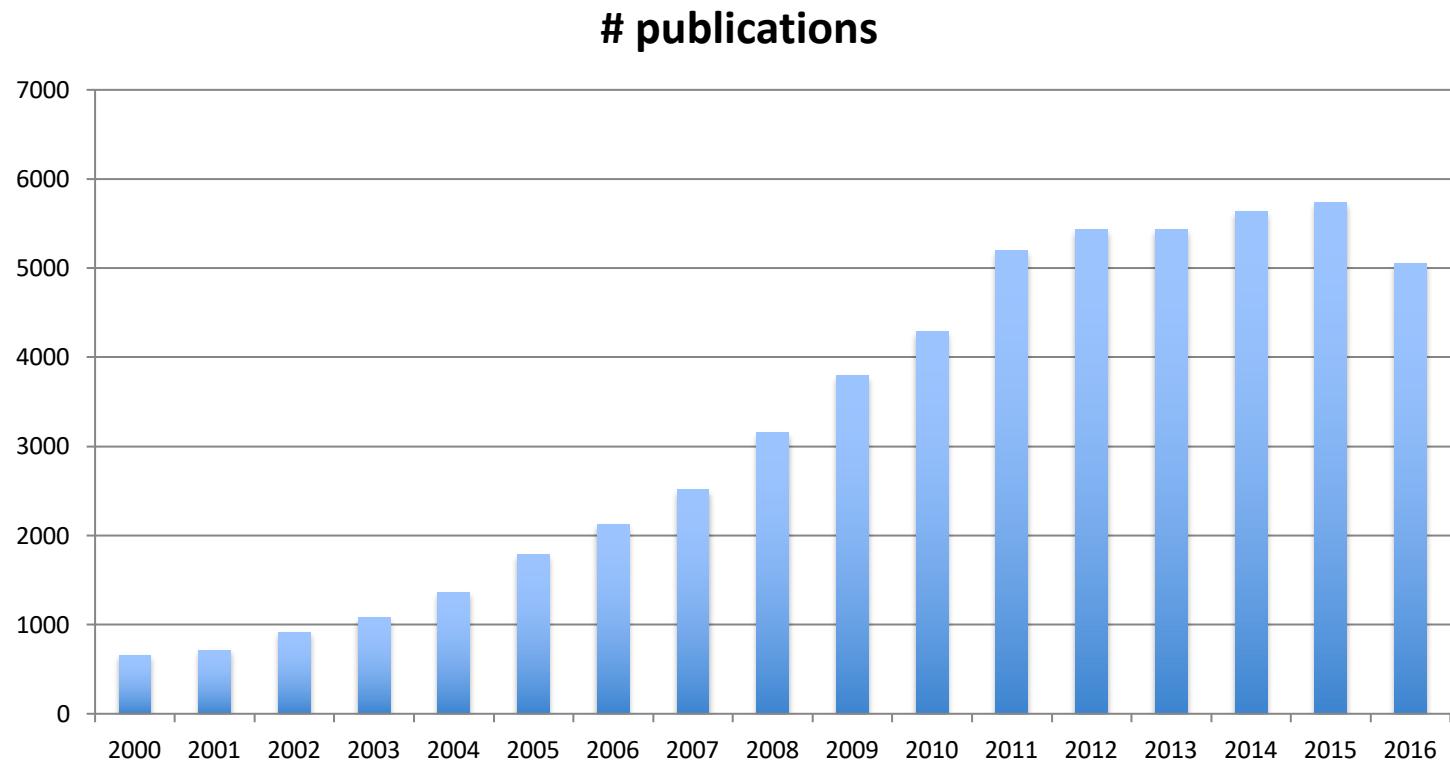


Coarse-Grain Models

What, Why, How...?

- **Definition of Coarse-Graining** (according to Saunders and Voth)
 - A method of reducing the complexity of a system by treating groups of atoms/molecules as single quasi particles
 - A bridge between highly detailed atomistic interactions and statistical mechanics

