Molecular Modelling

part of “Bioinformatik von RNA- und Proteinstrukturen”

Sonja Prohaska

Computational EvoDevo
University Leipzig

Leipzig, SS 2011

Protein Structure

- levels or organization
  - **Primary structure**: sequence of amino acids (from N- to C-terminus)
  - **Secondary structure**: β-sheets and α-helices
  - **Ternary structure**: arrangement of secondary structure elements into folds
  - **Quaternary structure**: assoziation of polypeptide changes into one protein
- proteins can be unfolded and refolded
- conformation is assumed to be the thermodynamically prefered one
- mathematical models should allow to predict protein structure from the protein sequence
- $\rightarrow$ **Molecular Modelling**

Anfinsen's dogma and Levinthal's paradox

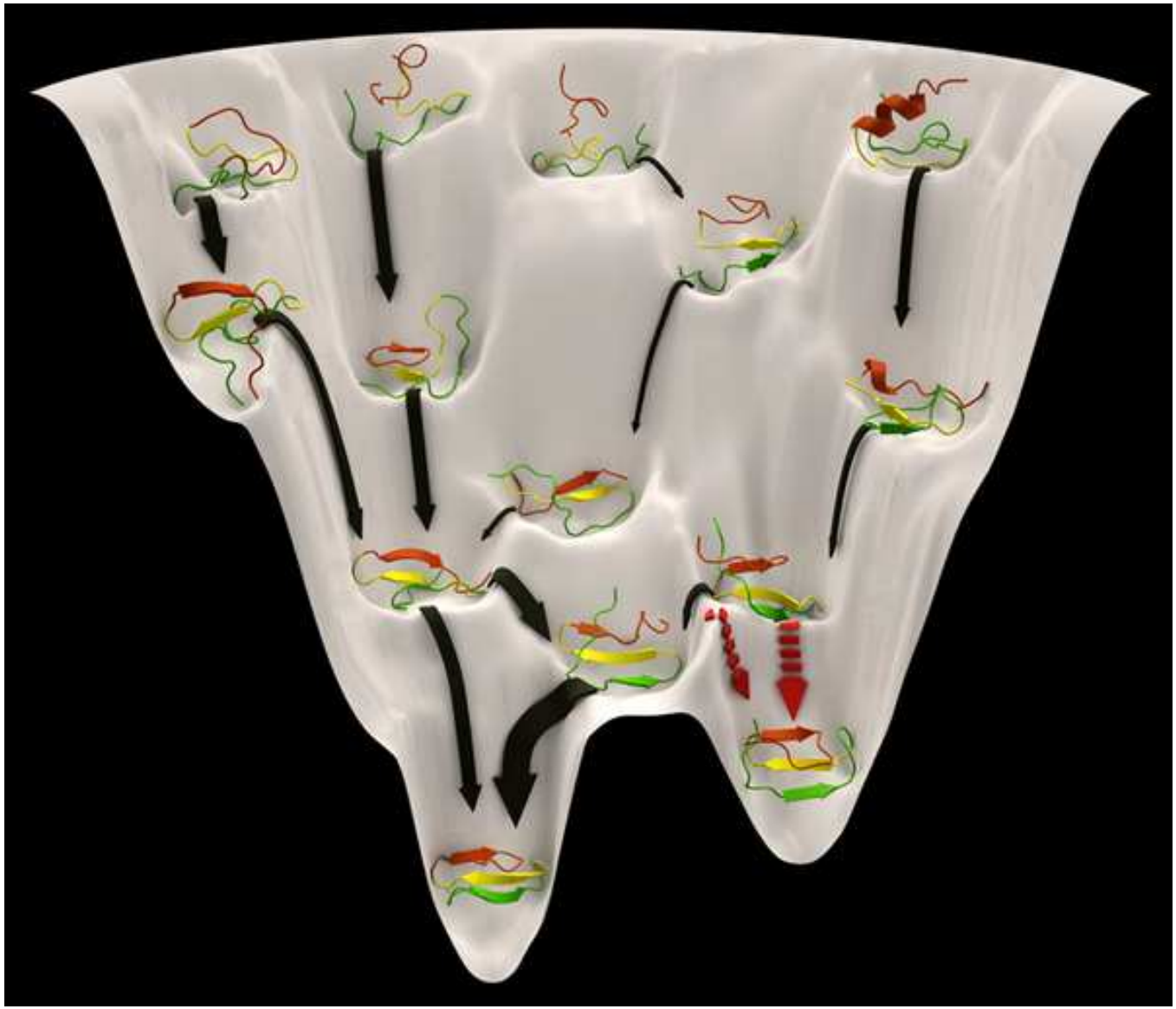
Anfinsen's dogma

