

2.1 – Macromolecular energetics

Bioinformatics - Biophysics

Outline

- Macromolecular energetics
 - Sources of data
 - Energy components
 - Mutational analysis
- Energy evaluation
 - Quantum mechanics methods
 - Classical methods
 - Solvation
 - Statistical Potentials
 - Knowledge based methods

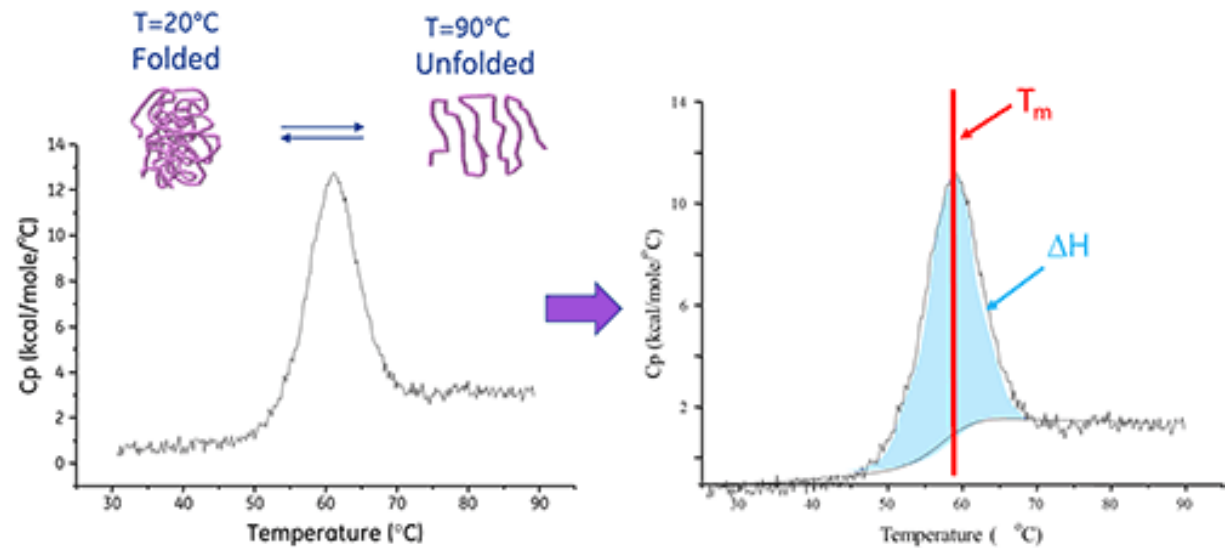
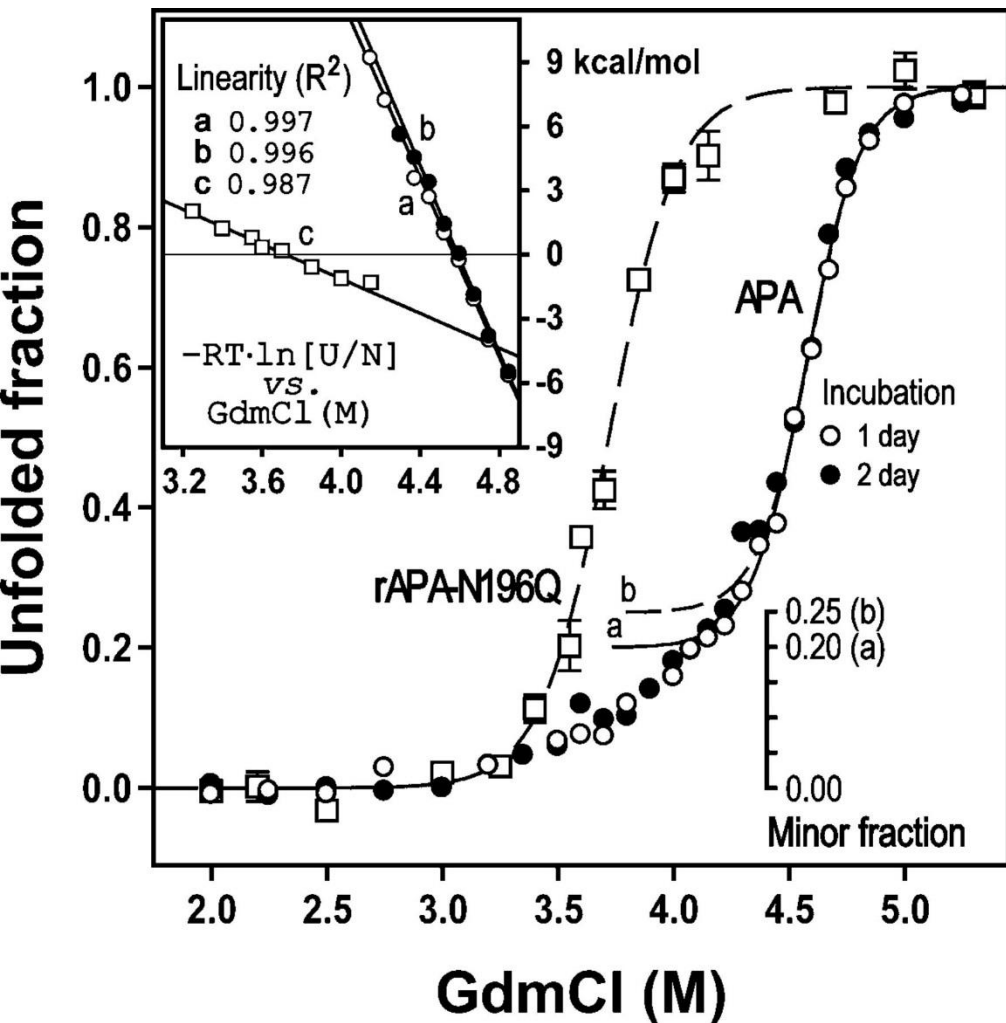
Objectives/strategies

- Understanding the **energetic basis** of macromolecular stability
 - Influence of residues/interactions (mutational analysis)
 - Stability improvement
 - Special proteins, special media
 - Special NAs (targeted therapies)
 - Annotation of sequence variants
 - pathology, target discovery
- Understanding folding and conformational evolution

Experimental approaches

$$\Delta G = \Delta G_w - m[D]$$

$$\Delta\Delta G = m \Delta[D_{1/2}]$$

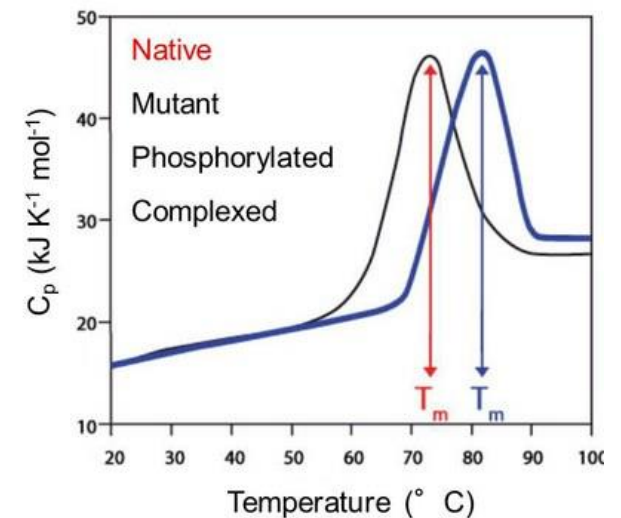


$$\Delta H(T) = \Delta H_m + \Delta C_P(T - T_m)$$

$$\Delta S(T) = \frac{\Delta H_m}{T_m} + \Delta C_P \ln(T/T_m)$$

$$\Delta\Delta H(T) \sim \Delta C_P \Delta T_m$$

$$\Delta\Delta S(T) \sim \Delta H_m \left(\frac{1}{T'_m} - \frac{1}{T_m} \right) + \Delta C_P \ln \left(\frac{T'_m}{T_m} \right)$$



Energy components (summary)

- Entropic
 - Conformational entropy (ΔG_{\nearrow})
 - Hydrophobic effect (ΔG_{\searrow}) (Solvation entropy)
- Enthalpy
 - Van der Waals (ΔG_{\searrow})
 - Hydrogen bonds (ΔG_{\sim})
 - Electrostatics (ΔG_{\searrow})
- Solvation Enthalpy (ΔG_{\searrow})
 - Electrostatics
 - VdW

