COARSE-GRAINED SIMULATIONS OF INTERACTING PROTEINS WITH DISCRETE MOLECULAR DYNAMICS

Agustí Emperador

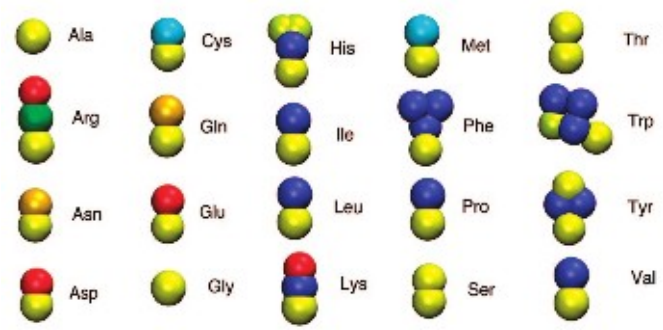
Institute for Research in Biomedicine, Barcelona

COARSE-GRAINING THE STRUCTURE

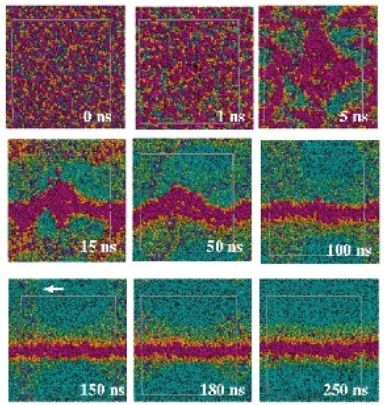
Mapping used by Marrink + atomistic representation of backbone

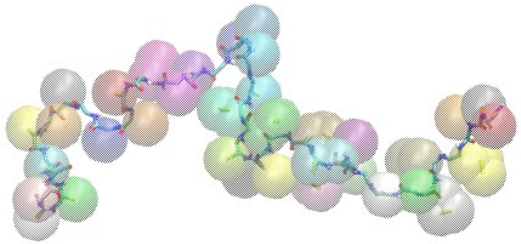
Marrink's model gave excellent results for lipid-lipid and lipid-protein interactions (BUT it is an **explicit solvent model** and does not consider *hydrogen bonding*)

Coarse-grained beads (particles) for the sidechains



Monticelli et al, *JCTC* 4, 819 (2008)





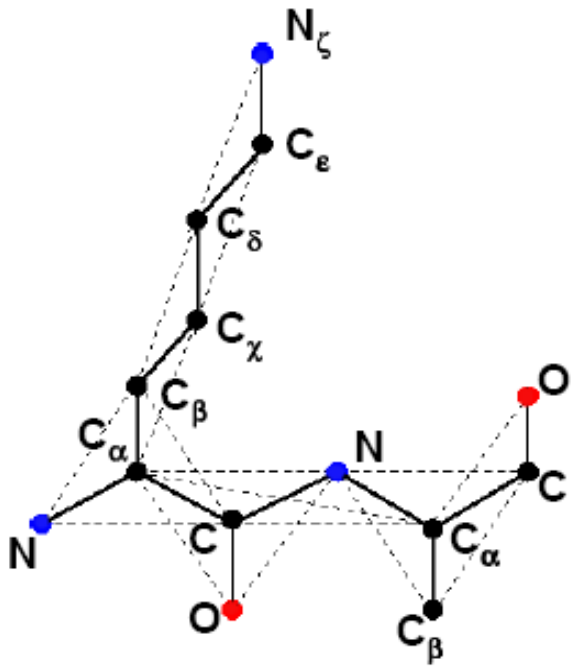
Our force field: **Van der Waals, electrostatic** and **implicit solvation** interaction between sidechain and C_α beads

Hydrogen bonding *only* between backbone atoms (*backbone not coarse-grained*)

