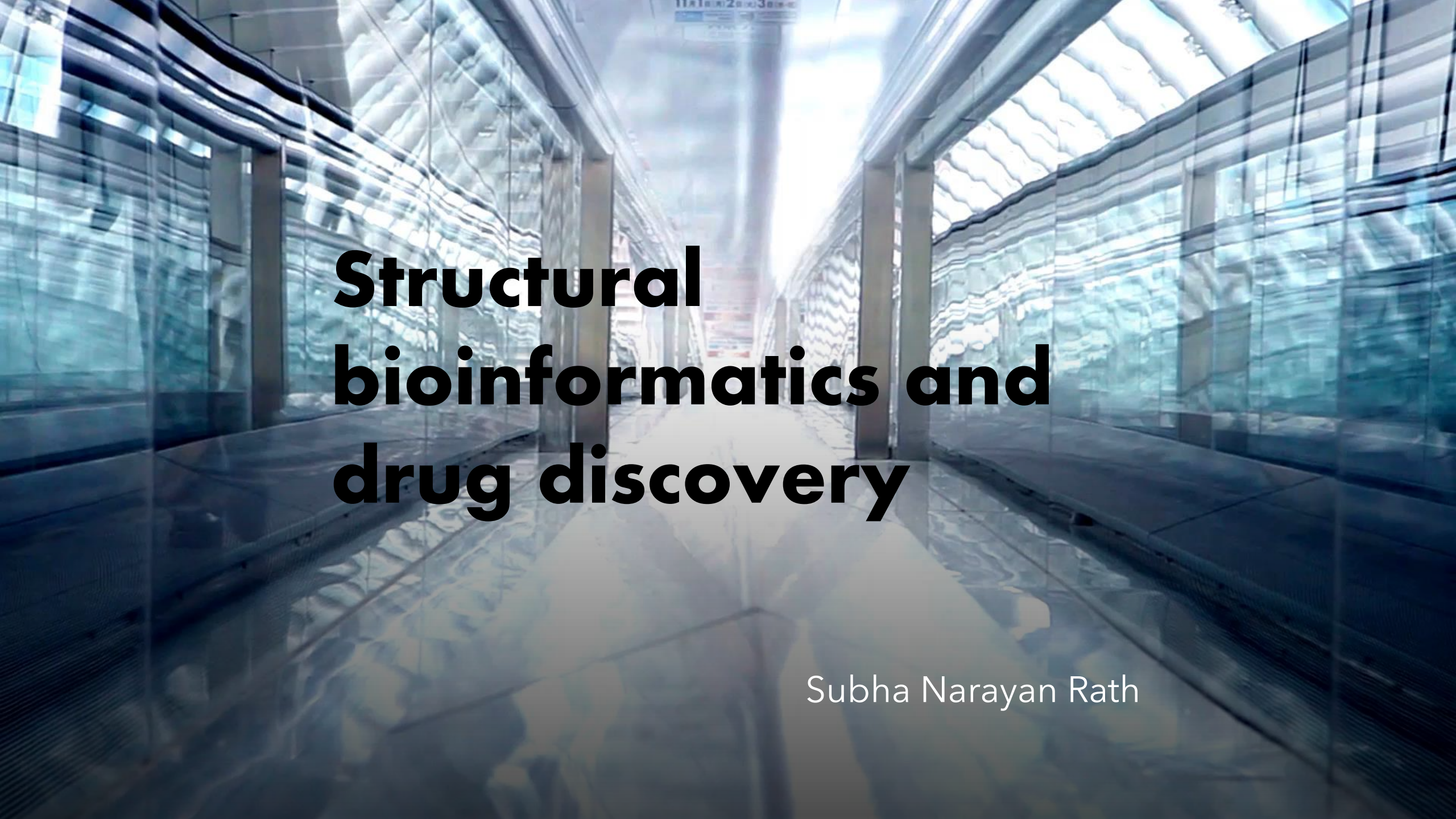


Structural bioinformatics and drug discovery

Subha Narayan Rath

Proteins have 3 profound properties:

- 1. great variety of structure and functions for all functions of cells and tissues
- 2. One place to synthesize ie. Ribosome and using only 20 different units
- 3. they fold spontaneously to the active native 3d state just by encoding the amino acid sequence
- At the end: HST forms to make the 3d structure, which is more informative
- Similar structure < > similar function
- PDB database is the main repository for 3d biological macromolecular structure data
- Source:
- 1. crystal structure
- 2. NMR models
- 3. others