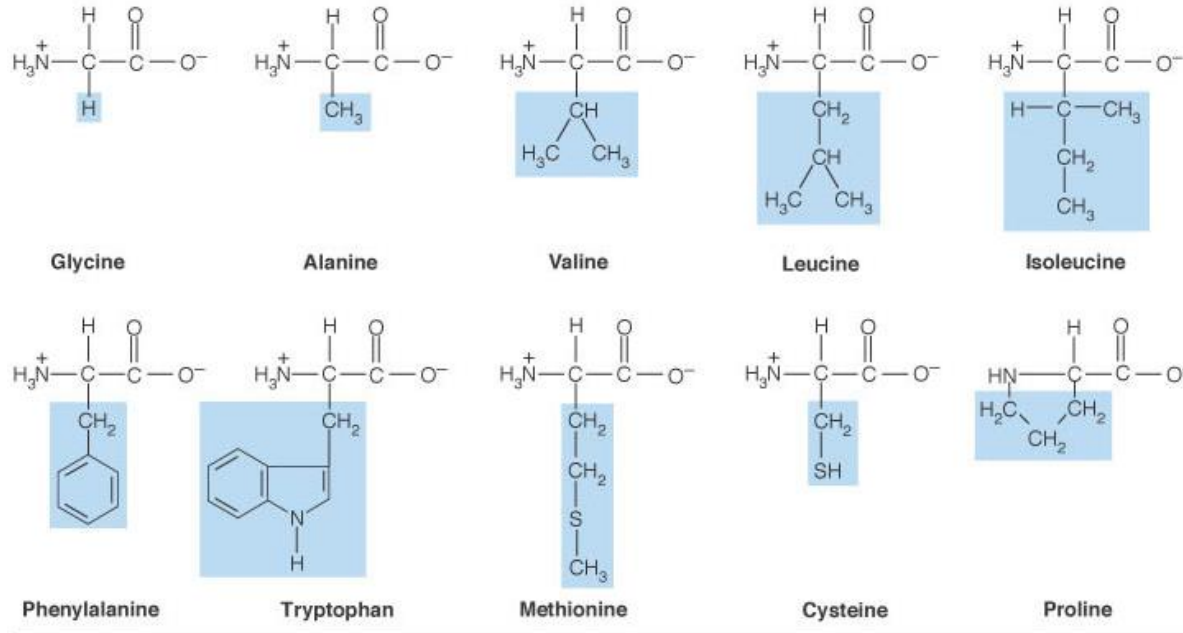
Enthalpy

$$\Delta G = \Delta H - T\Delta S$$

Entropy

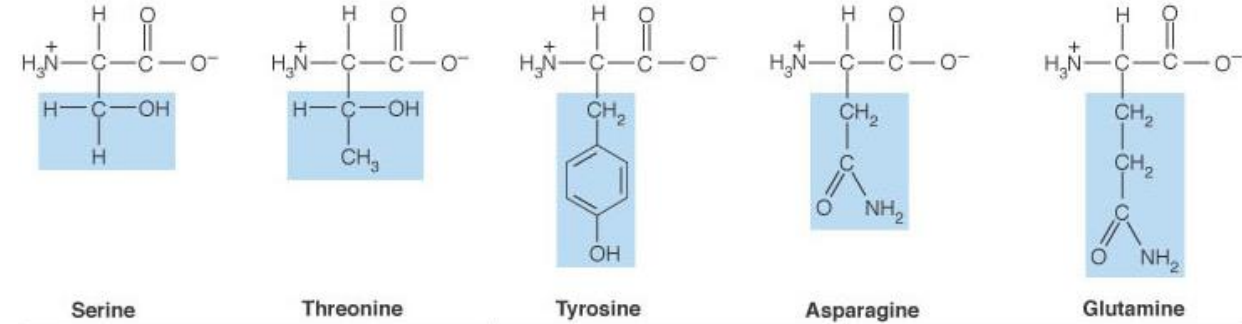
Electrostatics (dipolar)
Hydrogen bonds
Van der Waals

Neutral Nonpolar Amino Acids

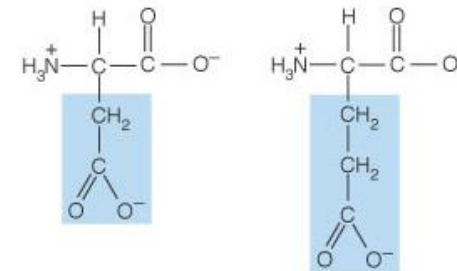


Hydrophobic effect
Van der Waals

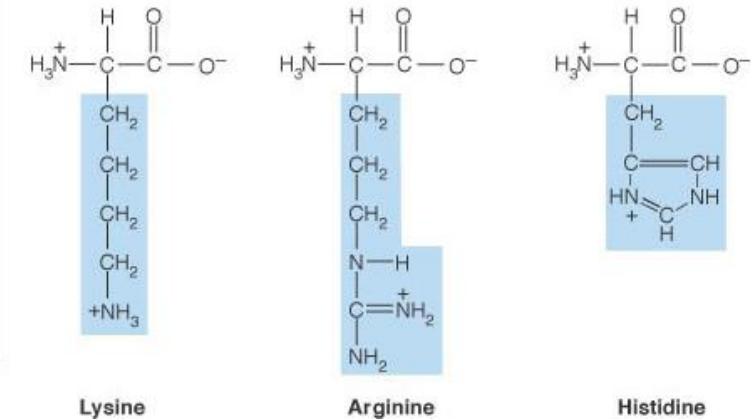
Neutral Polar Amino Acids



Acidic Amino Acids



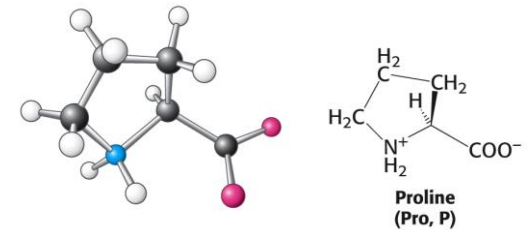
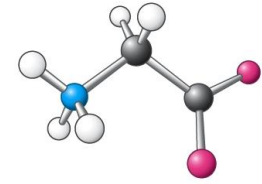
Basic Amino Acids



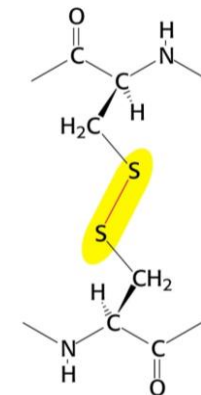
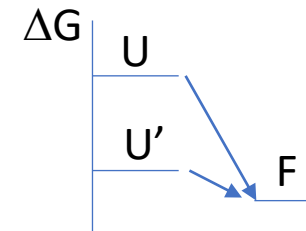
Electrostatics (Full charges)
Hydrogen bonds
Van der Waals

Conformational entropy

- Entropy decreases on folding
$$\Delta S = 2-6 \text{ cal/K.mol.res}$$
$$\Delta G = 1.2 \text{ kcal/mol-res (100 res } \Rightarrow 120 \text{ kcal/mol)}$$
- Folded structures are (almost) unique but unfolded ones are complex ensembles
 - Increased chain flexibility increases the number of “unfolded” microstates (entropy \nearrow)
 - Flexible residues (Gly) destabilize $\Delta S = 6.5 \text{ cal/K.mol.Gly}$
 - Rigid Amino acids or structures (Pro, -SS-) stabilize



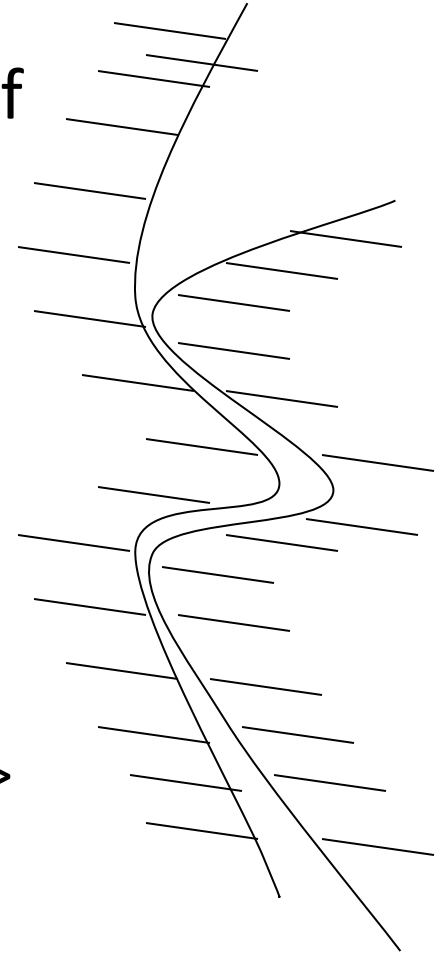
Pro (P)



Cys-SS-Cys

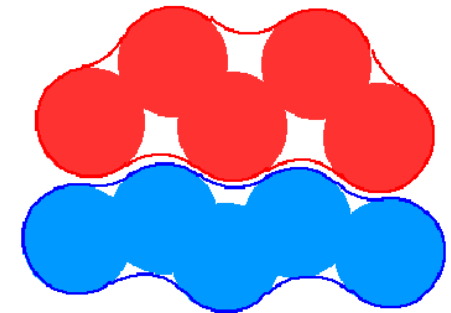
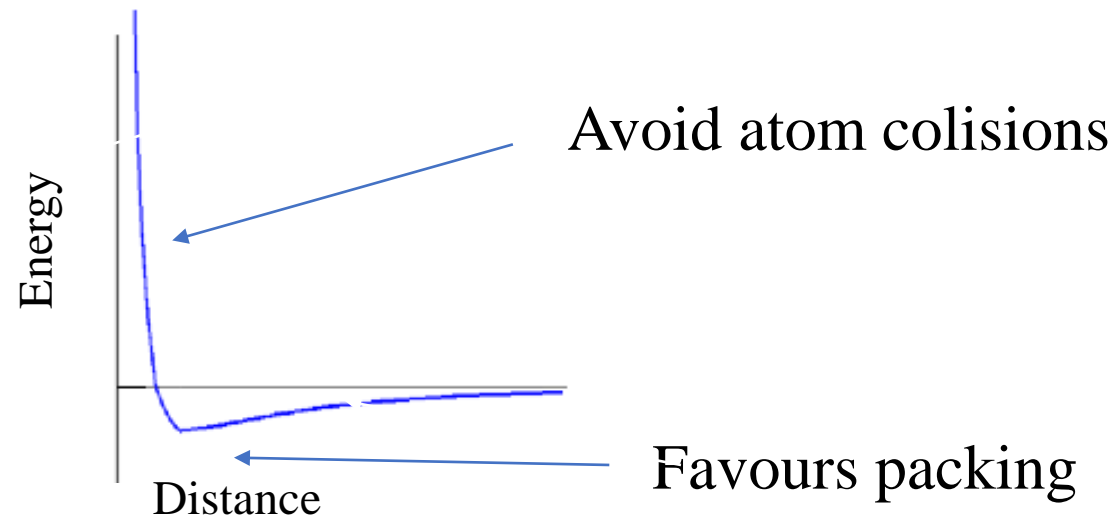
Hydrophobic effect

- Solvent entropy increases when the number of contacts with non-polar residues decrease
- Is slightly higher than conformational entropy.
- Driving force on protein folding or formation of protein complexes
- Roughly proportional to molecular surface
 - Aprox $30 \text{ cal}/\text{\AA}^2$ apolar surface (100 res $4,860 \text{ \AA}^2 \Rightarrow -145 \text{ kcal/mol}$)