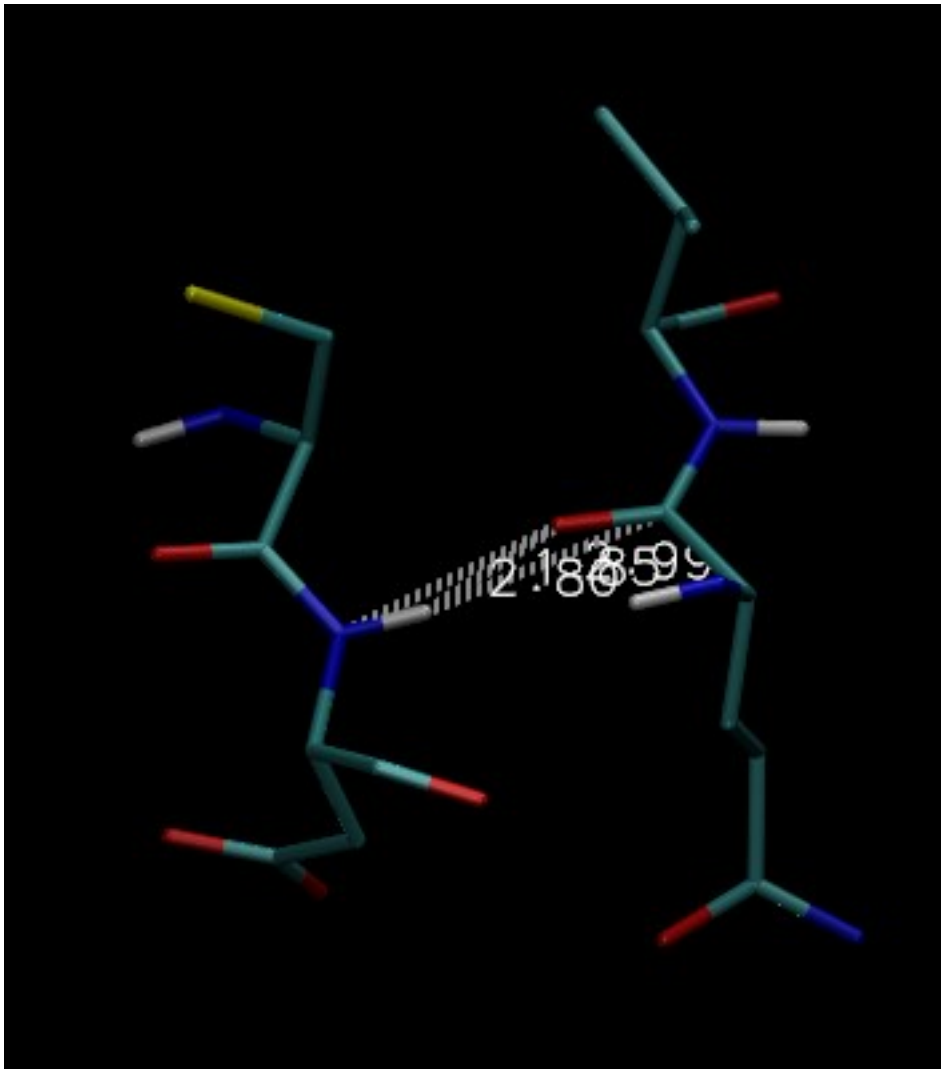
Backbone hydrogen bonding modeled like in atomistic Discrete Molecular Dynamics (DMD)

MOLECULAR STRUCTURE IN DMD

In DMD all the potentials depend on *particle-particle distances* (two-body potentials) It is not possible to define forces depending on angles or dihedrals. **Pseudobonds** have to be used to fix bond angles and dihedrals

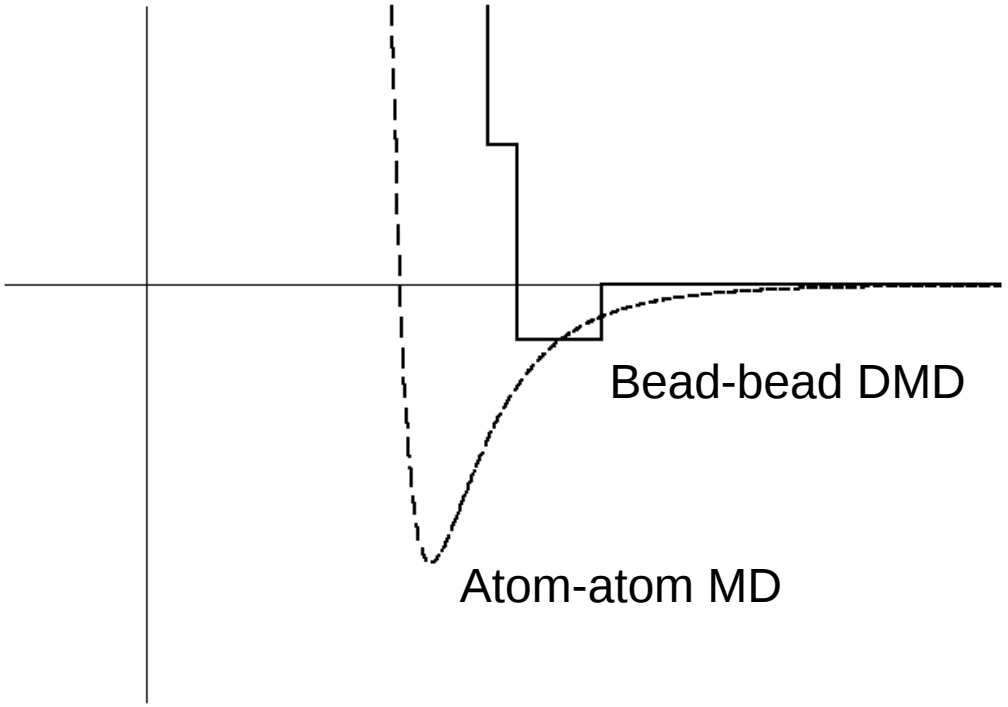


LYS-ALA



Hydrogen bonds are dipole-dipole interactions, therefore dependent on the angle. Pseudobonds are used to keep the correct angle.

INTERACTION POTENTIALS IN DISCRETE MOLECULAR DYNAMICS



DISCRETE MOLECULAR DYNAMICS

Discretization of the continuous potential

How much accuracy is lost?

