Evaluating Protein Structures

Note Title 4/17/2007

Protein Geometry

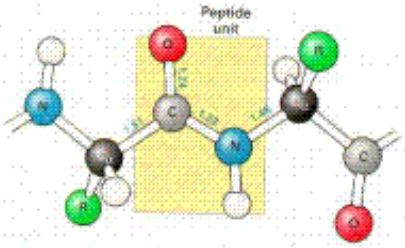
- Atom coordinates & radii (van der Waals)

vs.

- Bond angles & dihedrals (ϕ, ψ, χ, ω)
- Denavit-Hartenberg parameters
- Ramachandran plot

Energy potentials

- Energy terms from geometry
- Hydrogen bonds
- Packing or Contact analysis



Protein backbone geometry

Although PDB records coordinates
what matters locally is angles

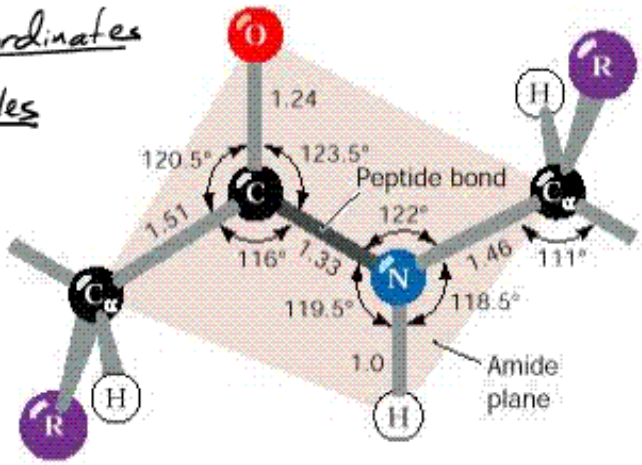
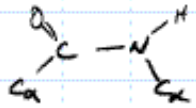


Figure 6-2. The trans peptide group. The bond lengths (in angstroms) and angles (in degrees) are derived from X-ray crystal structures.

cis places both C_α on same side



Protein back bone geometry

Although PDB records coordinates
what matters locally is angles

Parameters
Bond lengths : 2 atoms
Bond angles : 3 atoms
Bond dihedrals or torsions: 4

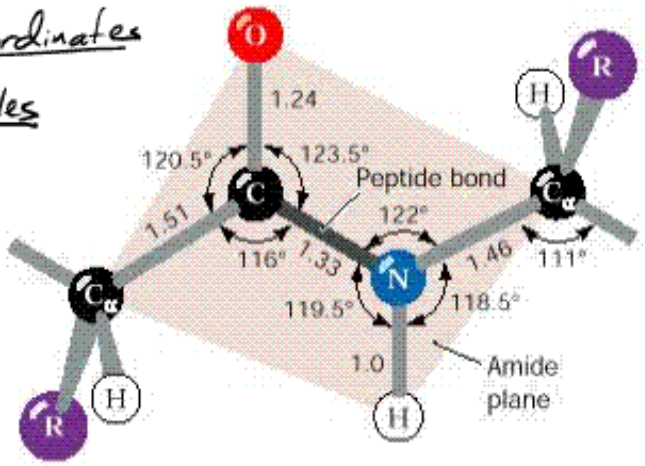
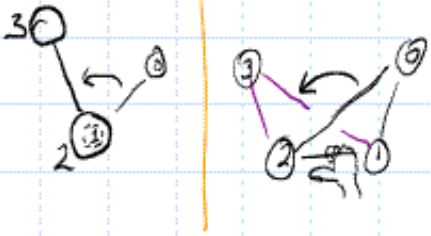
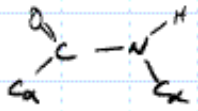


Figure 6-2. The trans peptide group. The bond lengths (in angstroms) and angles (in degrees) are derived from X-ray crystal structures.