- The native structure of a protein is determined only by the protein's amino acid sequence.
- The native structure is a **unique**, **stable** and **kinetically accessible** minimum of the free energy.

Levinthal's paradox

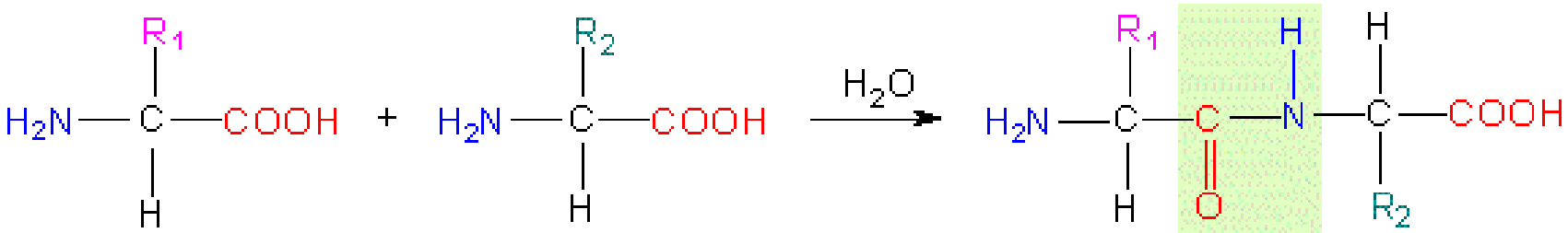
- The number of possible **conformations available** to a given protein is **astronomically large**.
- A polypeptide chain of 100 residues will have 99 peptide bonds, and therefore 198 different phi and psi bond angles. If each of these bond angles can be in one of three stable conformations, the protein may fold into a maximum of 3^{198} different conformations.

Folding Funnel



Primary Structure

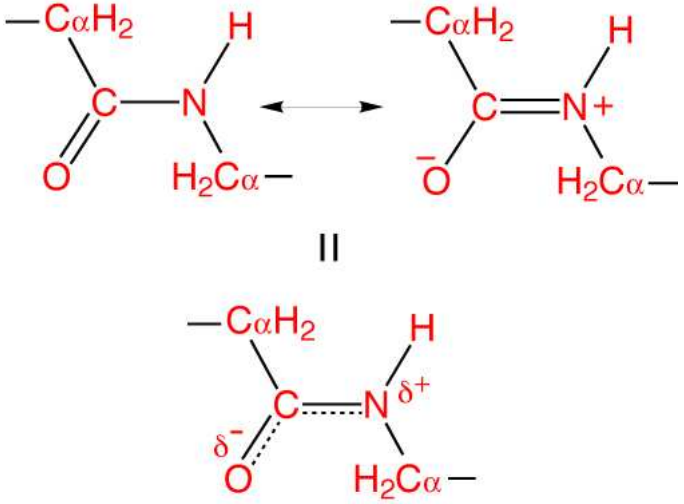
A polypeptide is a single linear **polymer chain** of **amino acids** bonded together by **peptide bonds** between the carboxyl and amino groups of adjacent amino acid residues.