	X-ray crystallography	NMR
Atomic resolution	Good	Reasonable
Hydrogens	Rarely determined	Determined
Molecule size	No restriction	Small proteins
Dynamics	Snapshot	Multi models
Membrane proteins	Problematic	
Procedure	Very long	long

Visualization: molecular graphics

- 1. rotation and translation
- 2. color specific parts of molecules
- 3. labelling of residues and atoms
- 4. Geometrical measurements (distances and angles)
- 5. schematic representation and structures to compare and alignment

Representation of molecules

- 1. stick and ball model
- 2. space filled model
- 3. Backbone: only connecting the C-alpha atoms
- 4. Schematic: helix: cylinder and strand: arrow
- 5. surface

Software to visualize them

- 1. Pymol
- 2. Rasmol
- 3. Chimera etc…..

- https://pymol.org/2/
- First download and install
- Pymol tutorial
- http://www.protein.osaka-u.ac.jp/rcsfp/supracryst/suzuki/jpxtal/Katsutani/en/interface.php

Structural analysis of proteins

- Examination of atomic interactions
- Examination of secondary structures
- Buried/exposed regions
- Analysis of ligands

Topics which can be solved

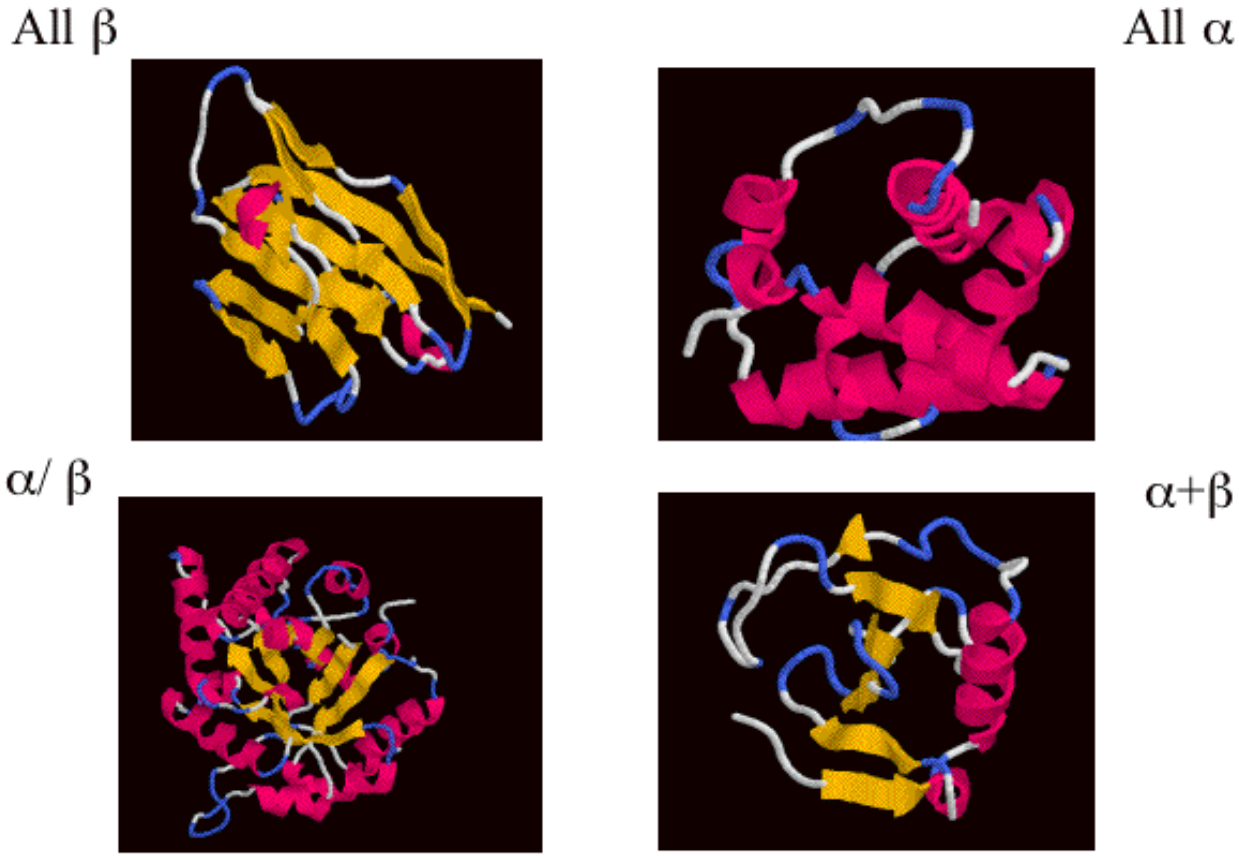
- Structural alignment
- Structural classification
- Secondary structure prediction or structure prediction
- Molecular docking
- Molecular dynamics