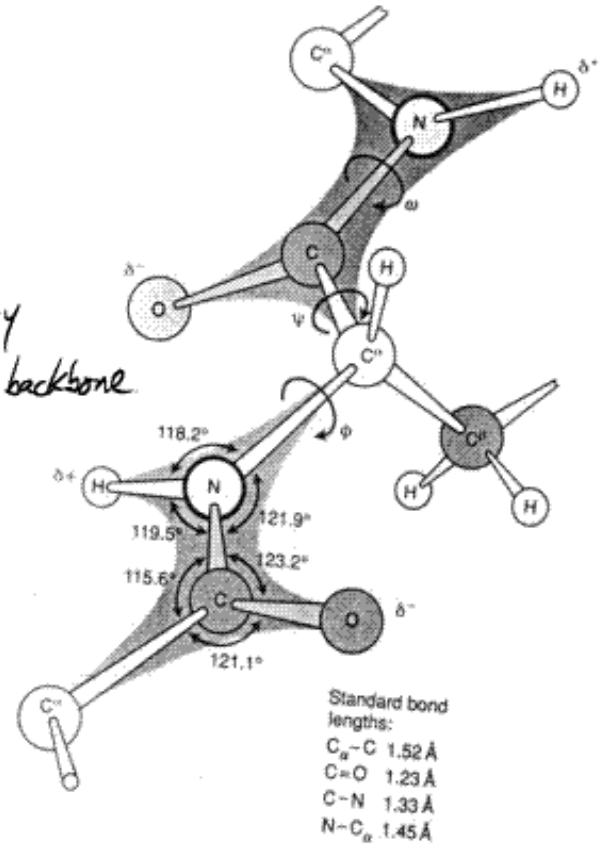


cis places both Cα on same side



ϕ, ψ angles

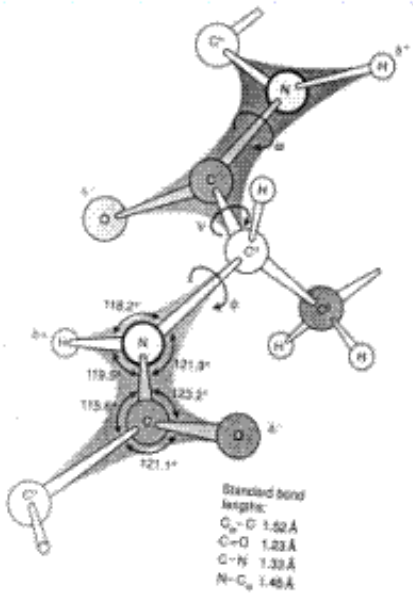
The torsion angles at Cα, ϕ and ψ, are the primary degrees of freedom on the backbone



Conversions

Coordinates → Angles
vector math & right hand rule

Angles → Coordinates
= Forward Kinematics
(Inverse Kinematics next week)
standard lengths (idealization)
rotation matrices or quaternions

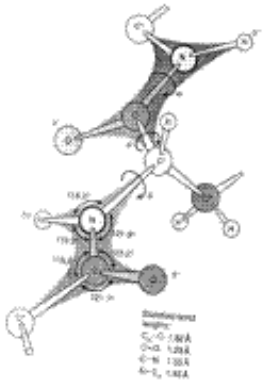


Project Ideas

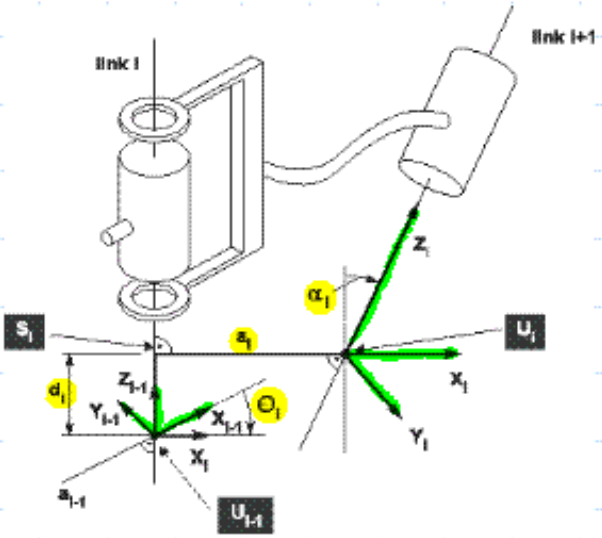
Idealize protein geometry while minimizing RMSD
Backbone motion shown in alternate conformations
Loop closure & Loop sampling