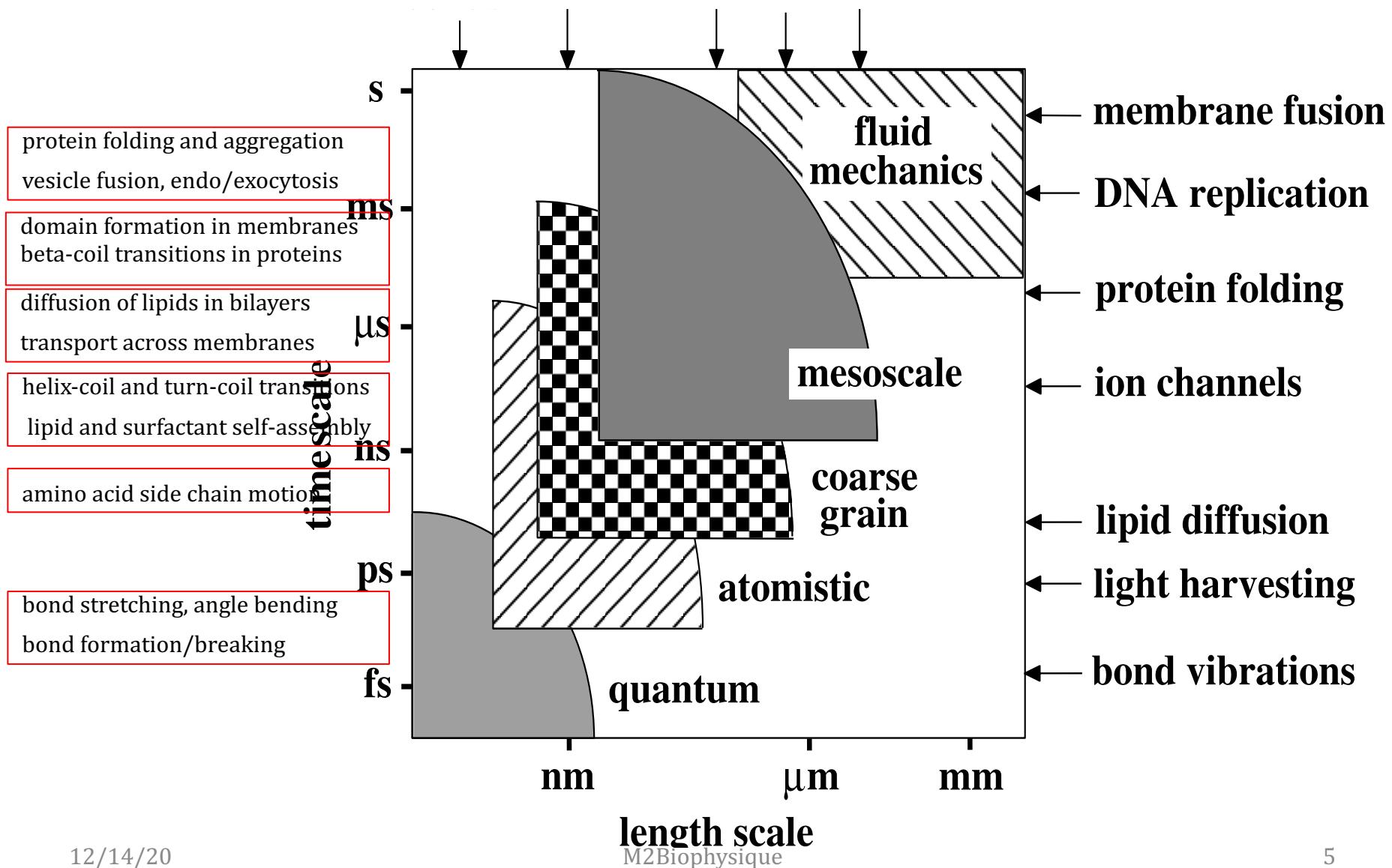
Importance



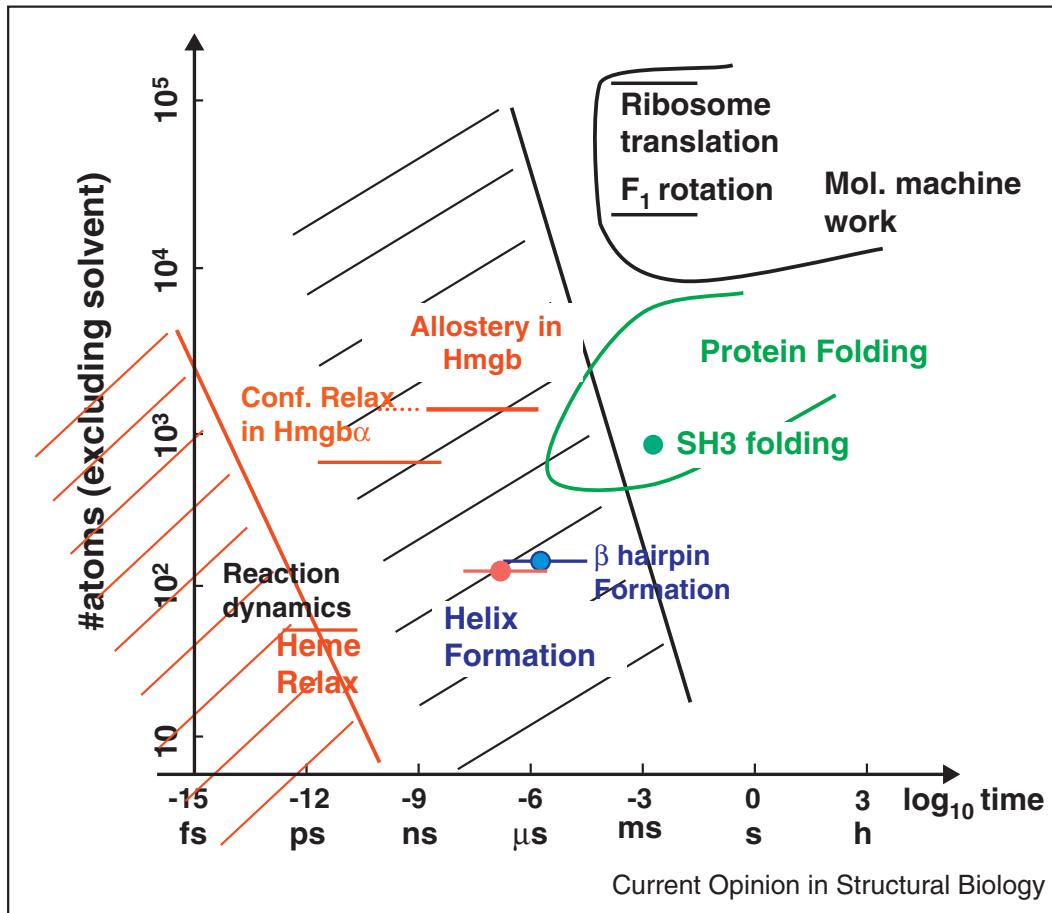
A challenge in Biology

Not only ! Chemistry

Spacio-Temporal Scales



Still Scales



Hierarchy in space and time of biomolecules. The red and black shaded areas roughly correspond to coverage of quantum mechanical simulations and all-atom MDs, respectively.