Much less than when using **implicit solvent** and **coarse-graining** the structure

Less steps → faster

The **well depth** is the value of the continuous potential at $R_{AB} = R_A + R_B$

Standard molecular dynamics (MD)

Integrate the equations of motion
at each timestep (Δt = 2 fs)

$$F = -\frac{dV}{dr} = ma$$

$$a = \frac{dv}{dt}$$

$$v = \frac{dr}{dt}$$

500.000.000 timesteps needed
to generate 1 μs of trajectory!

$$V = E_{bonded} + E_{non-bonded}$$

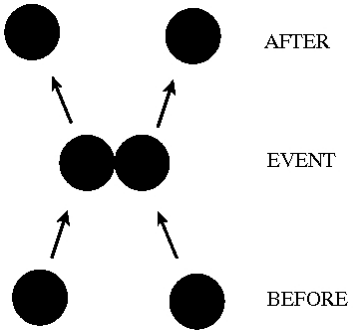
$$E_{non-bonded} = \sum_{a,b} \frac{Q_a Q_b}{r_{ab}} + \sum_{a,b} \left(\frac{C_{ab}}{r_{ab}} \right)^{12} - \left(\frac{D_{ab}}{r_{ab}} \right)^{6}$$

Discrete molecular dynamics (DMD)

Particles move with constant velocity…

$$r_i(t + t_c) = r_i(t) + v_i(t) t_c$$

…until a collision occurs



Transfer of linear momentum
upon a collision

$$m_i v_i = m_i v_i' + \Delta p$$
$$m_j v_j + \Delta p = m_j v_j'$$

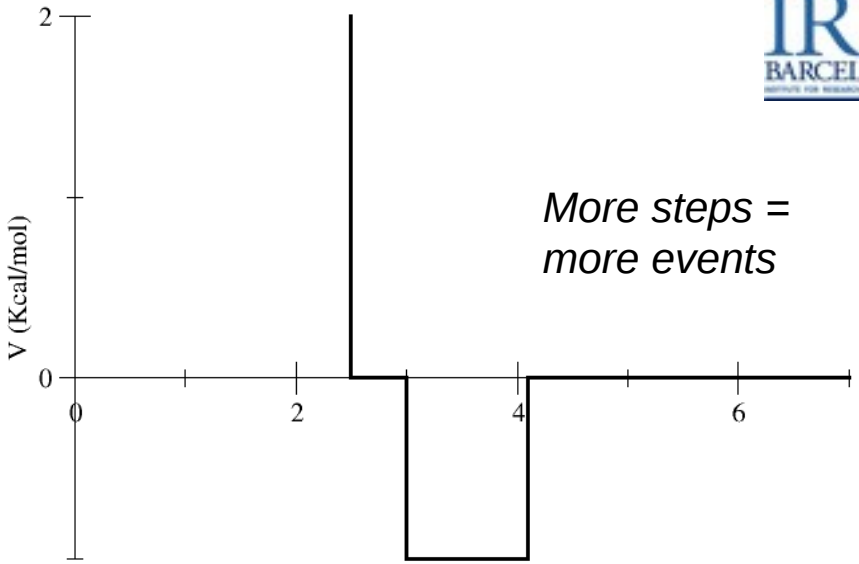
The velocities of particles i, j change

Discrete molecular dynamics

Conservation of **linear momentum**
Conservation of **energy** when entering a region with different potential energy
(Emperador et al, *Proteins* **78** 83 (2010))

$$m_i v_i + m_j v_j = m_i v_i' + m_j v_j'$$

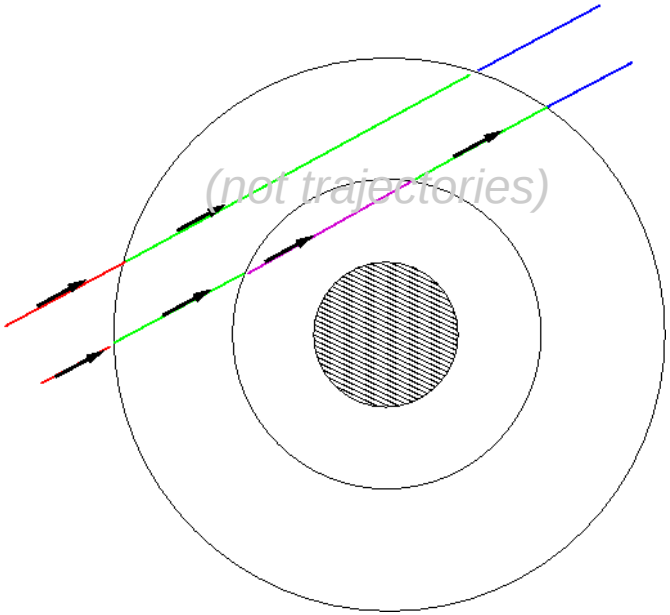
$$\frac{1}{2} m_i v_{i^2} + \frac{1}{2} m_j v_{j^2} = \frac{1}{2} m_i v_i'^2 + \frac{1}{2} m_j v_j'^2 + \Delta V$$



More steps = more events

Transferred linear momentum

$$\Delta p = \frac{m_i m_j}{m_i + m_j} \left\{ \sqrt{\left(v_j - v_i\right)^2 - 2\frac{m_i + m_j}{m_i m_j}\Delta V} - \left(v_j - v_i\right) \right\}$$