Denavit-Hartenberg Parameters

To move from one coordinate frame Y_{i-1} Z_{i-1} X_{i-1}
to another Z_i X_i Y_i



Find segment $\overline{S_i U_i}$ ⊥ rotation axis Z_{i-1} and Z_i

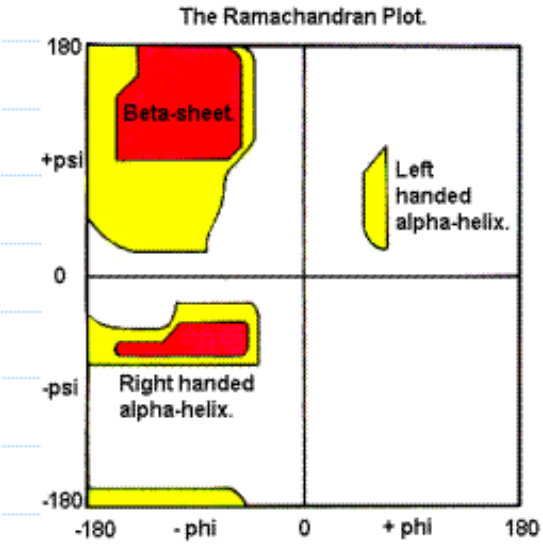


d_i = distance u_{i-1} to s_i on Z_{i-1}
θ_i = angle from X_{i-1} to X_i = $u_i - s_i$
a_i = distance s_i to u_i (along X_i)
α_i = angle from Z_{i-1} to Z_i around Y_i

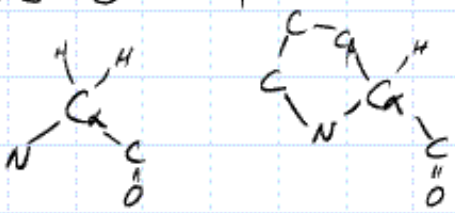
Ramachandran plot of ϕ vs. ψ

Derived originally from hard-sphere collisions;
Useful for checking structure & scoring.

http://www.biochem.umd.edu/biochem/kahn/teach_res/ramac

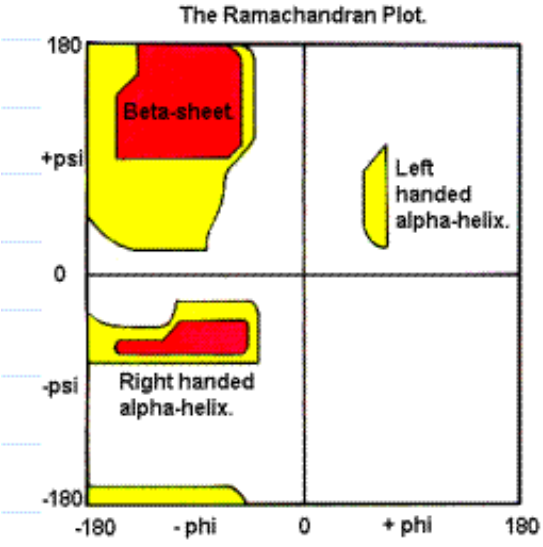


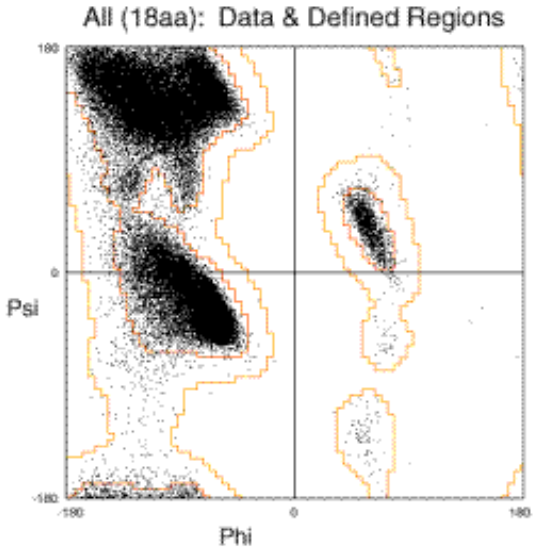
NB: GLY & PRO are special



Ramachandran plot of ϕ vs. ψ

Derived originally from hard-sphere collisions;
later by PDB analysis & quantum mechanics.
Show preferences for most amino acids - GLY & PRO are special.

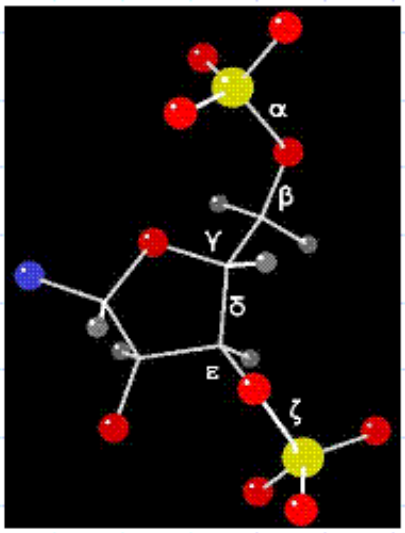


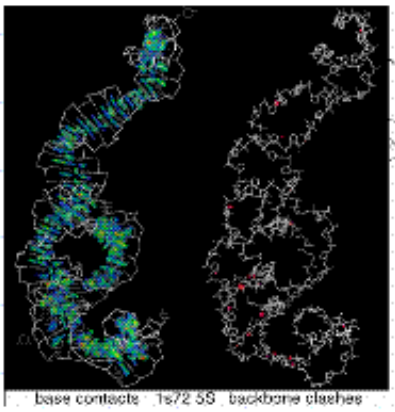


Project Idea: Ramachandran Plot for RNA?