Assumption

Some details can be stripped off but...

Assumption

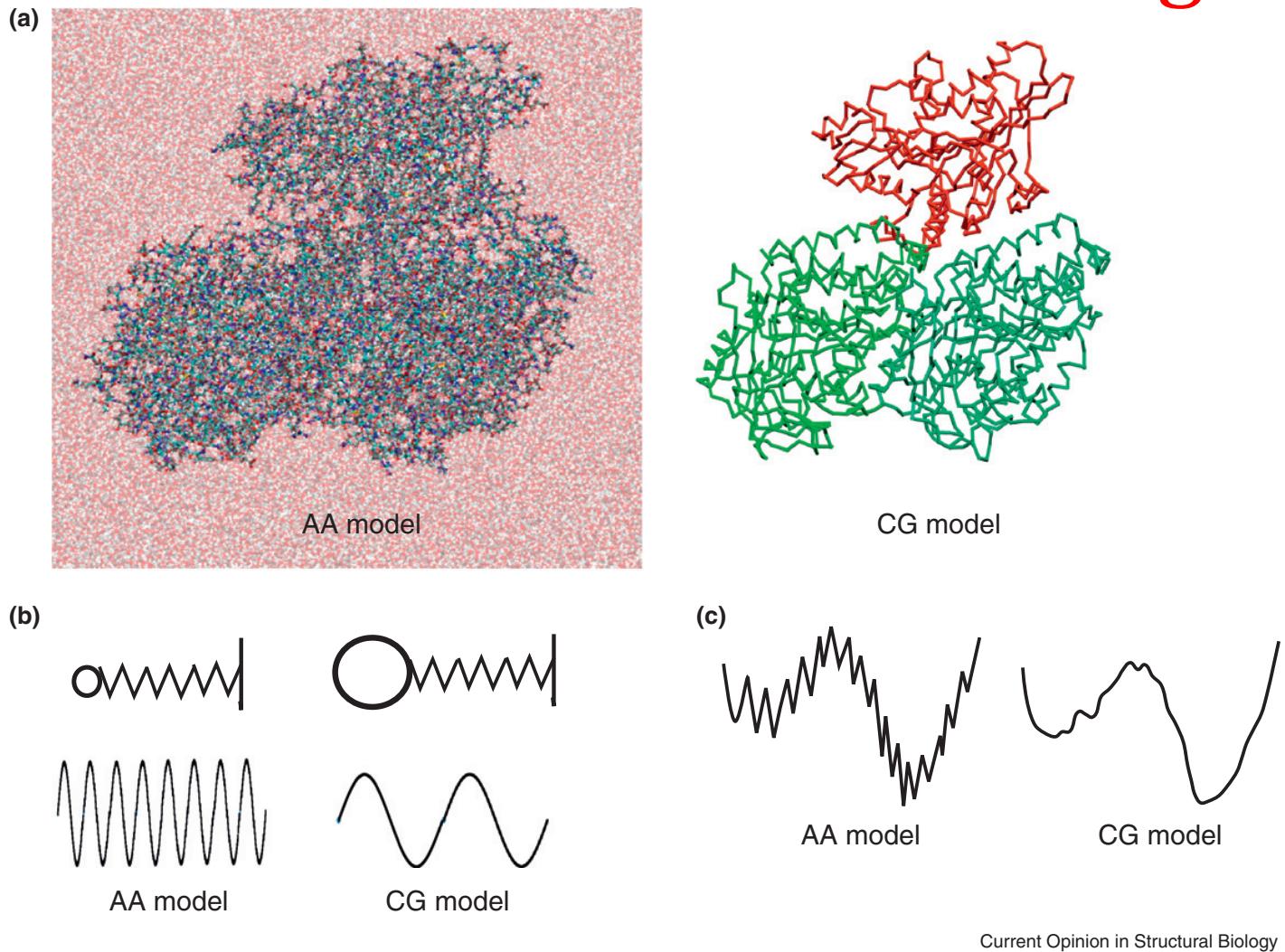
- Some details can be stripped off but..... It *depends on the problem* examined. Details: some degrees of freedom
e.g solvent details, internal deformation for large systems.
- Conditions for elimination:
 - Non essential for the process or property of interest
 - Large (number) or computationally intensive so that computational gain > loss in accuracy
 - Interactions governing these dof must be decoupled to the interactions governing the other dof: separation of frequencies
 - Elimination \Leftrightarrow efficient representation of the interactions governing the dof kept.