What properties of proteins used to detect structural similarity

- Sequence
- Type and number of secondary structures (HST)
- Structural arrangement of secondary structures
- Attributes of individual amino acids
- Distances between amino acids

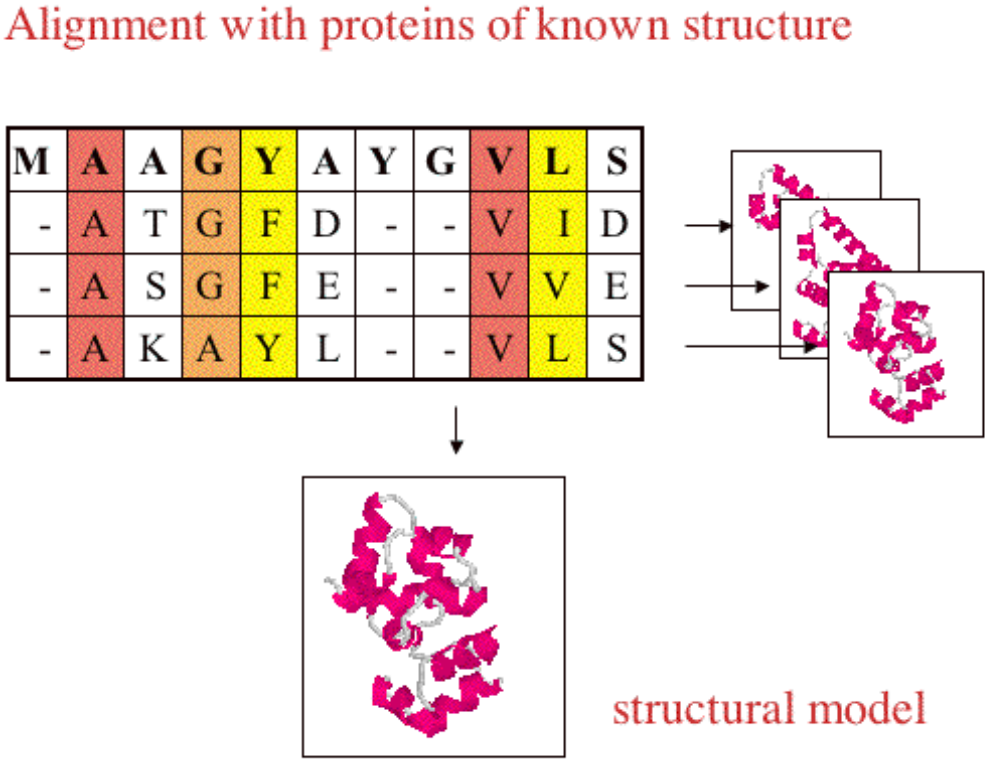
Structural classification

- All β
- All α
- α/β: β-α-β super secondary structures are present, could be linear or barrels
- α+β: both are separated at different parts of molecules
- The most common classification databases are:
- 1. SCOP
- 2. CATH