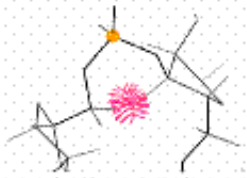
RNA backbone has 6 dihedrals, 5 of which are free.

RNA structures are poor at choosing backbone angles.




base contacts 1s72 SS backbone clashes


base contacts closeup


backbone clash closeup

BREAK

Geometric potentials (Empirical)

Use constants derived from known structures.

bond lengths

$$\sum_{bonds} K_r (r - r_{eq})^2$$

bond angles

$$\sum_{angles} K_\theta (\theta - \theta_{eq})^2$$

bond dihedrals (torsion)

$$\sum_{dihedrals} \frac{Vn}{2}\left(1 + \cos(n\phi - \gamma)\right)$$

Lennard-Jones + electrostatics

$$\sum_{i<j}\left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{D R_{ij}}\right]$$

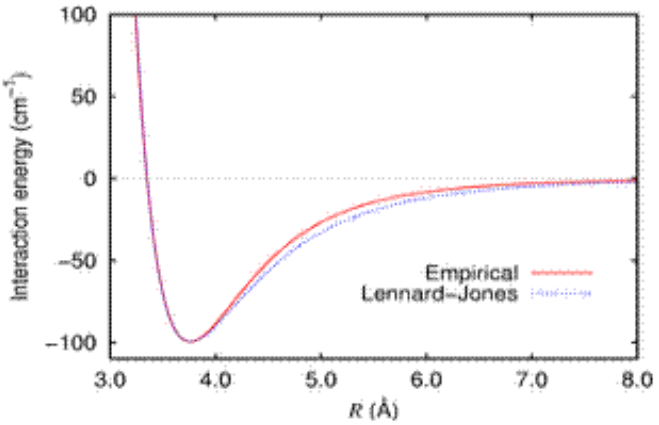
Distant charged interactions

Electrostatic $\oplus_{q_i} \longleftrightarrow^{r} \ominus_{q_j}$ $\frac{q_i q_j}{D \cdot r}$ D = dielectric constant

problematic: dielectric depends on environment
force (inverse square) balanced by increase in volume (r^2)

Lennard-Jones $\frac{A}{r^{12}} - \frac{B}{r^6}$

Eg. for Argon:



Geometric potentials

What is missing? packing, solvent, entropy
Hydrogen bonds

bond lengths $\sum_{bonds} K_r (r - r_{eq})^2$

bond angles $\sum_{angles} K_\theta (\theta - \theta_{eq})^2$

bond dihedrals (torsion) $\sum_{dihedrals} \frac{Vn}{2}(1 + \cos(n\phi - \gamma))$

Lennard-Jones + electrostatics $\sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{D R_{ij}} \right]$

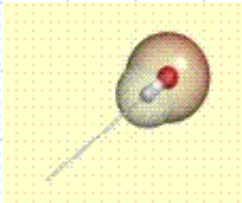
Hydrogen bonds

Crucial

- stabilize α helix, β sheet
- form with water on surface; nearly always satisfied inside.

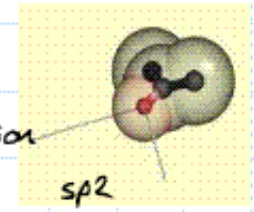
Donor

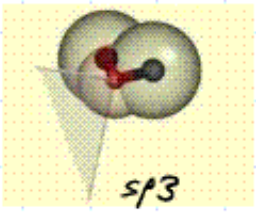
2.6 - 3.2 Å distant from Acceptor



Acceptor

depends on hybridization

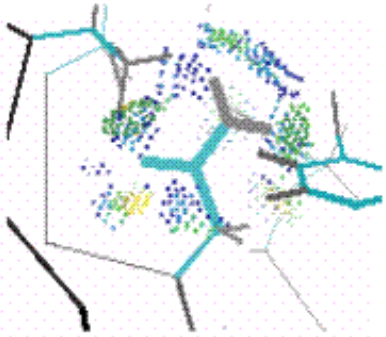




Remember: most PDB files don't list Hydrogens!