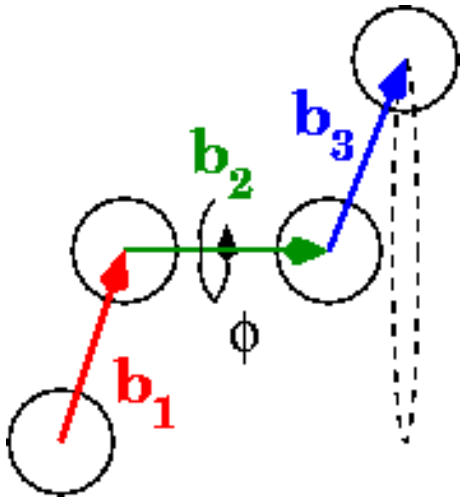
During the condensation reaction a water molecule is released per formed peptide bond.

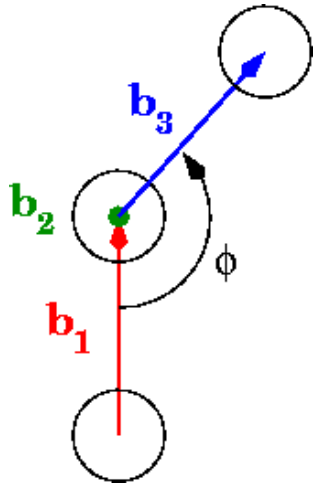
Polypeptide chain characteristics

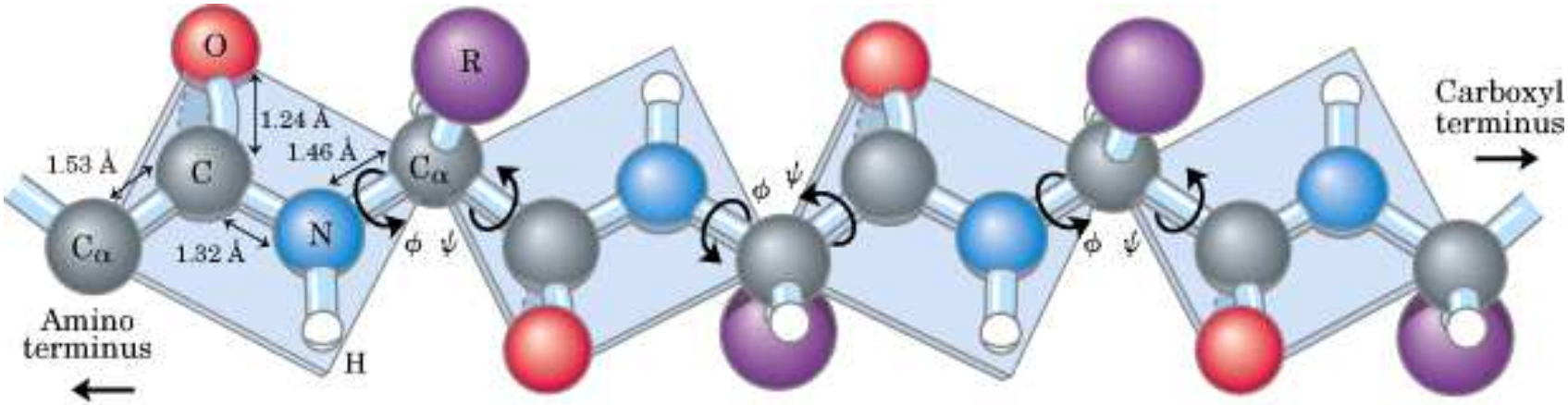
The peptide bond is **planar** due to its **double bond character**.