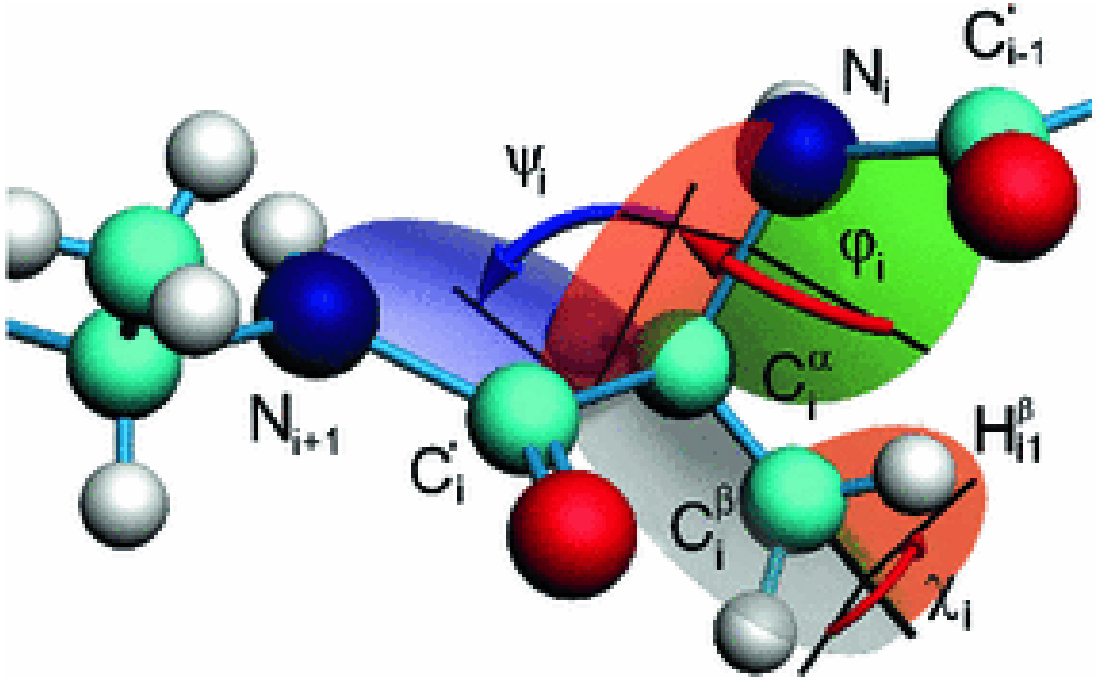
Secondary structure prediction

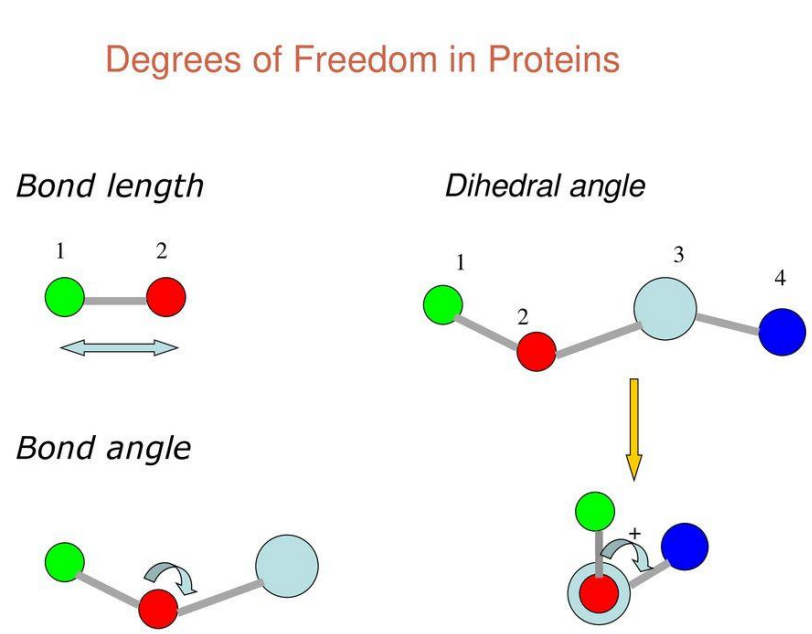
- Prediction of secondary structure is feasible and mostly used machine learning algorithm
- It is a bridge between linear and 3d strucures
- 1. Homology modelling: by aligning proteins of known structure (by SWISS MODEL)
- 2. Fold recognition (by known protein folds)
- 3. Ab initio method of modelling



Degrees of freedom:

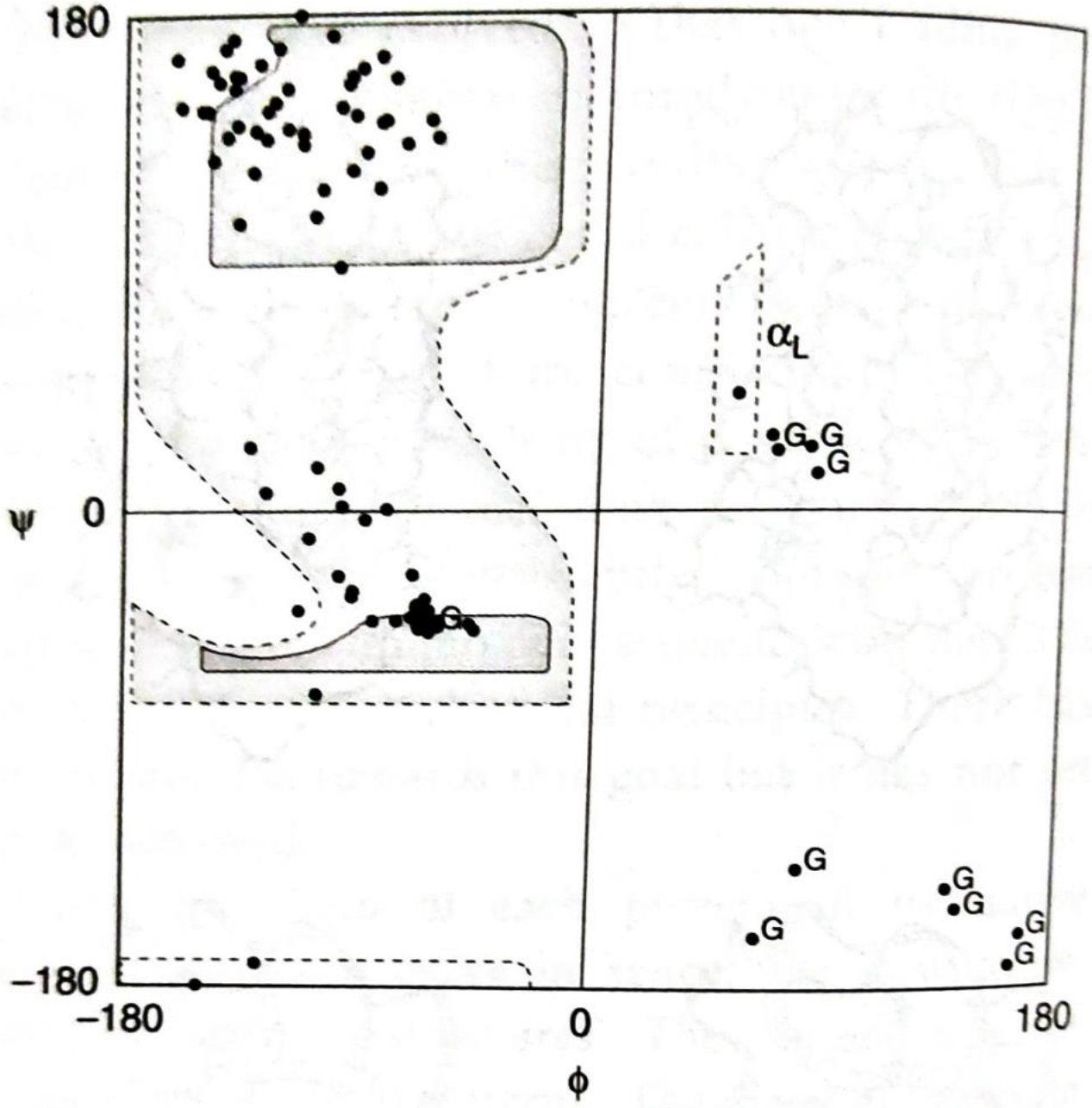
phi (φ) is the C(i-1),N(i),Ca(i),C(i) torsion angle and psi (ψ) is the N(i),Ca(i),C(i),N(i+1) torsion angle.





The Sasisekharan-Ramakrishnan-Ramachandran plot describes allowed mainchain conformations

- Rotation is permitted around the N-Ca and Ca-C single bonds of all residues (except Proline)
- The angles around these bonds and angle of rotation around the peptide bond define the confirmation of residues
- Peptide bonds are mostly planar with $\omega = 180$ degrees as they are in trans state
- Principle: two atoms can't occupy the same space limits the values of conformational angles
- The allowed range of ϕ, ψ with $\omega = 180$ fall into defined regions in a graph called Ramachandran plot
- Solid lines delimit energetically preferred regions of the angles: regions outside the broken lines are sterically disallowed
- Most amino acids falls into right handed helix or beta regions
- Glycine has additonal confirmation and can form left handed helix
- Only few are forced into energetically less favorable states
- Many but not all turns are short, surface exposed regions that contain charged or polar residues


180
0
−180
ψ
−180
0
180
ϕ
α_L
G
G
G
G
G
G
G
G
G
G

Role of sidechains

- 1. Size: glycine is smallest with only H and phynylalanine contains a benzene ring, one of the largest
- 2. Electric charge: acidic amino acids are negatively charged while basic ones are positively charged
- 3. Polarity: polar sidechains form hydrogen bonds to others and to water; others are electrically neutral and if unfavorable interactions with water: hydrophobic
- 4. Shape and rigidity

- E.g. D & E are similar and L & I are similar