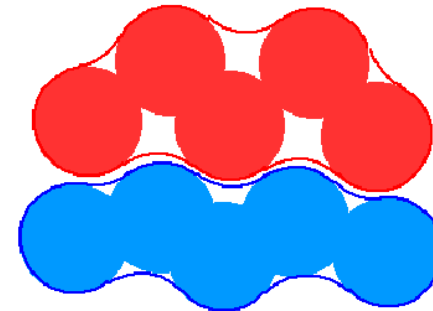
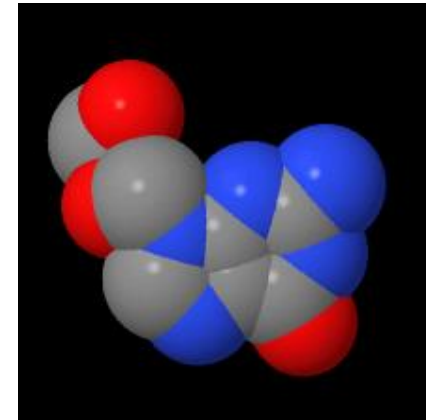
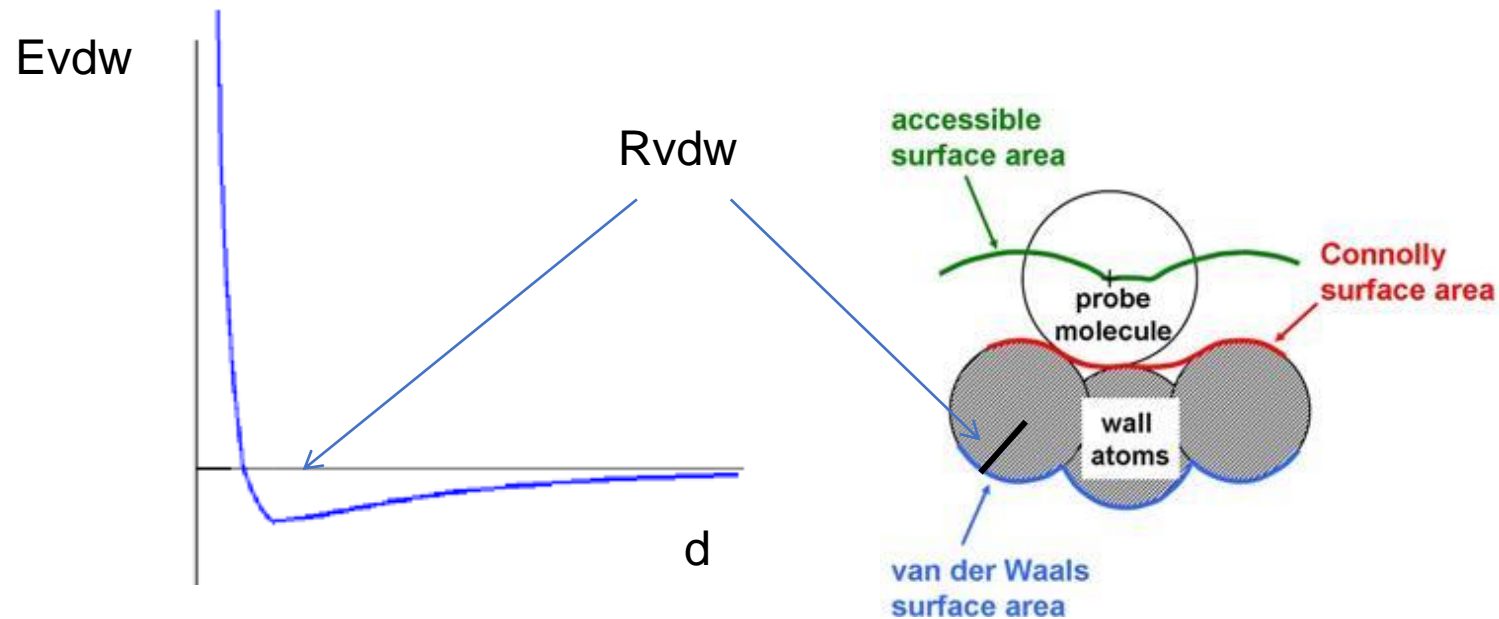
Van der Waals energies

- Caused by a combination of transient interactions between electronic clouds (“dispersion”), and repulsion between them.
- Small but cumulative (~ -1000 kcal/mol average protein)
- Control surface-surface interaction

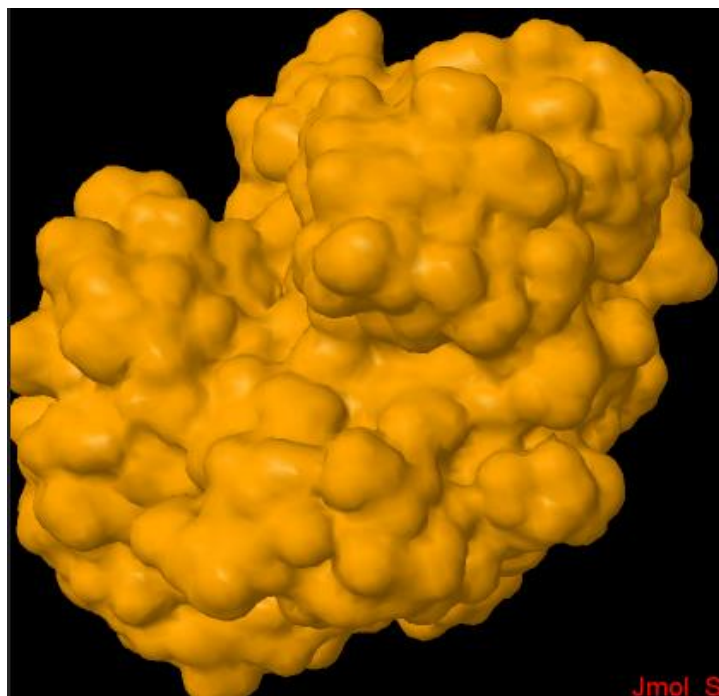


Molecular surface

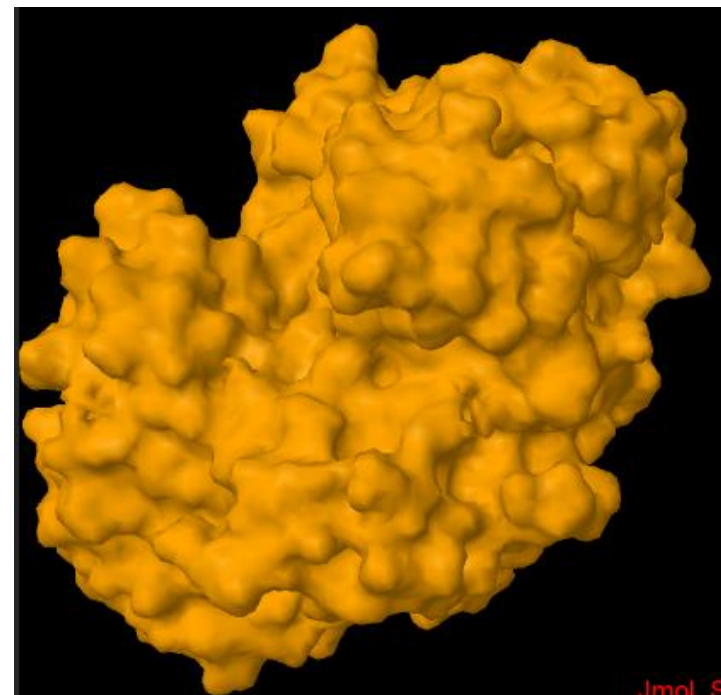
- Required to highlight interaction surfaces
- Based on VdW surfaces