Loren Williams, Georgia Tech







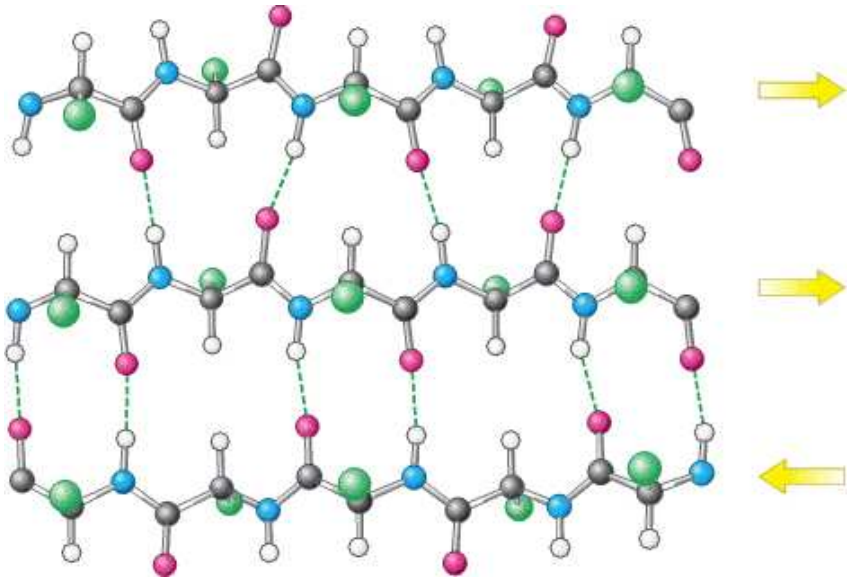
Polypeptide chain characteristics

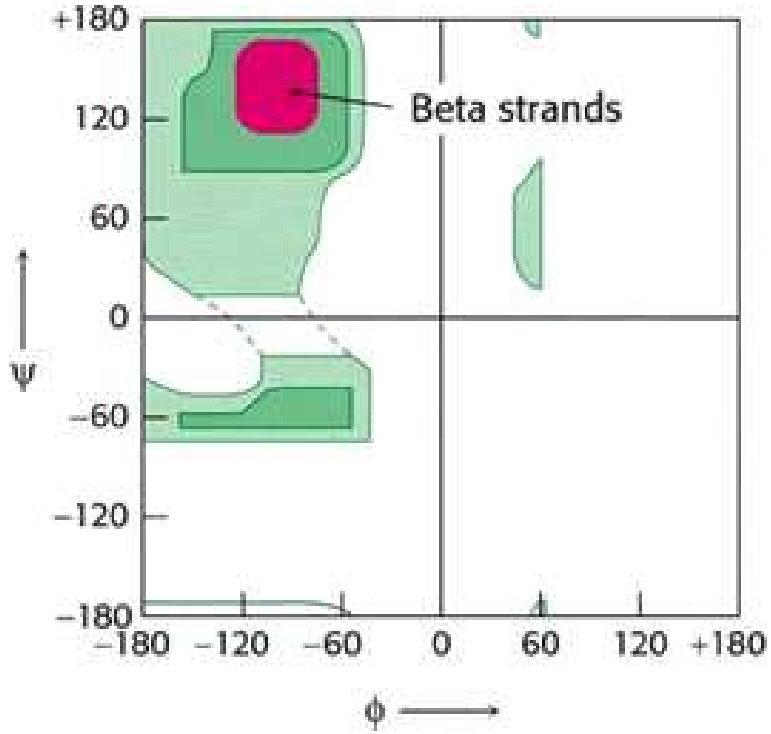
	dihedral/torsion angles		
bond	$N - C_\alpha$	$C_\alpha - C_1$	$C_1 - N$
angle	ϕ (*phi*)	ψ (*psi*)	ω (*omega*)
cis			$0°$
trans			$180°$
right-handed α-helix	$-57°$	$-47°$	
antiparallel β-sheet	$-139°$	$+135°$	
parallel β-sheet	$-119°$	$+113°$	

Possible combinations of ϕ and ψ values are visualized in **Ramachandran plots** (see next slides).