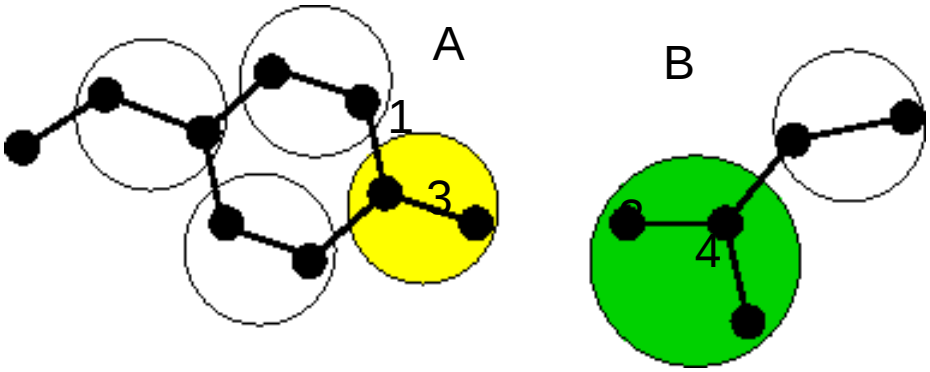


If ΔV>0, the particles can overcome the potential step as long as $\Delta V < \frac{m_1 m_2}{2\left(m_1 + m_2\right)}\left(v_j - v_i\right)^2$

Otherwise, the particles keep inside the well and $\Delta p = \frac{2 m_i m_j}{m_i + m_j}\left(v_i - v_j\right)$

CONSTRUCTION OF THE FORCE FIELD

The *depth* of the square well is the value of the interaction potential at $R_{AB} = R_A + R_B$



Pairwise **additive** atomic interactions
Collapse all atoms in the center of the bead

Value of the interaction potential at $R_{AB} = R_A + R_B$?
Interaction potential between beads A and B:

$$V_{AB} = V_{12} + V_{14} + V_{23} + V_{34}$$

Approximation: $r_{12} = r_{14} = r_{23} = r_{34} = R_{AB}$

Electrostatic potential

$$V_{AB} = \frac{q_1q_2 + q_1q_4 + q_2q_3 + q_3q_4}{4\pi\epsilon_0\epsilon_r R_{AB}} = \frac{(q_1 + q_3)(q_2 + q_4)}{4\pi\epsilon_0\epsilon_r R_{AB}} = \frac{Q_A Q_B}{4\pi\epsilon_0\epsilon_r R_{AB}}$$

Van der Waals potential

$$V_{AB} = (\epsilon_{12} + \epsilon_{14} + \epsilon_{23} + \epsilon_{34}) \left[\left(\frac{2}{N_A^{1/3} + N_B^{1/3}} \right)^{12} - 2 \left(\frac{2}{N_A^{1/3} + N_B^{1/3}} \right)^{6} \right]$$

$$V_{AB}^{VdW} = \sqrt{\epsilon_A \epsilon_B} \left[\left(\frac{2}{N_A^{1/3} + N_B^{1/3}} \right)^{12} - 2 \left(\frac{2}{N_A^{1/3} + N_B^{1/3}} \right)^{6} \right]$$

where $\epsilon_A = \epsilon_1 + \epsilon_3$, $\epsilon_B = \epsilon_2 + \epsilon_4$.

Solvation potential

EEF1 implicit solvation (Lazaridis & Karplus, *Proteins* **35** 133 (1999)) depending on the volume v_i and the solvation energy density $f_i \propto \Delta G_i$

$$V_{12} = -\int_{v_1} f_2 d\vec{r} - \int_{v_2} f_1 d\vec{r} \approx -f_1(r)v_2 - f_2(r)v_1$$

$$V_{AB}^{solv} \approx -C(\Delta G_1 v_2 + \Delta G_1 v_4 + \Delta G_3 v_2 + \Delta G_3 v_4 + \Delta G_2 v_1 + \Delta G_4 v_1 + \Delta G_2 v_3 + \Delta G_4 v_3)$$

$$V_{AB}^{solv} \approx -C[(\Delta G_1 + \Delta G_3)(v_2 + v_4) + (\Delta G_2 + \Delta G_4)(v_1 + v_3)]$$

$$V_{AB}^{solv} \approx -C(\Delta G_A v_B + \Delta G_B v_A)$$

TABLE I. Solvation Parameters[a]

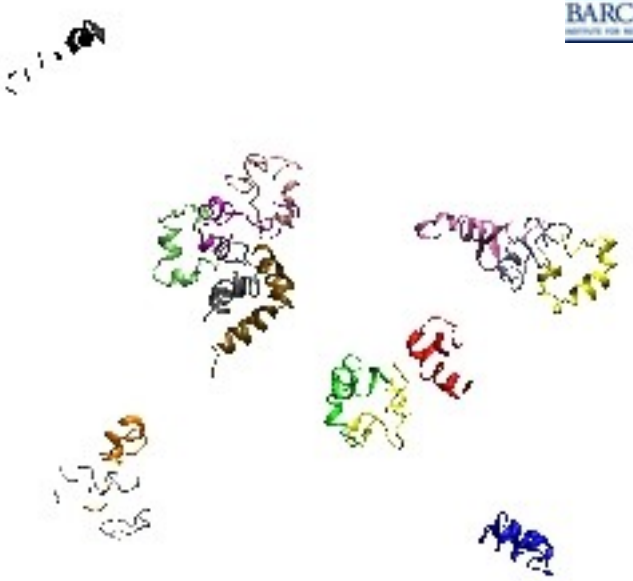
Atom types[a]	Volume	ΔG_i^{ref} [b]	ΔG_i^{free}	ΔH_i^{ref} [b]	ΔCp_i^{ref} [d]
C	14.7	0.000	0.00	0.000	0.00
CR	8.3	−0.890	−1.40	2.220	6.90
CH1E	23.7	−0.187	−0.25	0.876	0.00
CH2E	22.4	0.372	0.52	−0.610	18.60
CH3E	30.0	1.089	1.50	−1.779	35.60
CR1E	18.4	0.057	0.08	−0.973	6.90
NH1	4.4	−5.950	−8.90	−9.059	−8.80
NR	4.4	−3.820	−4.00	−4.654	−8.80
NH2	11.2	−5.450	−7.80	−9.028	−7.00
NH3	11.2	−20.000	−20.00	−25.000	−18.00
NC2	11.2	−10.000	−10.00	−12.000	−7.00
N	0.0	−1.000	−1.55	−1.250	8.80
OH1	10.8	−5.920	−6.70	−9.264	−11.20
O	10.8	−5.330	−5.85	−5.787	−8.80
OC	10.8	−10.000	−10.00	−12.000	−9.40
S	14.7	−3.240	−4.10	−4.475	−39.90
SH1E	21.4	−2.050	−2.70	−4.475	−39.90