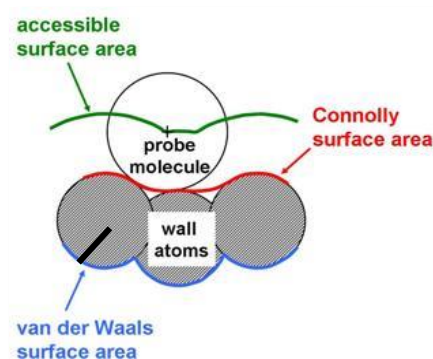
Molecular surfaces



SASA

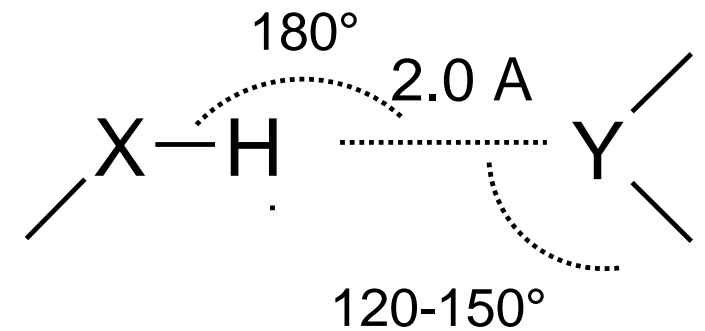


Connolly



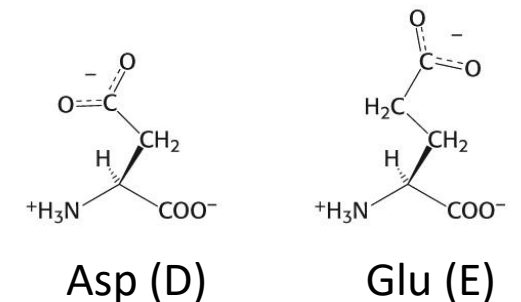
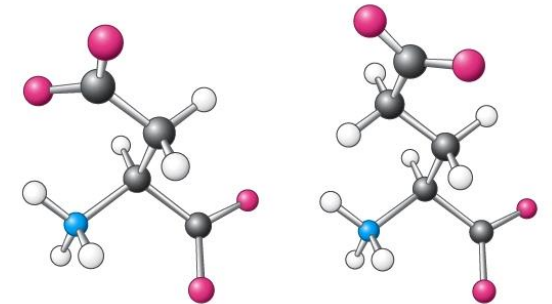
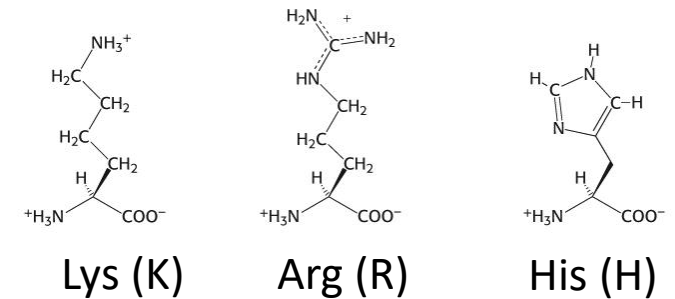
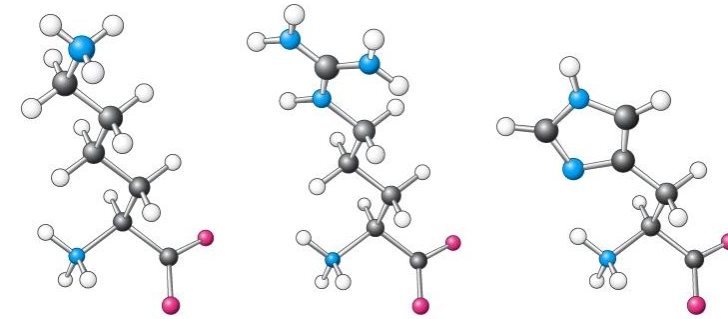
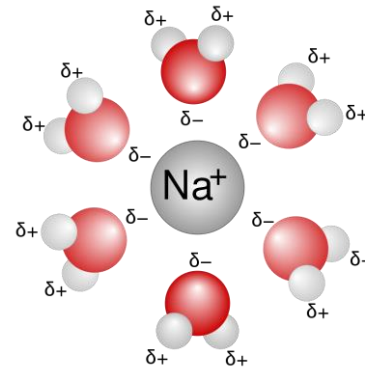
Hydrogen bonds

- Dipolar interactions (N, O, S)-H with electron pairs in N, O, S
- **Directional**: geometry defines the intensity of the interaction
- Usual energy around 1-5 kcal/mol
- Little neat estabilization
 - Unfolded state forms hydrogen bonds with solvent
- Define directionality to interactions
- All hydrogen bonds should be fulfilled in a folded structure
- Usually defined following a simple distance criteria
 - Proper analysis would require Quantum mechanics calculations



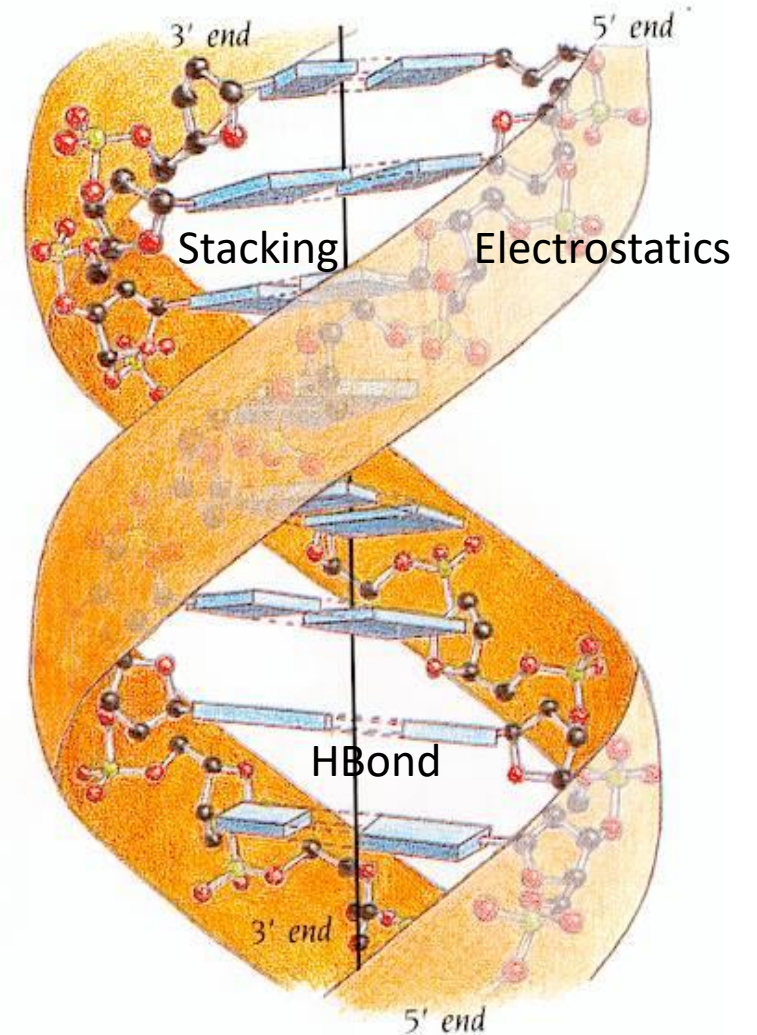
Electrostatics

- Interaction between **charged residues**
- **Dipolar** (directional) interactions (“charged” hydrogen bonds)
- Largest enthalpic component of **solvation** (Water relative dielectric constant ~ 80)
- Unpaired ionic residues are strongly destabilizers unless exposed to solvent
- Buried ionic pair are strong stabilizers
- Ionization states depend on environment
- Minor interactions:
 - Helix-dipoles, π -cation, π - π



Nucleic acids specifics

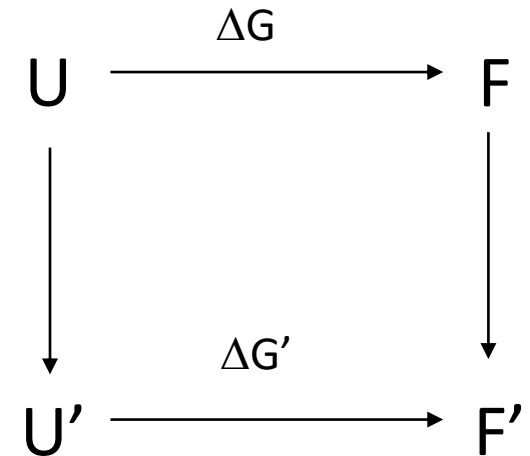
- Poly-anions: strong electrostatics
 - Repulsion between main-chain phosphates
 - Highly sensitive to ionic environment
- Base-base interactions
 - Hydrogen bonds
 - Basis of all recognition events (interchain, external)
 - Strong in gas phase, lower in solution
 - Stacking
 - Non specific
 - Major contribution to stability (~ -2 kcal/mol/step)
 - Highly cooperative



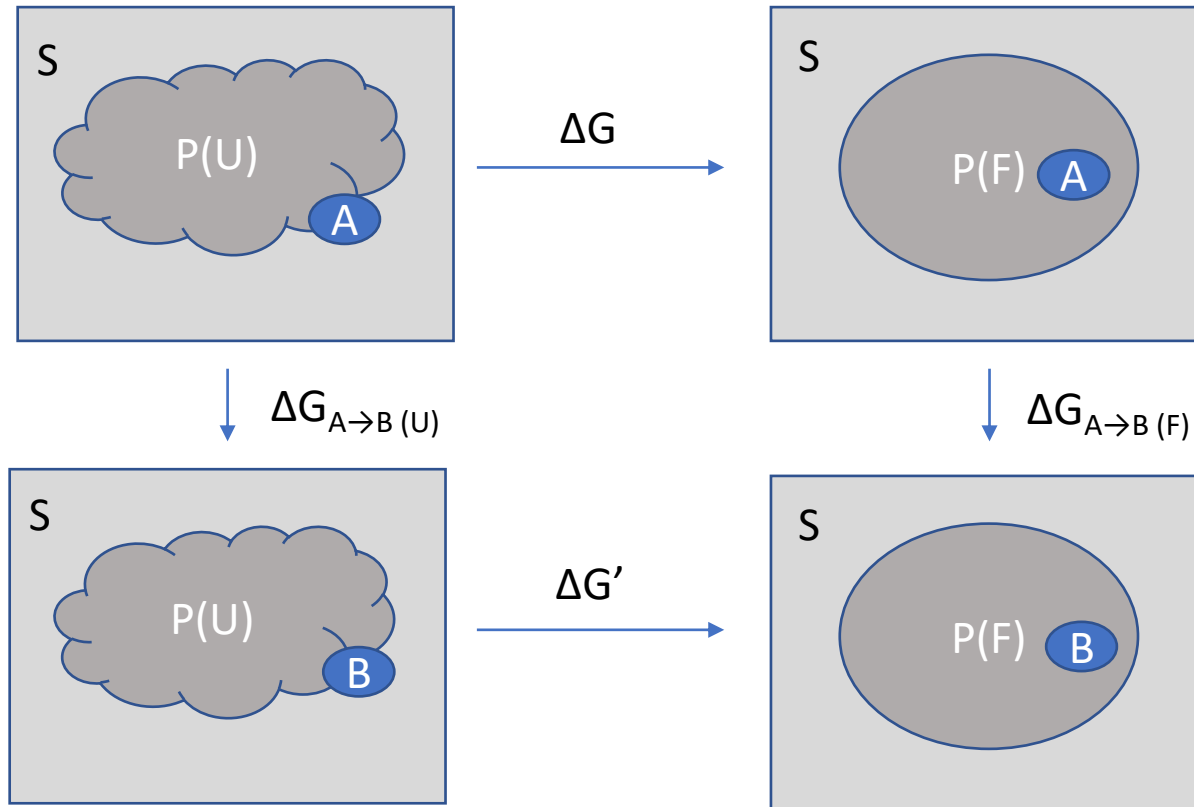
Mutational analysis

- Data derived from mutagenesis to change/remove interactions
- Relevant energy is the difference ($\Delta\Delta G$) between folded and unfolded states (including solvation contribution)
- Unfolded state
 - Exposed to solvent
 - Unstructured (or non-native structures)
 - Hydrogen bonds with water
- Folded state
 - Mostly structured
 - Internal interactions
 - Hydrophobic core buried

$$\Delta\Delta G = \Delta G' - \Delta G$$



Understanding mutational analysis



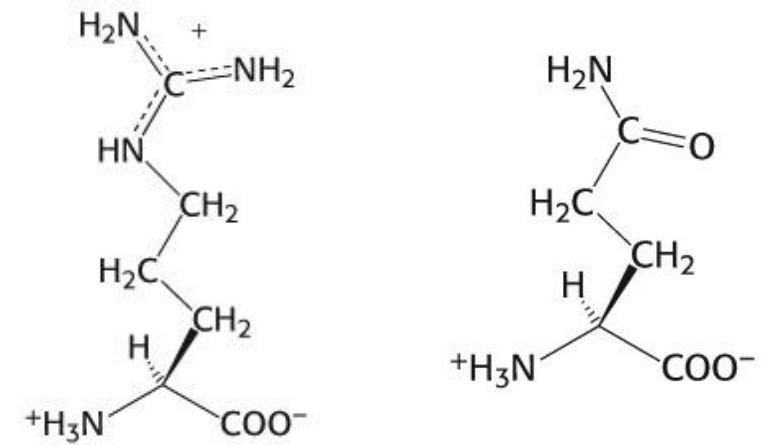
$$\Delta\Delta G = \Delta\Delta G_{A \rightarrow B}$$