Secondary Structure Elements I

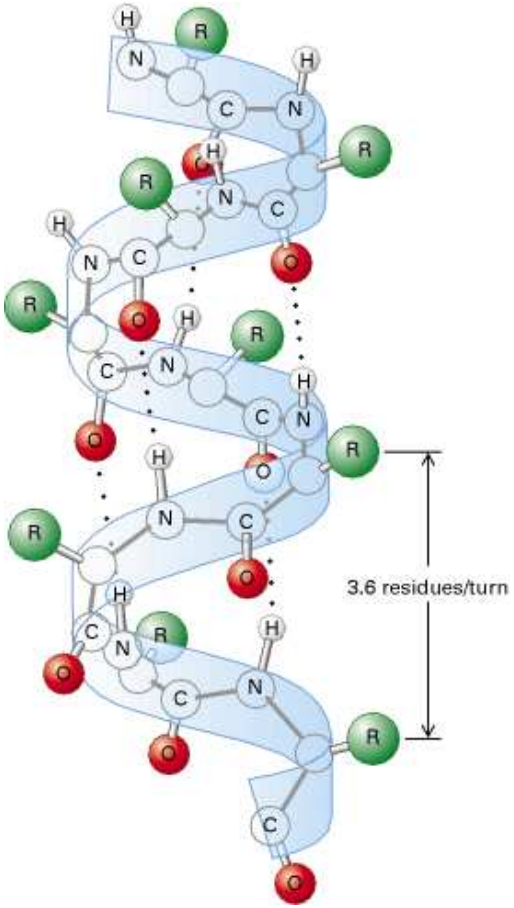
Parallel and anti-parallel β-strands can form **β-sheets**.

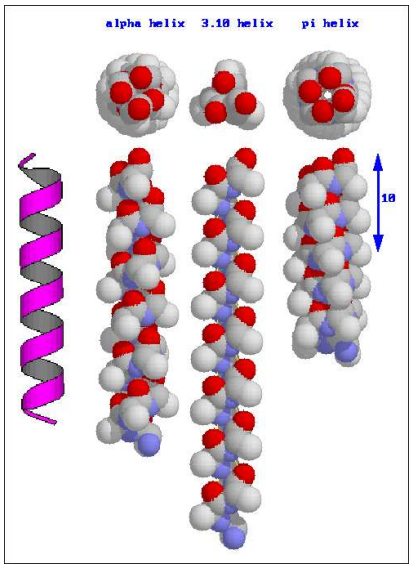


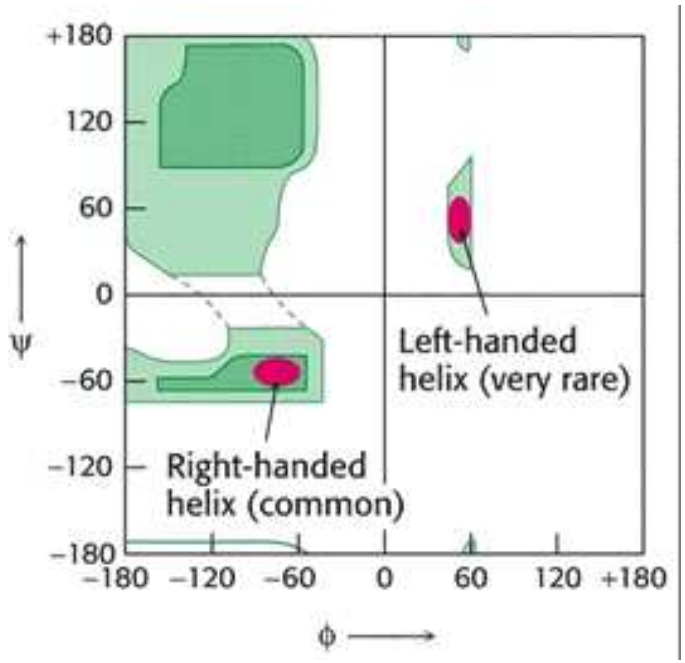


Secondary Structure Elements II

Hydrogen bonds between every N-H group and the oxygen of a C=O group four amino acids down the chain hold the α**-helices** in shape.