PARAMETRIZATION OF THE COARSE-GRAINED FORCE FIELD

The total **interaction potential** between nonbonded coarse-grained beads should be *adjusted* to give the correct behavior of the peptide solution

$$V = W_{solv} V_{solv} + W_{VdW} V_{VdW}$$

(plus electrostatic and hydrogen bond terms (fixed))



"GOLDEN RULE" FOR SIMULATING A SOLUTION OF PROTEINS:
equilibrium between associations and dissociations

Frequency of *association* depends on **concentration**

Frequency of *dissociation* depends on **bead-bead interaction**

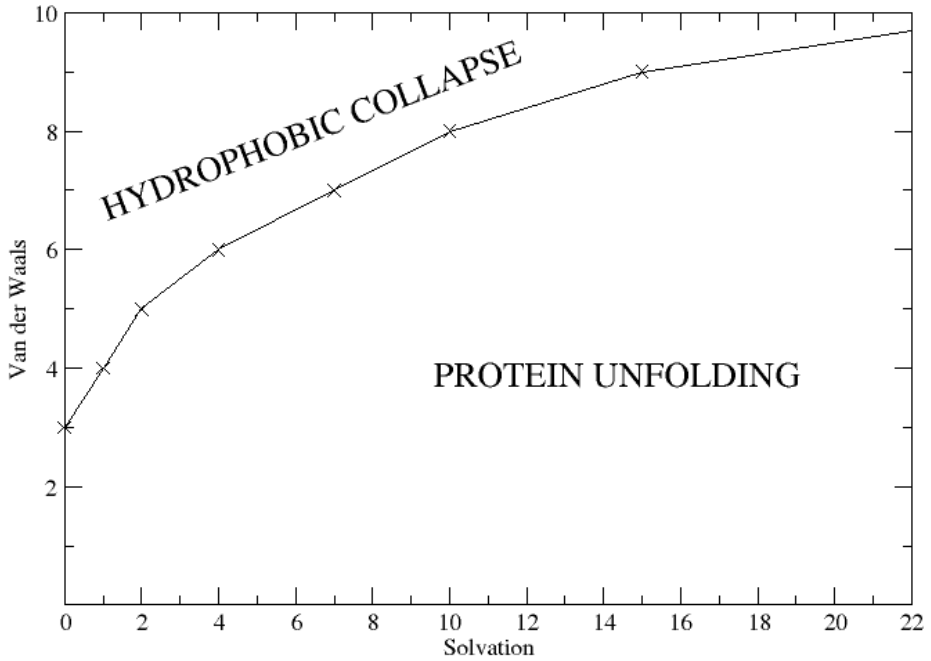
If association > dissociation, **hydrophobic collapse**. The solute precipitates

Interaction potential (between sidechain beads) $V = W_{solv} V_{solv} + W_{VdW} V_{vdW} + V_{el}$

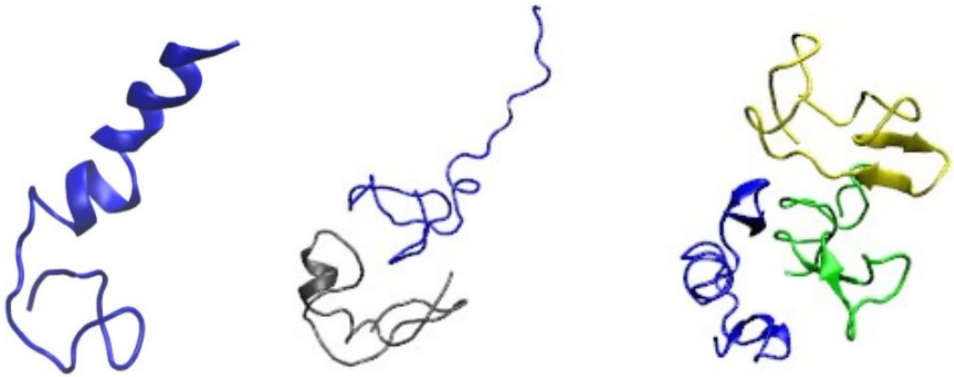
Van der Waals term *attractive*, solvation term mostly *repulsive*

No a priori knowledge of the values of W_{solv}, W_{VdW}
An exploration in the two-dimensional space W_{solv}, W_{vdW} should be made.