- Internal $A \rightarrow B$
- $\Delta\Delta G$ interaction (P) $A \rightarrow B$
- $\Delta\Delta G$ solvation $A \rightarrow B$
- $\Delta\Delta G$ internal P
- $\Delta\Delta G$ solvation P

$$\Delta\Delta G_{A \rightarrow B} = \Delta\Delta G_{A \rightarrow B(F)}^{internalAB} + \Delta\Delta G_{A \rightarrow B(F)}^{internalP} + \Delta\Delta G_{A \rightarrow B(F)}^{interP,AB} + \Delta\Delta G_{A \rightarrow B(F)}^{solvationP} + \Delta\Delta G_{A \rightarrow B(F)}^{solvationAB} - \Delta\Delta G_{A \rightarrow B(U)}^{internalAB} - \Delta\Delta G_{A \rightarrow B(U)}^{internalP} - \Delta\Delta G_{A \rightarrow B(U)}^{interP,AB} - \Delta\Delta G_{A \rightarrow B(U)}^{solvationP} - \Delta\Delta G_{A \rightarrow B(U)}^{solvationAB}$$

Example: R-E \rightarrow Q-E (buried)

- 1. Check aminoacid properties
 - R: Charged (HB, Electrostatic) Q: Polar (HB)
- 2. Assess energy balances (qualitative)



	R(F)	Q(F)	R(U)	Q(U)	$\Delta\Delta G(F)$	$\Delta\Delta G(U)$	$\Delta\Delta G$
Int. Polar	--	--	0	0	0	0	0
Int Elect	---	0	0	0	+++	0	+++
Int Vdw	-	-	0	0	0	0	0
Solvation	0	0	---	-	0	++	--
						TOTAL	+

$$\Delta\Delta G_{A \rightarrow B} = \Delta\Delta G_{A \rightarrow B(F)}^{interP,AB} + \Delta\Delta G_{A \rightarrow B(F)}^{solvationAB} - \Delta\Delta G_{A \rightarrow B(U)}^{interP,AB} - \Delta\Delta G_{A \rightarrow B(U)}^{solvationAB}$$

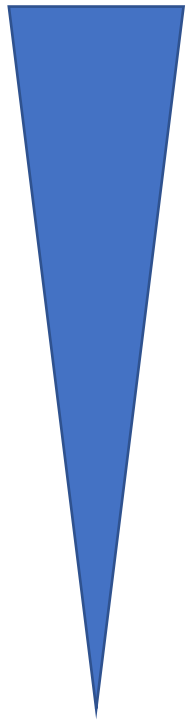
Energy evaluation

Energy from structure.

- Aim: To obtain an “energy” estimation from molecular structure.
- Energy can be used as scoring function when dealing with structural data (prediction, simulation, ...)
- Energy does not need to correspond to a “physical” energy.

Structure to energy (Summary)

- Different approaches: different levels of accuracy



Quality

- Quantum mechanics
 - Rigorous, allow chemical processes, limited to small systems
- Classical mechanics
 - Less rigorous but accurate, no size limit, does not allow for chemical transformations
- Statistical/coarse-grained potentials
 - Knowledge-based, accurate for low resolution and non-realistic representations
- Empirical scores
 - Simple parameter set fitted to experiment

Quantum mechanics

- Based in first principles.
- Rigorous, Universal

$$\hat{H} |\Psi(t)\rangle = i\hbar \frac{d}{dt} |\Psi(t)\rangle = \frac{\hat{\vec{p}}^2}{2m} |\Psi(t)\rangle + V(\hat{\vec{r}}, t) |\Psi(t)\rangle$$

$$\hat{H} |\varphi_n\rangle = E_n |\varphi_n\rangle$$

- Too costly
- Can be applied only to small systems (active sites, chemical transformations)



Erwin Schrödinger
Nobel Prize 1933

Quantum mechanics usage

- Parameterization of other methods
- Characterization/Parameterization of new molecules
- Distorted structures
- Fine analysis of energy terms (vdw, HB)
- Kinetic mechanisms (bond reorganization)
 - Usually combined with classical methods (QMMM)
- “Quantum” effects: Polarization, tunnel, ...

Classical mechanics

- Empirical approximations (require parameterization)
- Computationally very efficient
- No size limit
- Less accurate than quantum mechanics
- Solvation and entropy cannot be included!!

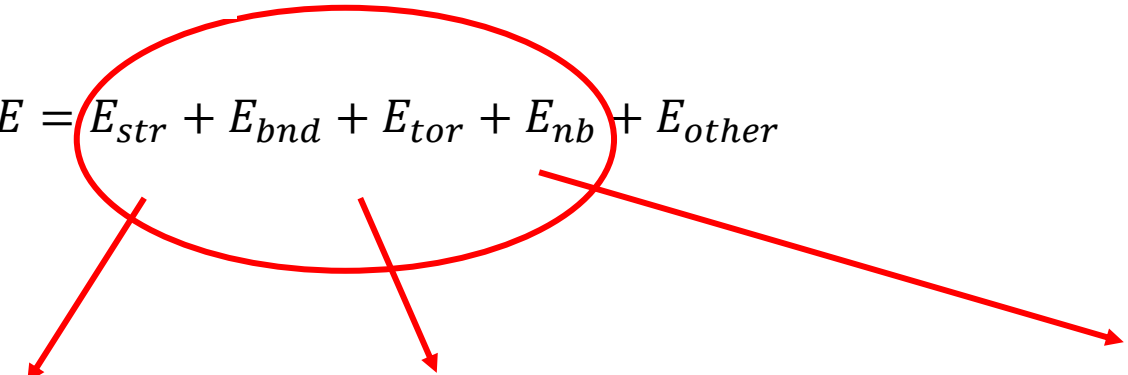


Martin Karplus
Nobel Prize 2013

Bonded-terms Non bonded-terms Other restrains

$$E = \underbrace{E_{str} + E_{bnd} + E_{tor}}_{\text{Bonded-terms}} + \underbrace{E_{nb}}_{\text{Non bonded-terms}} + \underbrace{E_{other}}_{\text{Other restrains}}$$

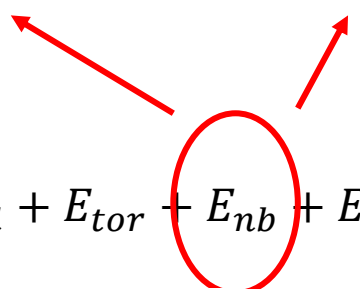
Bonded-terms

$$E = E_{str} + E_{bnd} + E_{tor} + E_{nb} + E_{other}$$


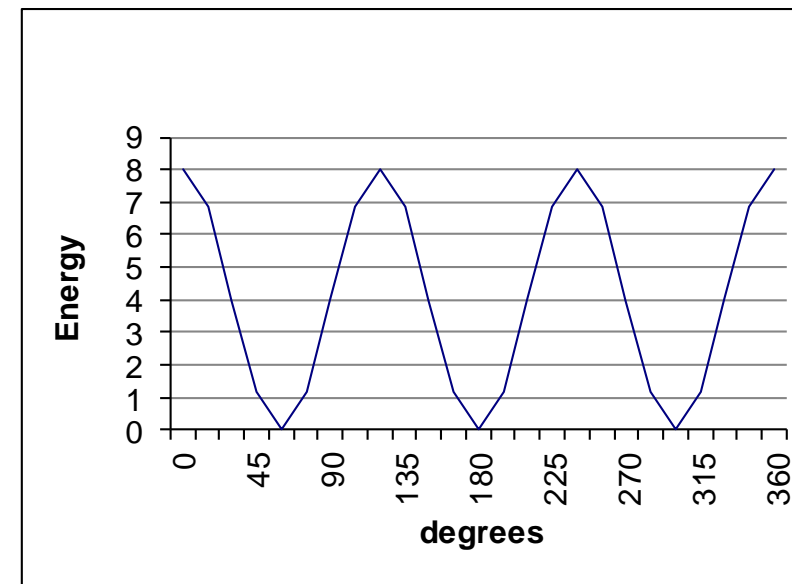
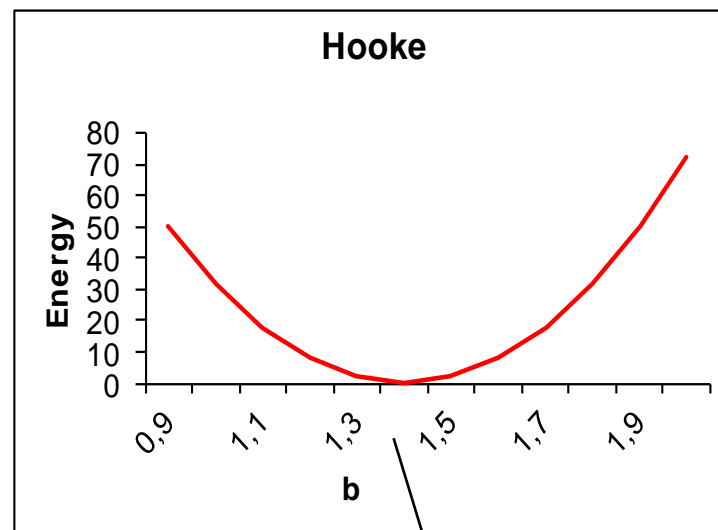
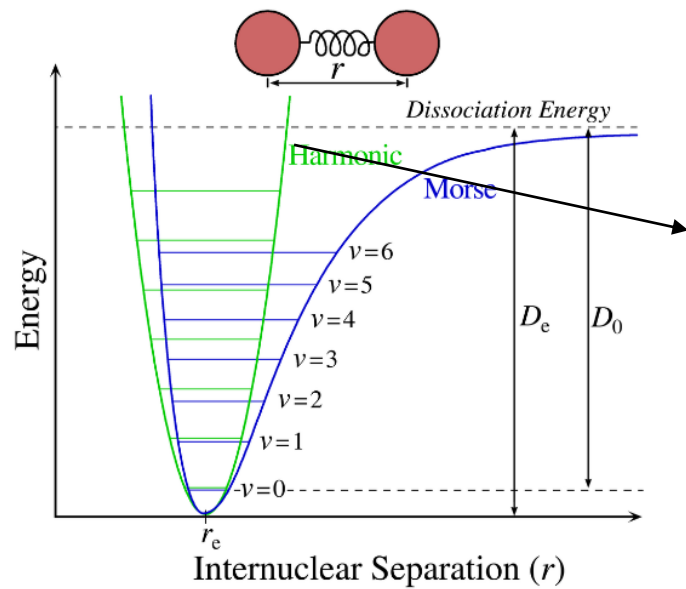
$$E_{pot} = \sum_{bonds} \frac{D_b}{2} (r - r_0)^2 + \sum_{angles} \frac{H_\theta}{2} (\theta - \theta_0)^2 + \sum_{dihedrals} \frac{V_N}{2} (1 + \cos(PN \varphi - PH)) + \\ + \sum_i \sum_j \varepsilon_{ij} \left[\left(\frac{A_{ij}}{r_{ij}} \right)^{12} - \left(\frac{C_{ij}}{r_{ij}} \right)^6 \right] + \sum_i \sum_j \frac{q_i q_j}{\varepsilon r_{ij}} + E_{other}$$