Molecular Modelling

- atoms are point charges with an associated mass
- collective mathematical expression is the potential function
- subsumes energetic contributions from multiple physical forces
- and between (all) pairs of atoms (“pair potential”)
- reprecented as force field
- **molecular mechanics** minimizes the (static) potential energy
- **molecular dynamics** models the behaviour of the system over time

Forces that determine protein structure

- bonded energy contribution
  - bond lengths
  - bond angles
  - torsion angles
- non-bonded energy contribution
  - van der Waals forces
  - electrostatic interactions
- conjugated system distortion
- hydrogen bonds
- solvent

$$E^{total} = E^{bond} + E^{angle} + E^{torsion} + E^{Waals} + E^{coulomb}$$

Bond length, angle and torsion

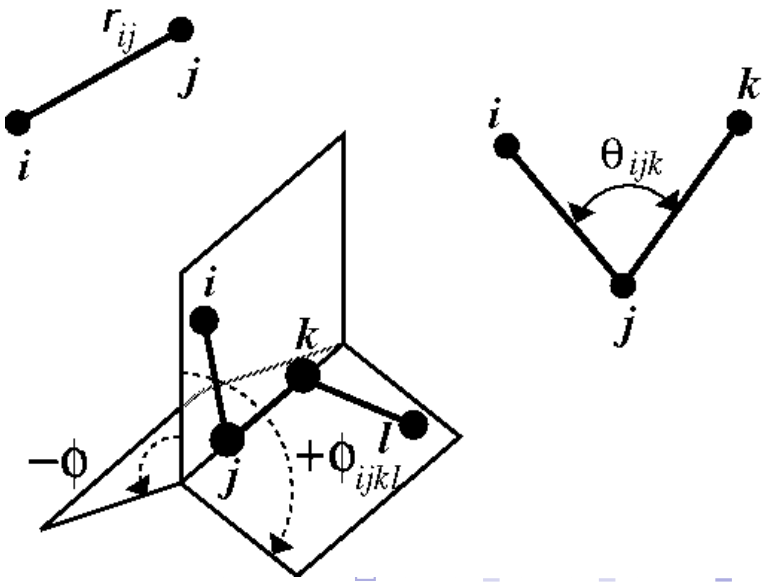
Penalty for molecule geometry deviating from optimal.

$$E_{ij}^{bond} = \frac{1}{2} k_{ij}^{bond} (r_{ij} - r_{ij}^0)^2 \quad (1)$$

$$E_{ijk}^{angle} = \frac{1}{2} k_{ijk}^{angle} (\theta_{ijk} - \theta_{ijk}^0)^2 \quad (2)$$

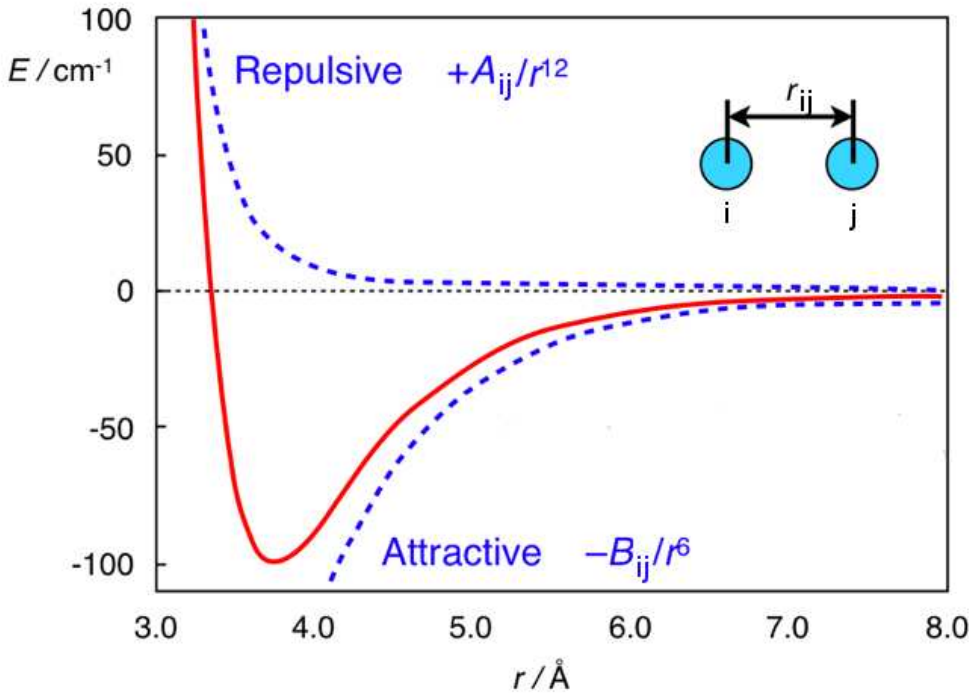
$$E_{ijkl}^{torsion} = \frac{1}{2} V_n (1 + \cos n\phi_{ijkl}) \quad (3)$$