PHASE DIAGRAM OF A PROTEIN SOLUTION



Grid density: 1x1 (10x20 points)
Four independent simulations of a box with four amyloid beta peptides, Aβ40 at concentration 30 μM

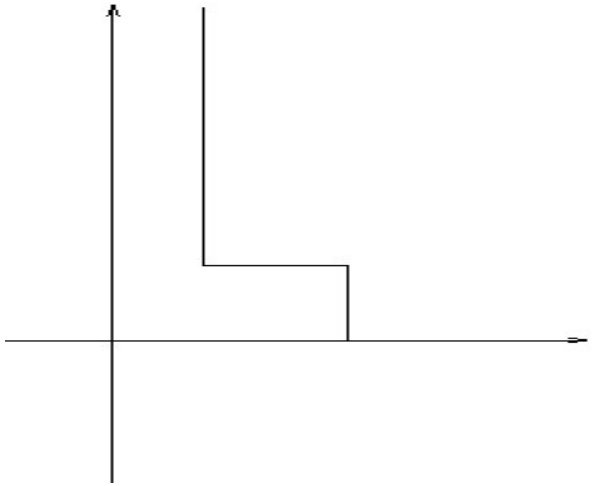


Any point in the phase transition line gives equilibrium associations/dissociations

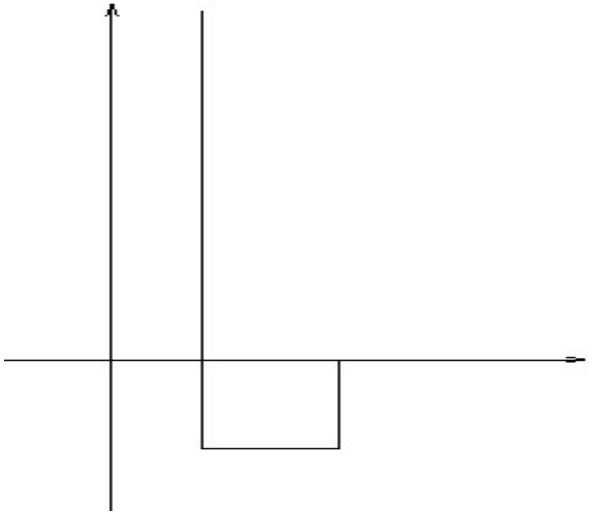
EXAMINING THE SOLVATION POTENTIAL

Implicit solvation potential mimics the effect of **water** on the solute particles. On average, the solvation term is *repulsive*

Hydrophilic particles are strongly solvated, surrounded by water molecules. They are kept apart → *repulsive* term



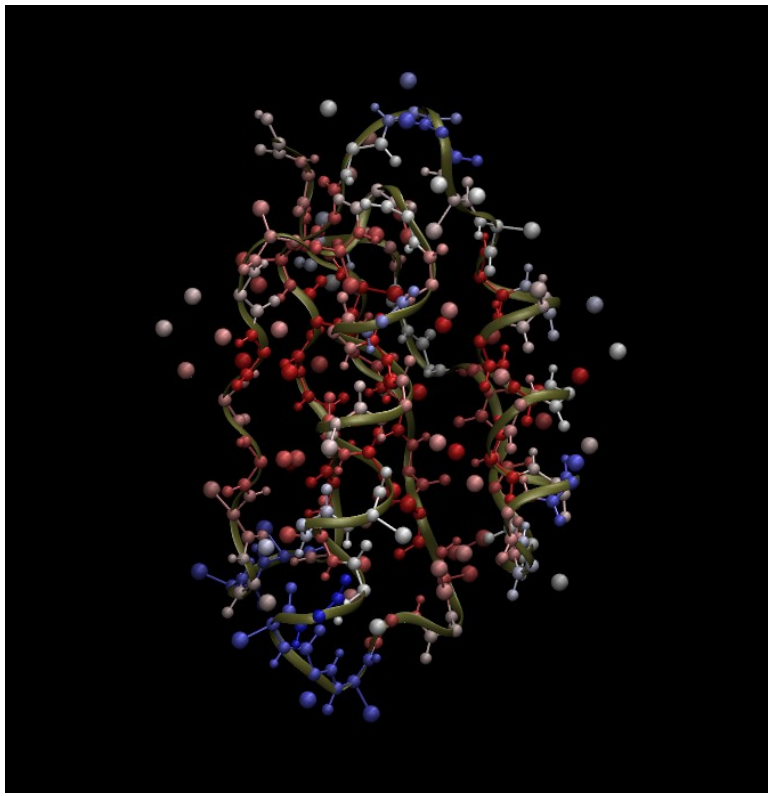
Hydrophobic particles are excluded by water, tend to join and form hydrophobic cores → *attractive* term