Van der Waals

Electrostatics

$$E = E_{str} + E_{bnd} + E_{tor} + E_{nb} + E_{other}$$


Non bonded-terms



$$E_{pot} = D_e (1 - e^{-a(r-r_0)})^2$$

$$E_{bond} = \sum_{bonds} \frac{D_b}{2} (r - r_0)^2 + \sum_{angles} \frac{H_\theta}{2} (\theta - \theta_0)^2 + \sum_{dihedrals} \frac{V_N}{2} (1 + \cos(PN \varphi - PH))$$

Force constant

r_{min}

Force constant

θ_{min}

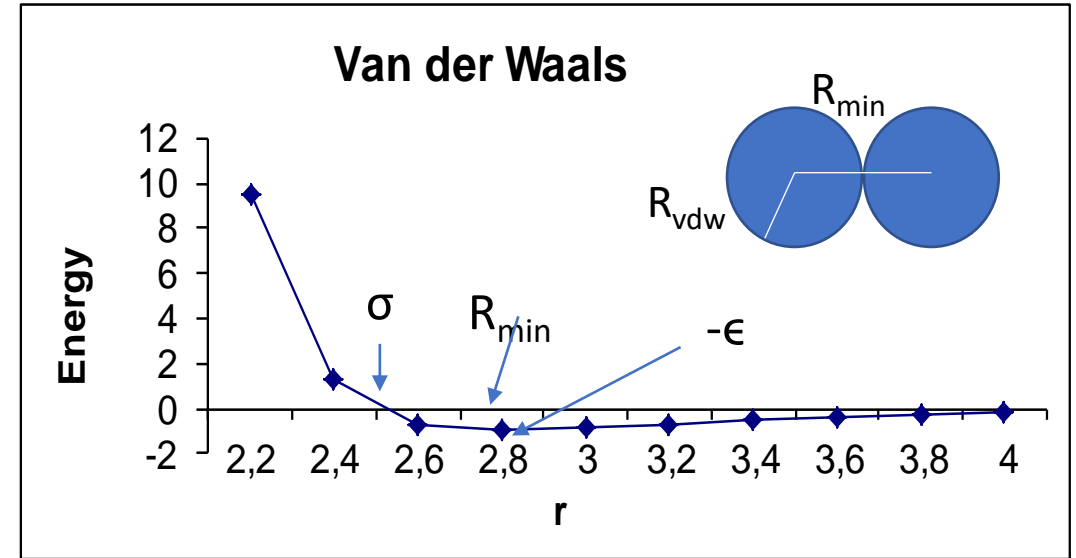
Energy Barrier

Periodicity

Phase

Van der Waals

- Empirical
- “Lennard-Jones” or “6-12” potentials
- Several alternatives



$$E_{vdw} = 4\epsilon \left(\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right) \quad E_{vdw} = \epsilon \left(\left(\frac{R_{min}}{r} \right)^{12} - 2 \left(\frac{R_{min}}{r} \right)^6 \right) \quad E_{vdw} = \frac{A}{r^{12}} - \frac{C}{r^6}$$

- Parameters for atom pairs obtained from combination of atomic ones

$$\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$$

$$\sigma_{ij} = \sqrt{\sigma_i \sigma_j}$$

$$\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$$

$$R_{min_{ij}} = R_{vdw_i} + R_{vdw_j}$$

$$A_{ij} = (\epsilon_i \epsilon_j)^{1/2} (R_{vdw_i}^{12} + R_{vdw_j}^{12})$$

$$C_{ij} = 2 \times (\epsilon_i \epsilon_j)^{1/2} (R_{vdw_i}^6 + R_{vdw_j}^6)$$

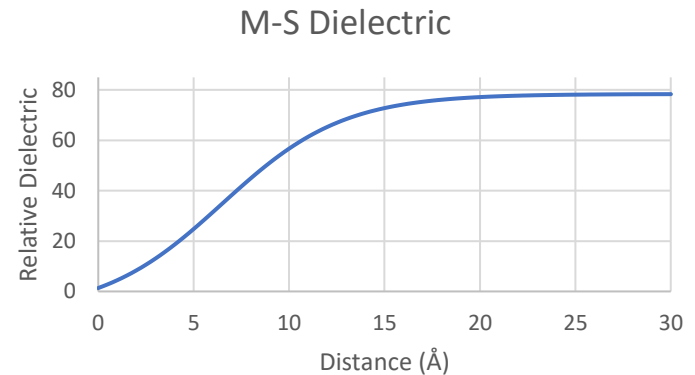
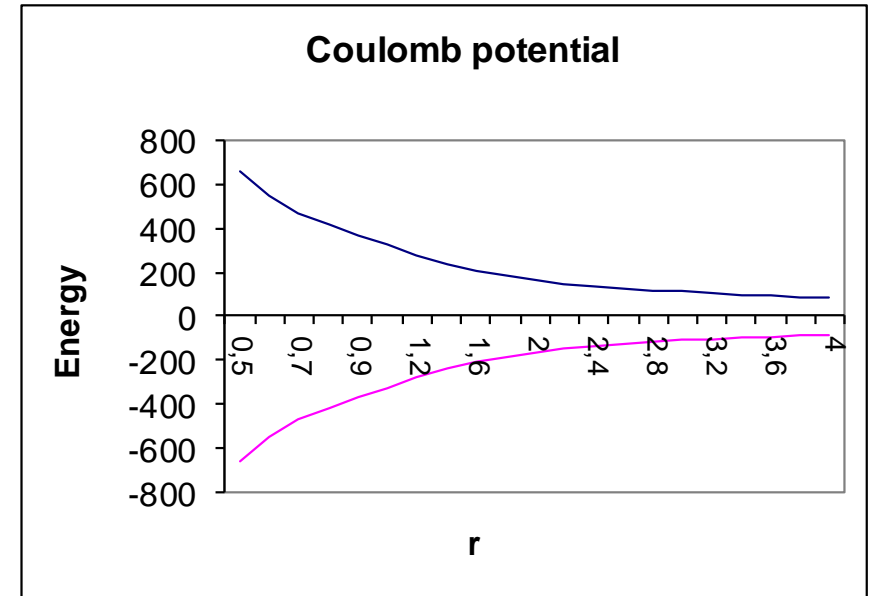
Electrostatics

- Stronger interaction
- All atoms have a partial charge
 - Dipolar interactions (directional)
 - Hydrogen bond energies
- Coulomb's Law

$$E_{elec} = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{\epsilon_r r} = 332.16 \frac{q_i q_j}{\epsilon_r r}$$

- q_i, q_j : Partial charges
 - ϵ_r : Relative dielectric constant
 - r : distance
- Dielectric strategies
 - Fixed value: 1 for vacuum, 80 for fully solvated
 - Linear: $\epsilon_r = k \cdot r$
 - Mehler-Solmajer