r_{ij}^0 ... optimal bond length
θ_{ijk}^0 ... optimal valence angle
V_n ... height of the torsional barrier
ϕ_{ijkl} ... torional angle
k ... constants



van der Waals forces

The **Lennard-Jones potential** is a simple model to approximates the interaction between a pair of neutral atoms.



$$E_{ij}^{vanderWaals} = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \quad (4)$$

electrostatic interactions

Coulomb's law describes the electrostatic interaction between electrically charged atoms. The dipole interactions of amide groups along the protein backbone has a significant influence on the overall conformation.

$$E_{ij}^{coulomb} = \frac{1}{k_{ij}^{coul}} \frac{q_i q_j}{r_{ij}} \tag{5}$$

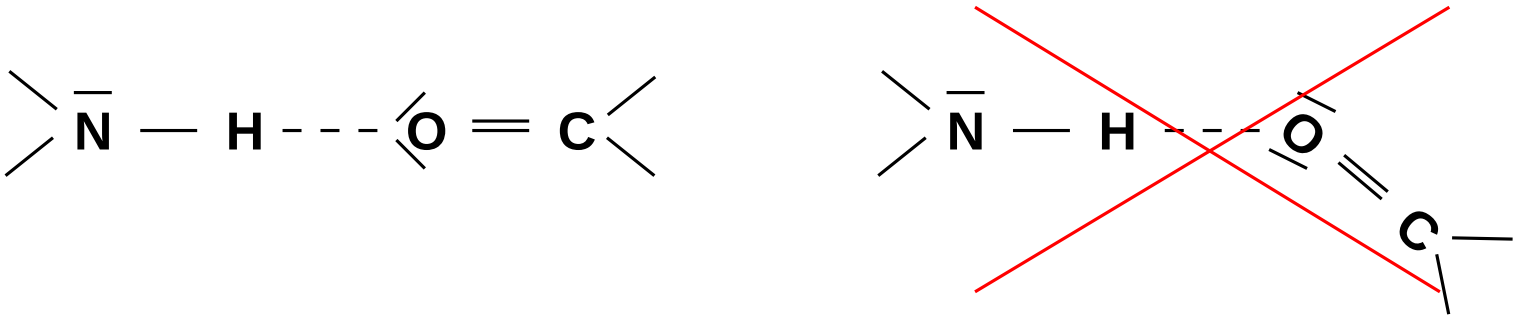
r_{ij}	...	distance between atoms i and j
q_i ,q_j	...	partial changes on atoms i and j
k_{ij}^{coul}	...	constant

AMBER Force Field

Assisted Model Building with Energy Refinement (AMBER) – represents a functional form of force fields as well as a molecular dynamics software package.

$$E^{total} = \sum E_{ij}^{bond} + \sum E_{ijk}^{angle} + \sum E_{ijkl}^{torsion} + \sum E_{ij}^{vanderWaals} + \sum E_{ij}^{coulor}$$

Energy terms for Hydrogen bonds



$$E_{ij}^{hydro} = \frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \quad (6)$$

The equation takes only the distance of i (H on H-Donor) and j (O on H-Aceptor) into account but not the angle.

Literature

Biophysical Chemistry, Part I: The conformation of biological macromolecules by Charles R. Cantor and Paul R. Schimmel; W. H. Freeman and Company, 12. edition, 2002; Chapter 5: “Conformational analysis and forces that determine protein structure”, p252-309.