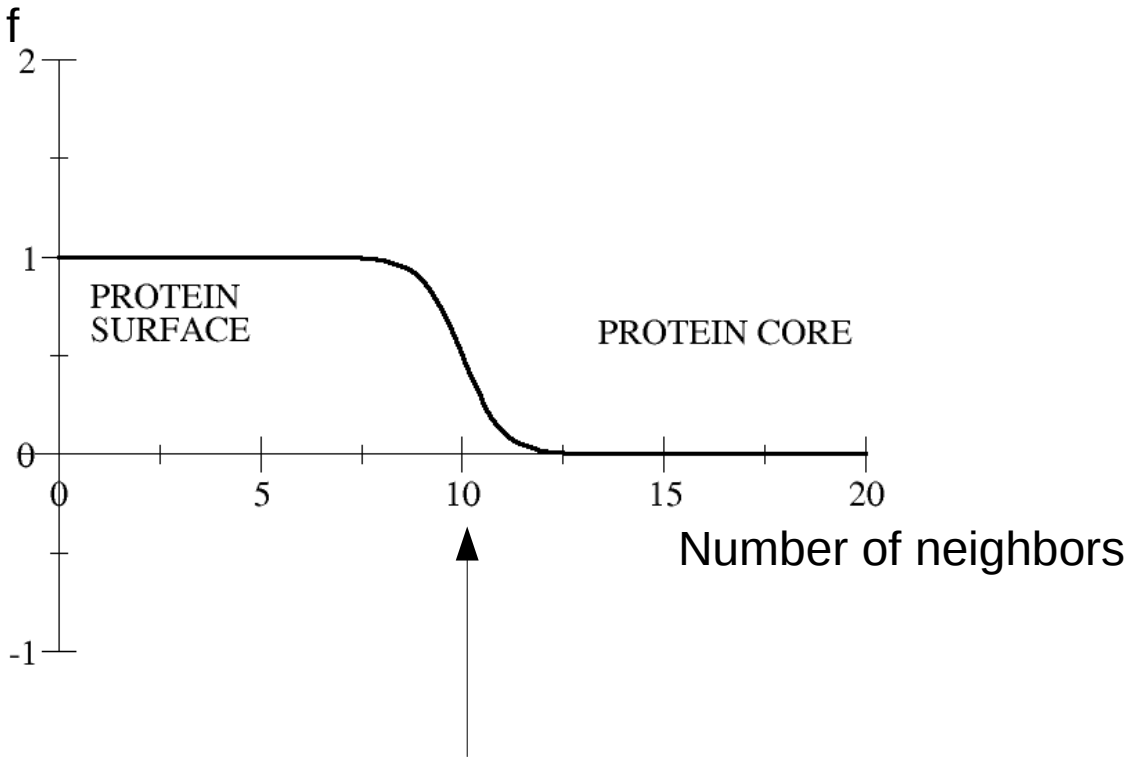


The implicit solvation model we use is *strictly valid* as long as all the particles of the protein are **completely exposed** to the solvent

The **implicit solvation** term has been modulated by a factor f that depends on the *level of exposition* of each particle to the solvent.

This factor becomes 0 for *no exposition* to the solvent (buried atoms) and is 1 at the *surface* of the protein (exposed to the solvent)





Number of neighbors of each atom

exposed/buried
limit value

Few neighbors

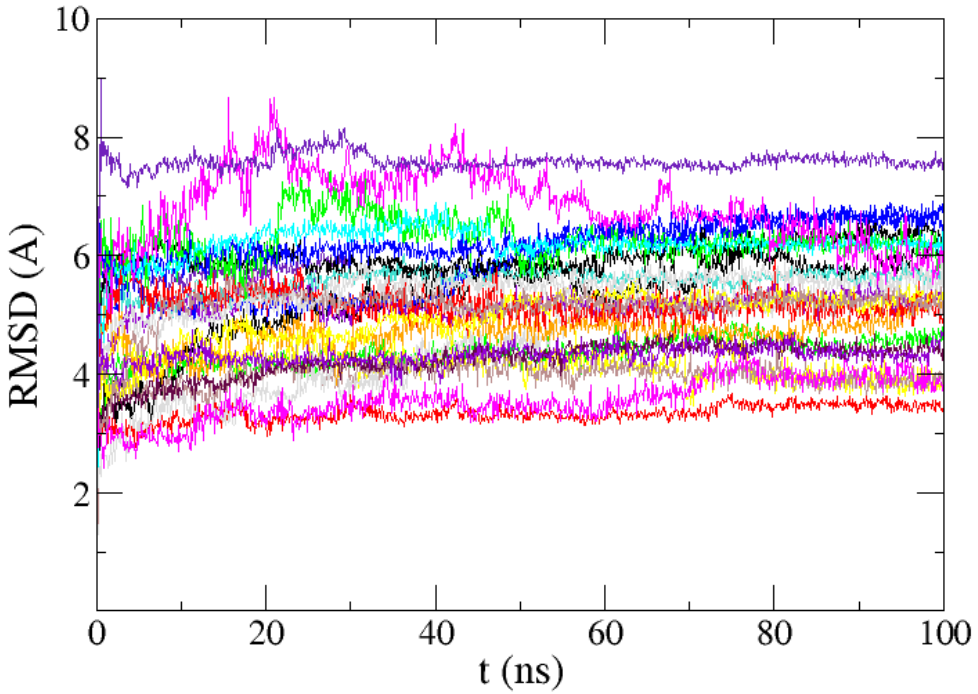
Many neighbors

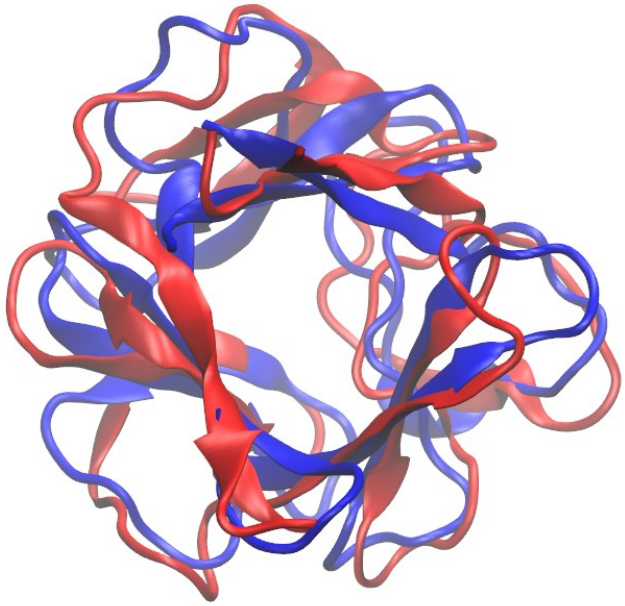
The coarse-grained model and force field was adjusted for a solution of Aβ40

Does it work correctly for folded proteins?