$$\epsilon_r = \frac{86.9525}{1 - 7.7839 e^{-0.3153 r}} - 8.5525$$



Examples of experimental restraints

- X-Ray refinement

$$E = E_{str} + E_{bnd} + E_{tor} + E_{nb} + E_{other}$$

Other restraints

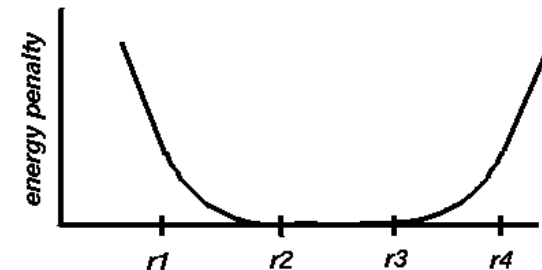


$$E_{X-Ray} = \sum_{hkl} (|F_o| - k|F_c|)^2$$

– F_o , F_c : Observed and calculated structure factors

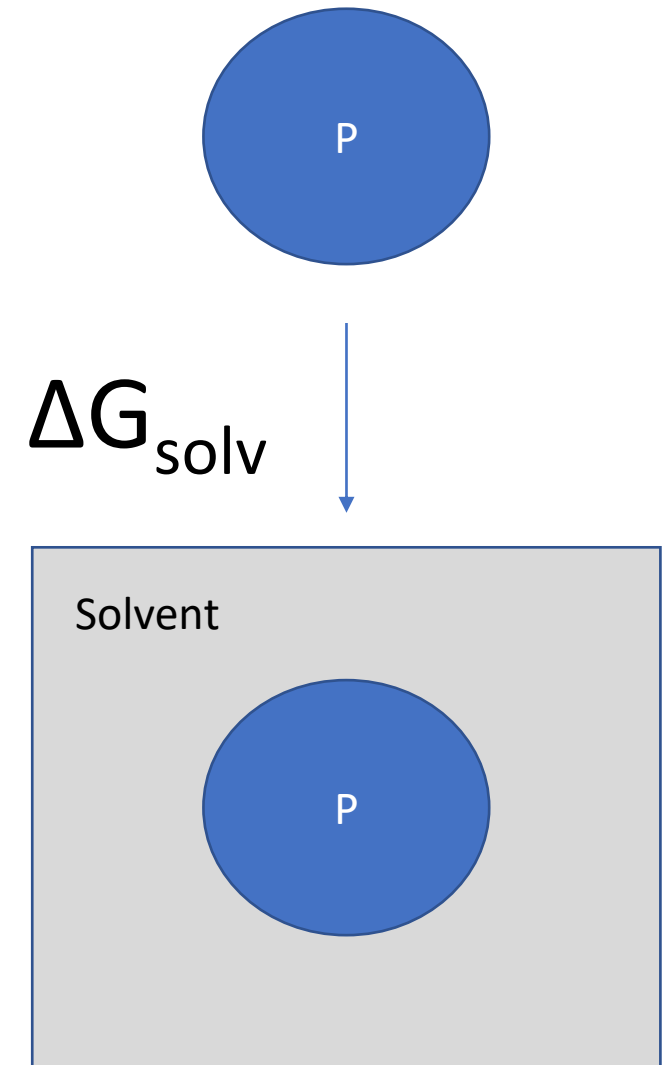
- NMR refinement

$$\begin{aligned} r_1 < r < r_2 &: k_2(r - r_2)^2 \\ r_2 < r < r_3 &: 0 \\ r_3 < r < r_4 &: k_3(r - r_3)^2 \end{aligned}$$



Solvation energy

- Defined as the energy to move a solute from vacuum to solvent
- Two major components
 - Electrostatic (Enthalpy)
 - Hydrophobic (Solvent entropy)
 - Entropy cannot be calculated in single system
- Key to understand macromolecular stability and binding



Approaches to the electrostatic component

- Explicit solvent
 - Solvent molecules are present in the system

$$E_{solv} = \sum_i^{solute} \sum_j^{solvent} \frac{q_i q_j}{\epsilon r_{ij}}$$

- However, solvent molecules reorient, a single value is meaningless, values should be **averaged over time**

$$E_{solv} = \left\langle \sum_i^{solute} \sum_j^{solvent} \frac{q_i q_j}{\epsilon r_{ij}} \right\rangle$$

Approaches to Electrostatic Solvation

- Implicit solvent. Classical continuous models

$$\nabla[\epsilon(r)\nabla\Phi(r)] = -4\pi\rho(r)$$

$$\rho = \rho_{\text{int}} + \rho_{\text{ext}}$$

$$\rho_{\text{ext}} = -\epsilon(\kappa^2)\sinh[\Phi(r)]$$

$$\kappa = \sqrt{\frac{8\pi N_a e^2 I}{10^3 \epsilon kT}}$$

$$\Delta G_{\text{solv}} = \Delta G_{\text{elec}}^{\epsilon=\text{sol}} - \Delta G_{\text{elec}}^{\epsilon=\text{vac}}$$

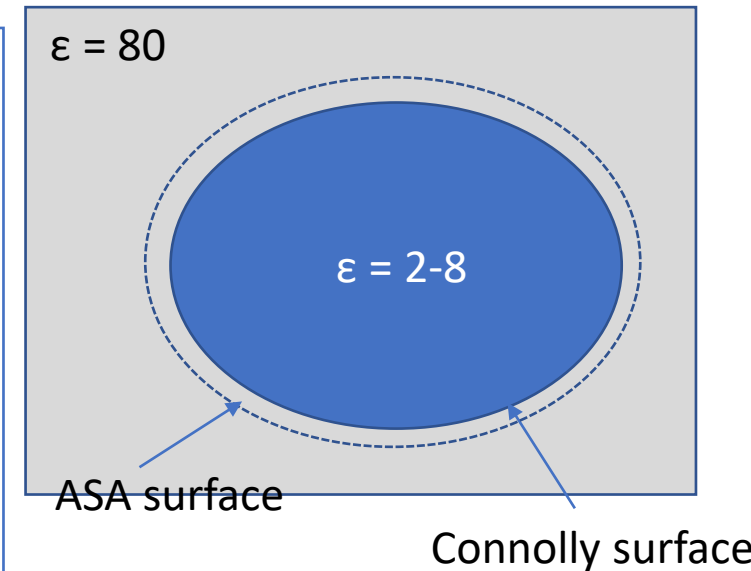
Poisson-Boltzmann

$$\Delta G_{\text{solv}} = \frac{1}{8\pi\epsilon_0} \left(1 - 1/\epsilon\right) \sum_{i,j} \frac{q_i q_j}{f_{GB}}$$

$$f_{GB} = \sqrt{r_{ij}^2 + a_{ij}^2} e^{-D}$$

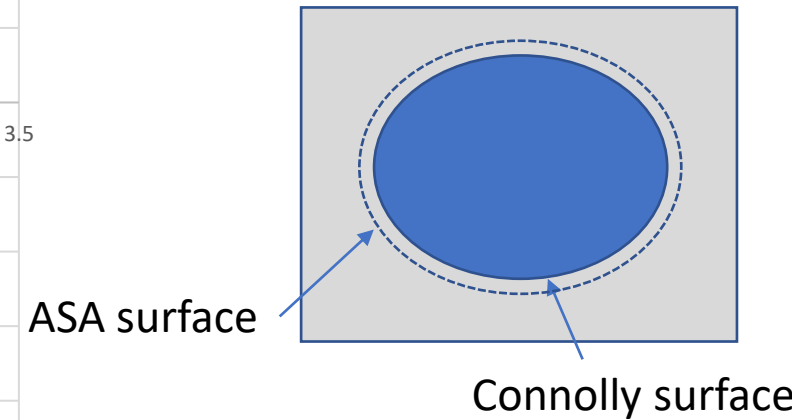
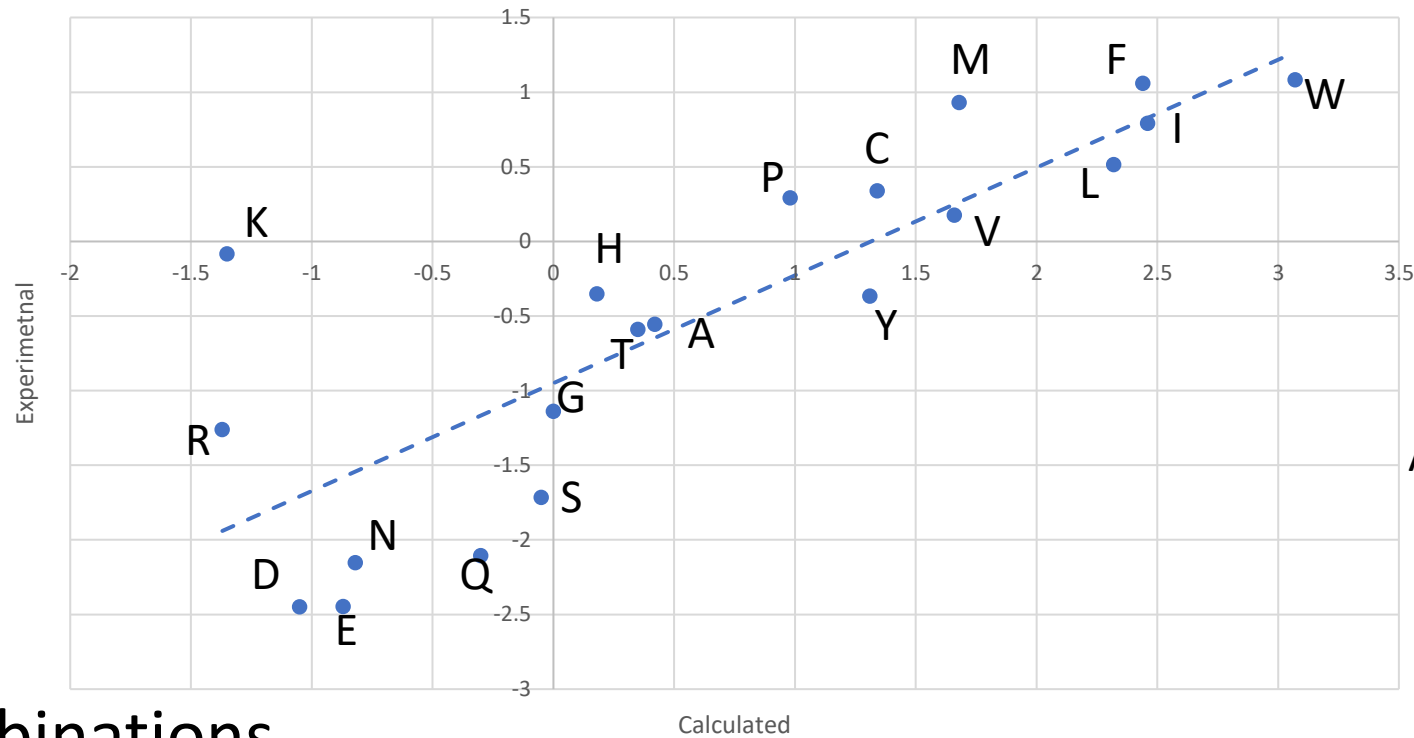
$$D = \left(\frac{r_{ij}}{2a_{ij}}\right)^2 \quad a_{ij} = \sqrt{a_i a_j}$$

