HOMOLOGY MODELING

- Homology modeling, also known as comparative modeling of protein refers to constructing an atomic-resolution model of the "*target*" protein from its amino acid sequence and an experimental three-dimensional structure of a related homologous protein (the "*template*").

- The method of homology modeling is based on the observation that protein tertiary structure is better conserved than amino acid sequence.

- The sequence alignment and template structure are then used to produce a structural model of the target.

The various steps in Homology Modeling are:

- Template recognition and initial alignment
- Alignment correction
- Backbone generation
- Loop modeling
- Side chain modeling
- Structure energy minimisation
- Structure validation

3D STRUCTURE VALIDATION

Characteristics of a good model

- Makes chemical sense: normal bond lengths and angles, correct chirality ,flat aromatic rings, flat sp2-hybridised carbons.
- Makes physical sense: no non-bonded atoms, favourable crystal packing, related atoms display similar thermal disorder, occupancies of alternative conformations add up to one.
- Makes statistical sense: the model is the best hypothesis to explain available experimental data
- Makes structural sense: the model has a reasonable Ramachandran plot, not too many unusual side-chain conformations, or buried charges, residues are happy in their environment

•The conformation of the backbone of non-terminal amino acid residues is determined by three torsion angles

→ Phi Ci-1-Ni-CAi-Ci
→Psi Ni-CAi-Ci-Ni+1
→Omega CAi-Ci-Ni+1-CAi+1

•Due to the peptide bond's partial double-bond character, the omega angle is restrained to values near 0° (cispeptide) and 180° (trans-peptide)

•Cis-peptides are relatively rare and usually (but not always) occur if the next residue is a proline.

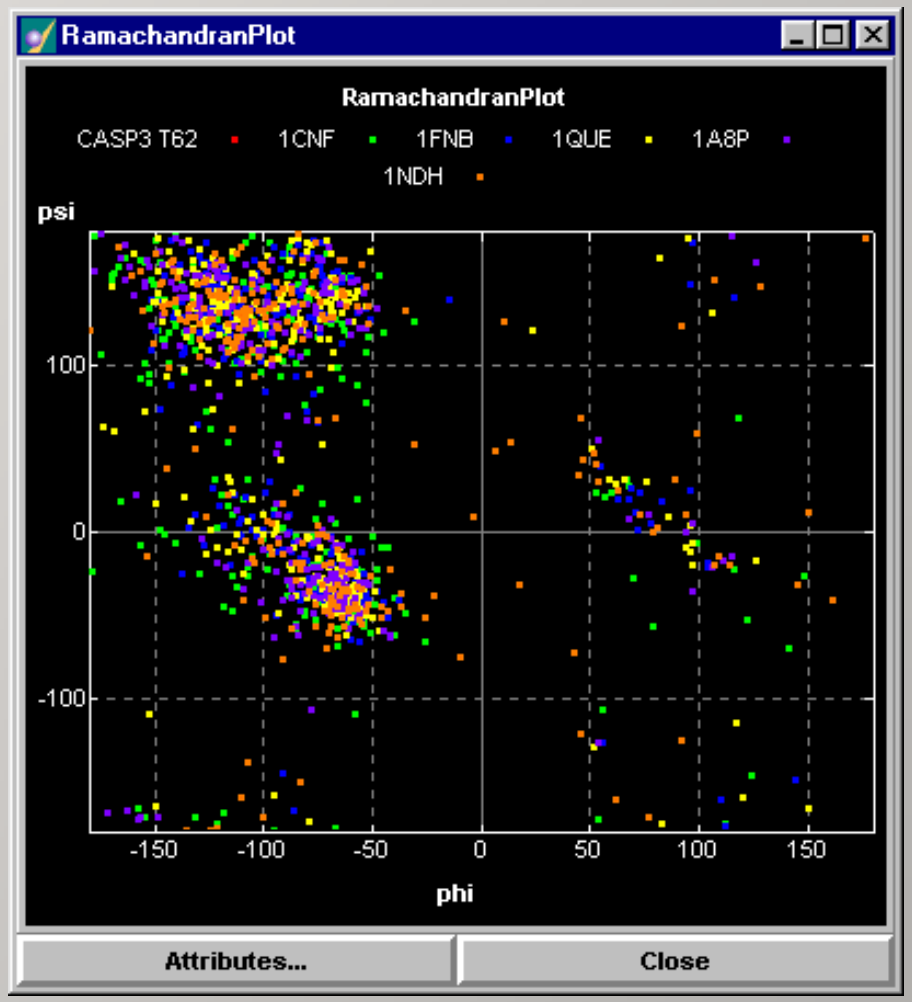
•The omega angle has little to offer as a validation check, although values in the range of ±20 to ±160 degrees should be treated with caution.