Performance of standard CG models for folded proteins: RMSD around 4 A

Simulation of a test set of proteins:





Average RMSD: 5 A

Included electrostatic interaction with **dielectric constant** ε_r=20

PARAMETERS OF THE PROGRAM TO BE ADJUSTED IN THE OPTIMIZATION OF THE FORCE FIELD

FVDW <----> W_{VdW}

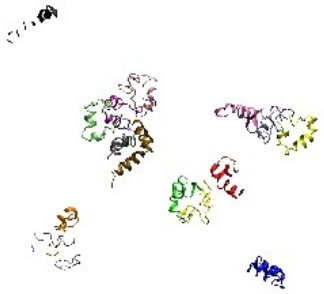
FSOLV <----> W_{solv}

ASOLV <----> position of the step in the modulation of solvation term

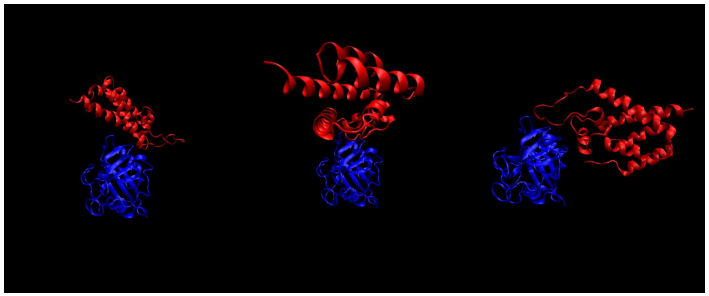
EPS <---> 332 /ε_r (kcal A)/(mol e^2)

APPLICATIONS:

Aggregation of peptides



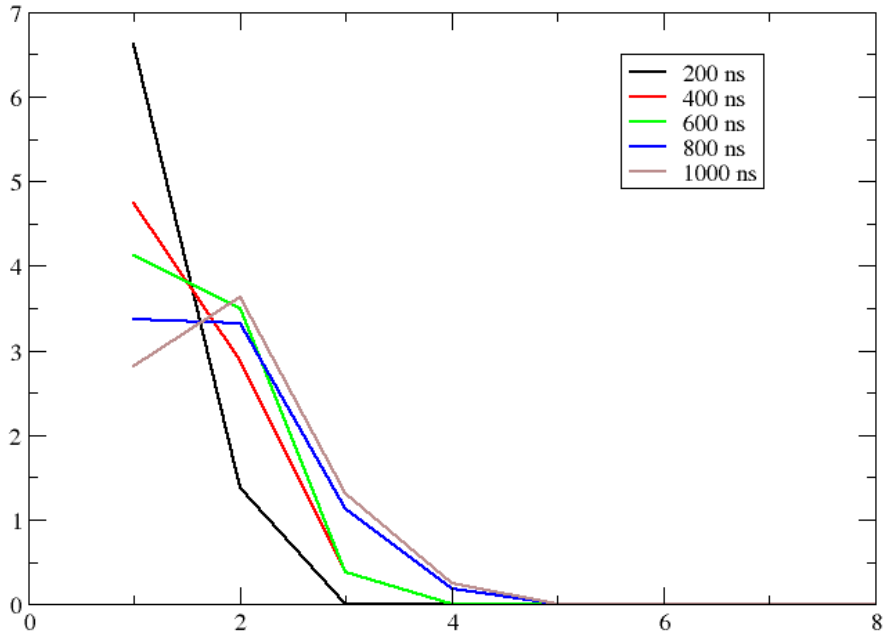
Protein-protein docking: which is the correct receptor-ligand configuration?



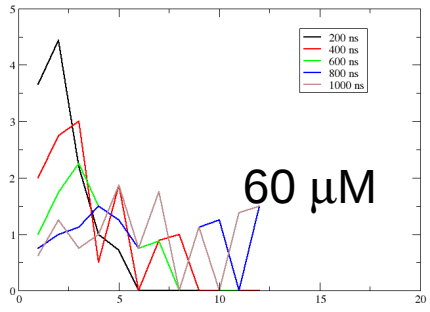
EXAMPLE: OLIGOMERIZATION OF THE Aβ40 PEPTIDE

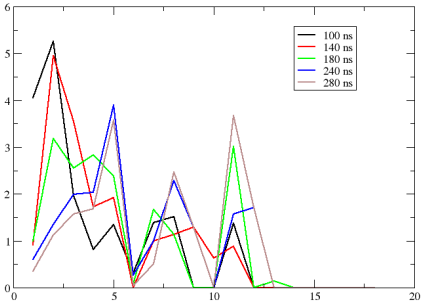
Comparison with experimental measurements for a 30 μM Aβ40 solution
8 peptides in the simulation box (80 nm side). 16 independent simulations

Oligomer size distribution



At higher concentrations, tendency to collapse





140 μM