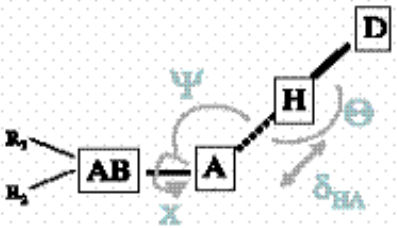
Two evaluations of Hydrogen bonds

Reduce (Richardson lab Duke)



function of overlap volume
with simple visualization

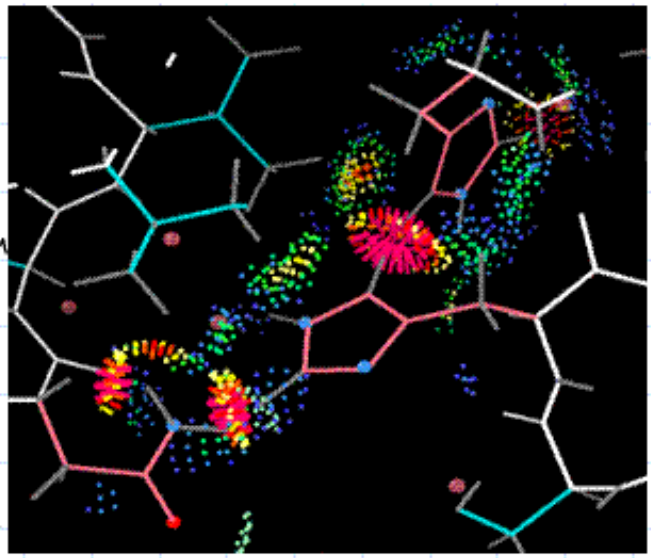
Rosetta (Baker lab U Wash)



function of AHdist, θ, ψ
with statistically-derived parameters

Molprobity

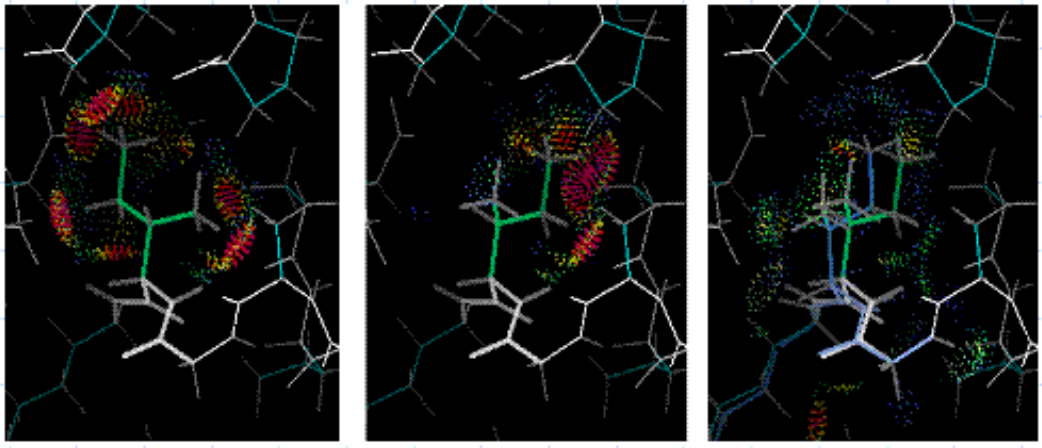
Analyze "Contact Dots"



Simply place dots on atom surfaces, and color blue/green if good (van der Waals attraction or Hydrogen bonds) and yellow/red if bad (clash overlap)
Indicates error in fitting to electron density, usually.
Molprobity web service performs this analysis. (Exersize Sess.

Sidechain refinement with added Hydrogens

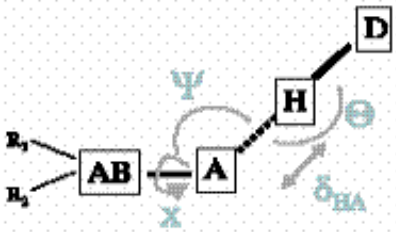
An example of structure repair.



original | flip | flip + adjustment resolves clashes

Project Idea: Find candidate repairs automatically
(I can help with probe/reduce source code.)

Project Idea:



By comparing geometry from PDB & ab initio folding, improve Hydrogen bond statistics → functions for Rosetta

pick one
- for charged donors (current bonds are too long)
- for bifurcated bonds (two acceptors shouldn't count twice)
- for χ angles (not used at present)
- for aromatic rings
- for bond networks

I have lots of initial work - see hbonds-rosetta06.ppt on line.