RAMACHANDRAN PLOT

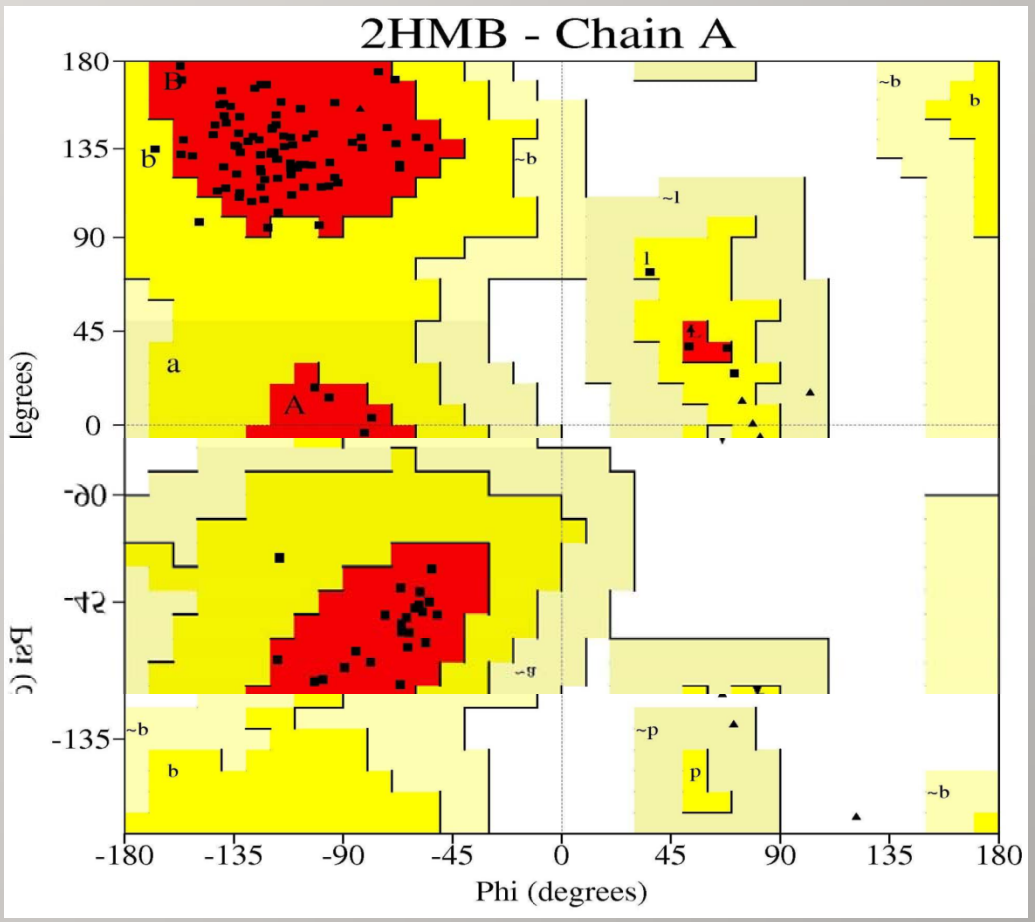
- The phi and psi torsion angles are less restricted, but due to steric hindrance, there are several preferred combinations of phi, psi values.

- A scatter plot of phi,psi values for all residues in a protein model is called a Ramachandran plot.

- This holds even for proline and glycine residues, although their distributions are atypical. Glycine - no side chain, adopts < and = angles in all 4 quadrants of Ramachandran plot



- Good models have most of the residues clustered tightly in the most-favoured regions with very few outliers
- Good, but low-resolution models, may have less pronounced clustering, but still have few outliers
- Poor models have no clustering and there are many outliers



A - Core alpha
L - Core left-handed alpha
a - Allowed alpha
l - Allowed left-handed alpha
~a - Generous alpha
~l - Generous left-handed alpha
B - Core beta
p - Allowed epsilon
b - Allowed beta
~p - Generous epsilon
~b - Generous beta

Most favoured regions [A,B,L]	113	95.8%
Additional allowed regions [a,b,l,p]	5	4.2%
Generously allowed regions [~a,~b,~l,~p]	0	0.0%
Disallowed regions [XX]	0	0.0%
Non-glycine and non-proline residues	10	

Good indicators of stereochemical quality

- planarity
- chirality
- phi/psi angles
- chi angles
- non-bonded contact distances
- unsatisfied donors and acceptors

- Bond Lengths : In protein structure, internal bond lengths should conform to known stereochemical values. Model bond lengths should lie within a few standard deviations of the mean bond length observed in the PDB.

- Bond Angles : Angles are normally distributed about the mean angles observed in the PDB.

  Internal bond angles and statistical outliers of the angles C-N-CA, NCA-C, N-CA-CB, CB-CA-C, CA-C-N, CA-C-O, and O-C-N.

- Dihedral Angles : Internal dihedral angles phi, omega, psi, and chi1 angles and the OC-CO dihedral. Kabsch & Sander secondary structure assignment for the residue.

- Nonbonded Contacts : Non-bonded and non-hydrogen bonded contacts where the van der Waals radii overlap by more than a prescribed value

OTHER TOOLS

- **Verify3D** →A comparison of the model to its own amino-acid sequence, using a 3D profile computed from atomic coordinates of the structure.

- **SCWRL 3.0** →A program for adding sidechains to a protein backbone based on the backbone-dependent rotamer library.

- **PROCHECK** → provide an idea of the stereochemical quality of all protein chains in a given PDB structure. They highlight regions of the proteins which appear to have unusual geometry and provide an overall assessment of the structure as a whole.

REFERENCES

- *www.proteinstructures.com/Structure/.../**Ramachandran-plot**.html*
- *www.nrbsc.org/pdf/**Model**ler_tutorial03.pdf*
- *www.colorado.edu/Outreach/BSI/pdfs/hearts**protein**.adv.pdf*
- *www.mendeley.com/.../**protein-threedimensional-structure validation** - United States*