Generalized Born



Approaches to the hydrophobic component

- **Experimental** ΔG_{solv} is (almost) linearly correlated with ASA Surface



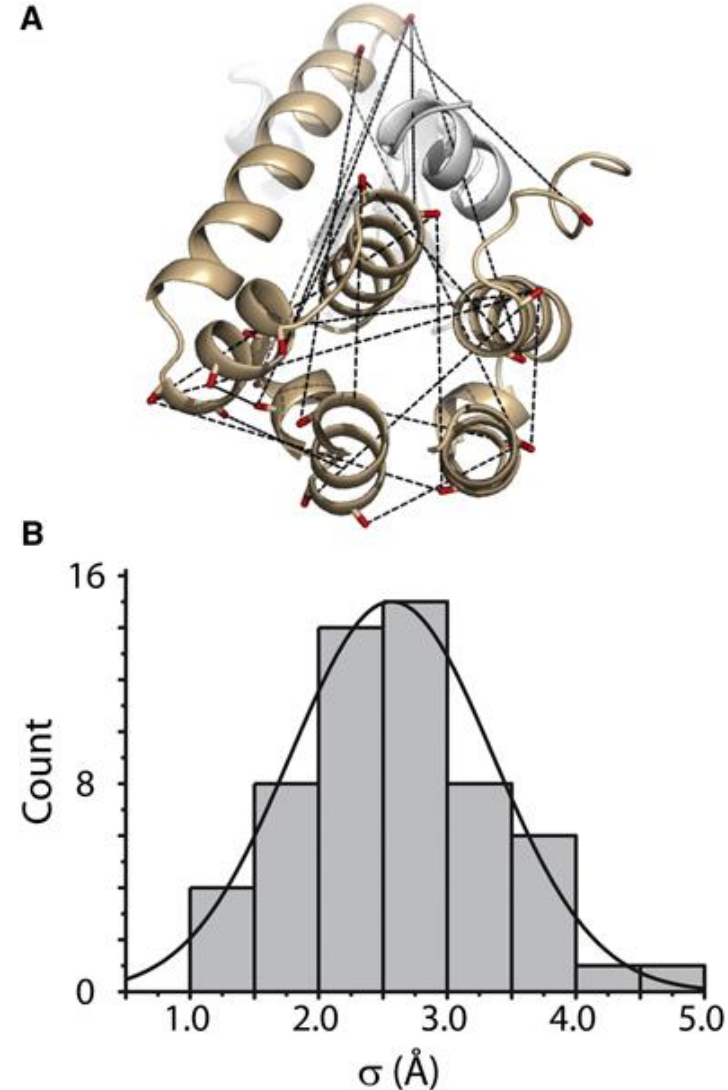
- Combinations
 - PB-SA, GB-SA

$$\Delta G_{solv} = \sum_i \sigma_i ASA_i$$

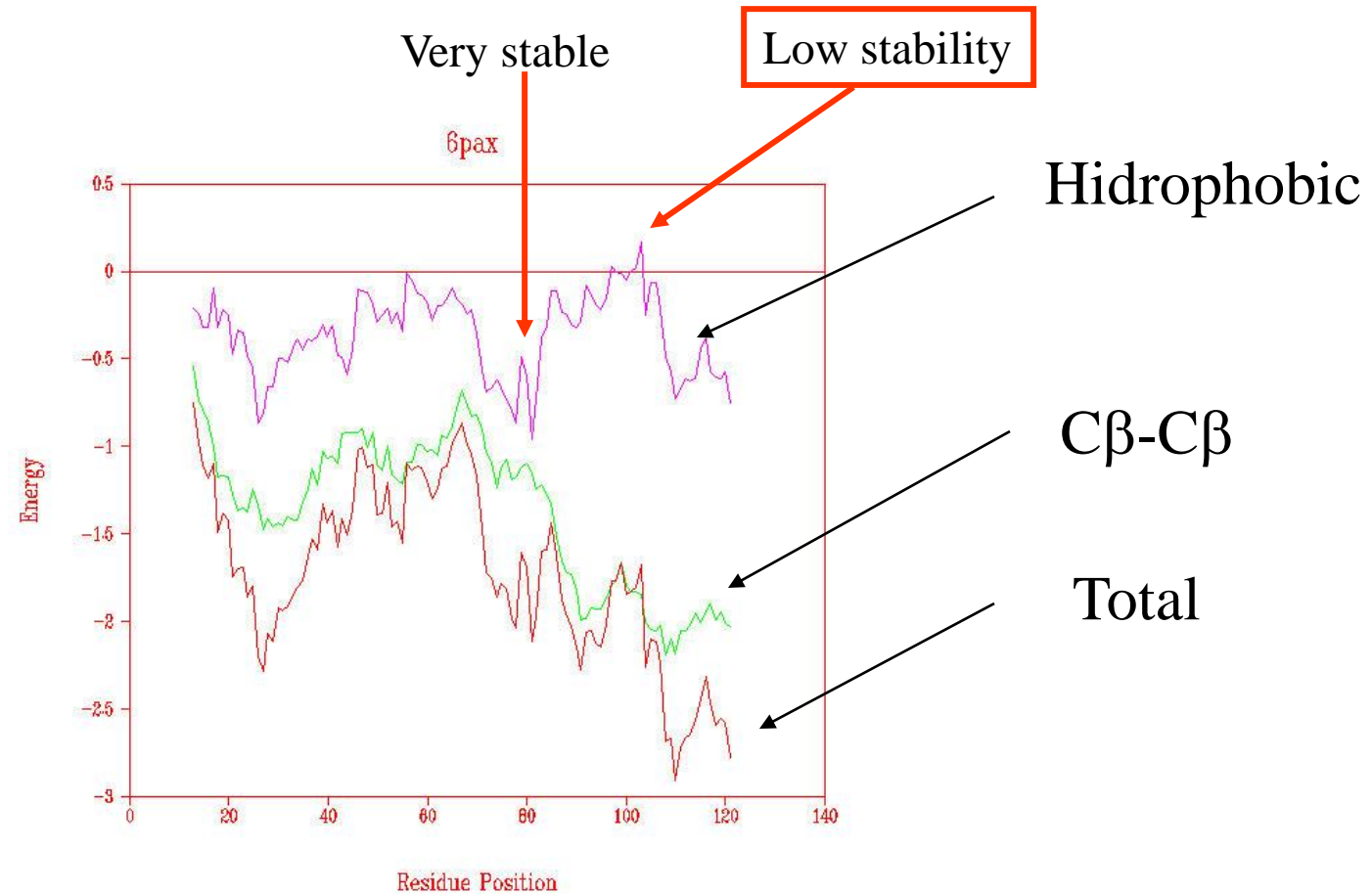
Statistical potentials

- Aims to reproduce known structures
 - Based on position and distances of aminoacid residues
- Can include preferences or other structure features
- Used to evaluate stability of 3D structures (experimental or modelled)
- Required on simplified Systems
- Energies are obtained from Boltzmann Distribution:

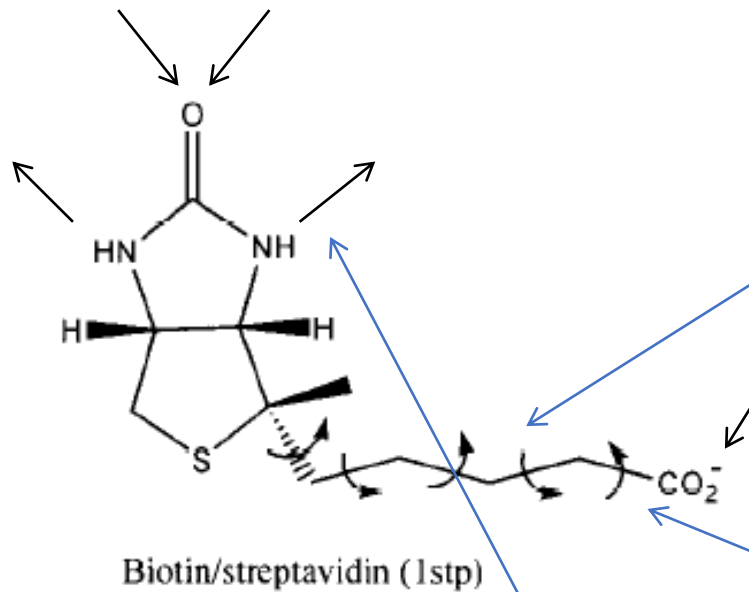
$$E = -RT \ln p$$



PROSA Example



Empiric scores



$$\Delta G = \Delta G_0 + \Delta G_{\text{rot}} \times N_{\text{rot}} \quad (1)$$

$$+ \Delta G_{\text{hb}} \sum_{\text{neutral H-bonds}} f(\Delta R, \Delta \alpha) \quad (2)$$

$$+ \Delta G_{\text{io}} \sum_{\text{ionic int.}} f(\Delta R, \Delta \alpha) \quad (3)$$

$$+ \Delta G_{\text{aro}} \sum_{\text{aro int.}} f(\Delta R, \Delta \alpha) \quad (4)$$

$$+ \Delta G_{\text{lipo}} \sum_{\text{lipo. cont.}} f^*(\Delta R) \quad (5)$$

FlexX

$$\text{HPScore} = C_{0,1} + C_{\text{VDW},1}(\text{VDW}) + C_{\text{HB},1}(\text{H-Bond}) + C_{\text{HP}}(\text{Hydrophobic Pair}) + C_{\text{RT},1}(\text{Rotor})$$

$$\text{HMScore} = C_{0,2} + C_{\text{VDW},2}(\text{VDW}) + C_{\text{HB},2}(\text{H-Bond}) + C_{\text{HM}}(\text{Hydrophobic Match}) + C_{\text{RT},2}(\text{Rotor})$$

$$\text{HSScore} = C_{0,3} + C_{\text{VDW},3}(\text{VDW}) + C_{\text{HB},3}(\text{H-Bond}) + C_{\text{HS}}(\text{Hydrophobic surface}) + C_{\text{RT},3}(\text{Rotor})$$

Xscore (Empirical fit)

Which level and when?

- Quantum Mechanics
 - Very accurate energies
 - Chemical processes
 - Distorted structures
- Classical Mechanics
 - Near-equilibrium structures
 - Conformational Changes, Folding
 - Refinement of experimental structures
- Statistical/knowledge based
 - Non natural (simplified) representations
 - Scoring based on similarity
 - Uses “non-energy” terms
- Empiric scores
 - Closely related series of compounds