Some numbers

Level	Particles	Size of bead/nm	Scaling effort	CG reduction N_{df}	CG reduction comp. effort
I	Nucleons + electrons	10^{-6}	$N_n^{\geq 3}$		
II	Nuclei + electrons	10^{-6} – 10^{-5}	$N_e^{\geq 3}$	10–100	$> 10^3$
III	Atoms	0.03–0.3	N_a^{1-2}	10–100	$> 10^3$
IV	Supra-atomic beads	0.5–1.0	N_b^{1-2}	2–5	2–25
V	Supra-molecular beads	0.5–1.0	N_b^{1-2}	2–10	2–100

Rather Limited

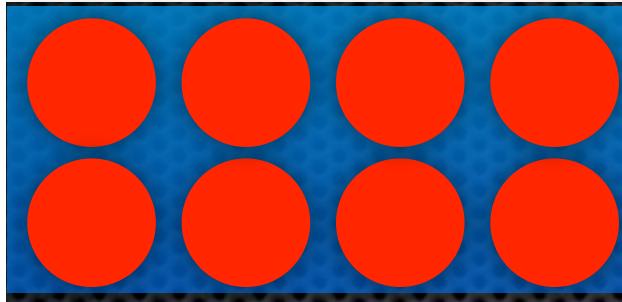
Reasons for Coarse Graining



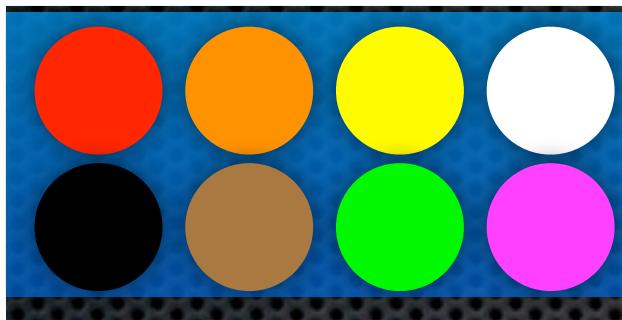
Coarse-Grained Models

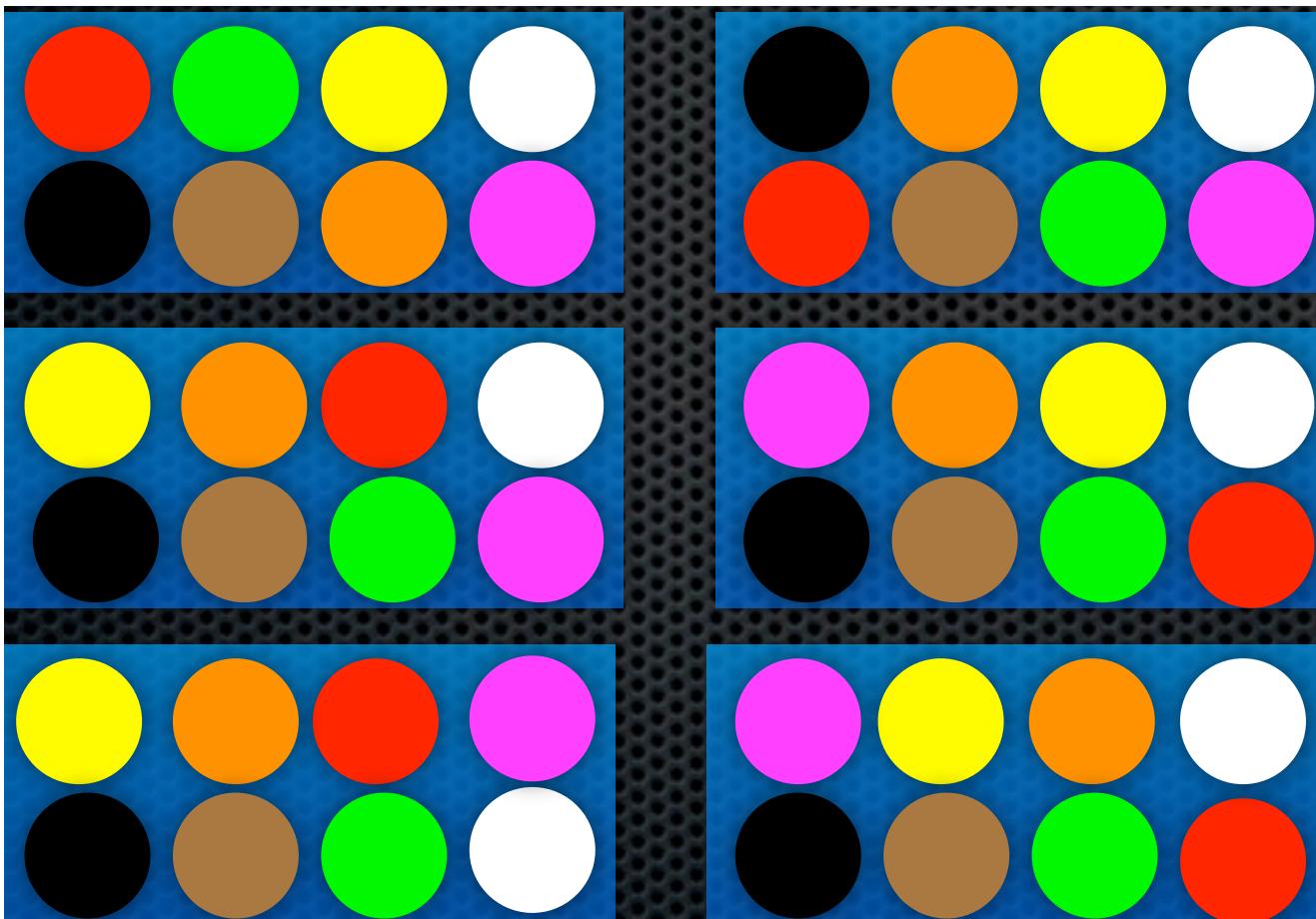
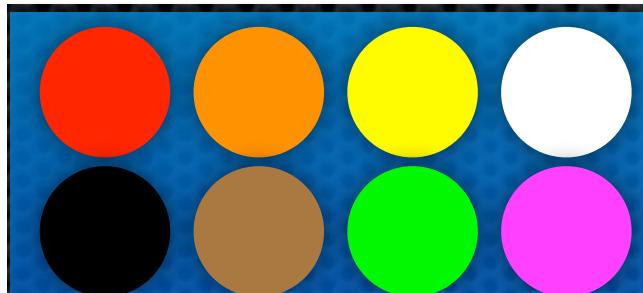
- Main points:
 - Series of atoms grouped in a pseudo-particle, Simplification of Solvent (important)
 - Need to (re)calibrate the potential function (the force field)
- Some molecules are better adapted to coarse-graining
 - Polymer with repeat units
 - Lipids because of the simplicity of tails
- Proteins, DNA and RNA are difficult:
 - units are different
 - Role of entropy

CG and Free Energy



CG and Free Energy

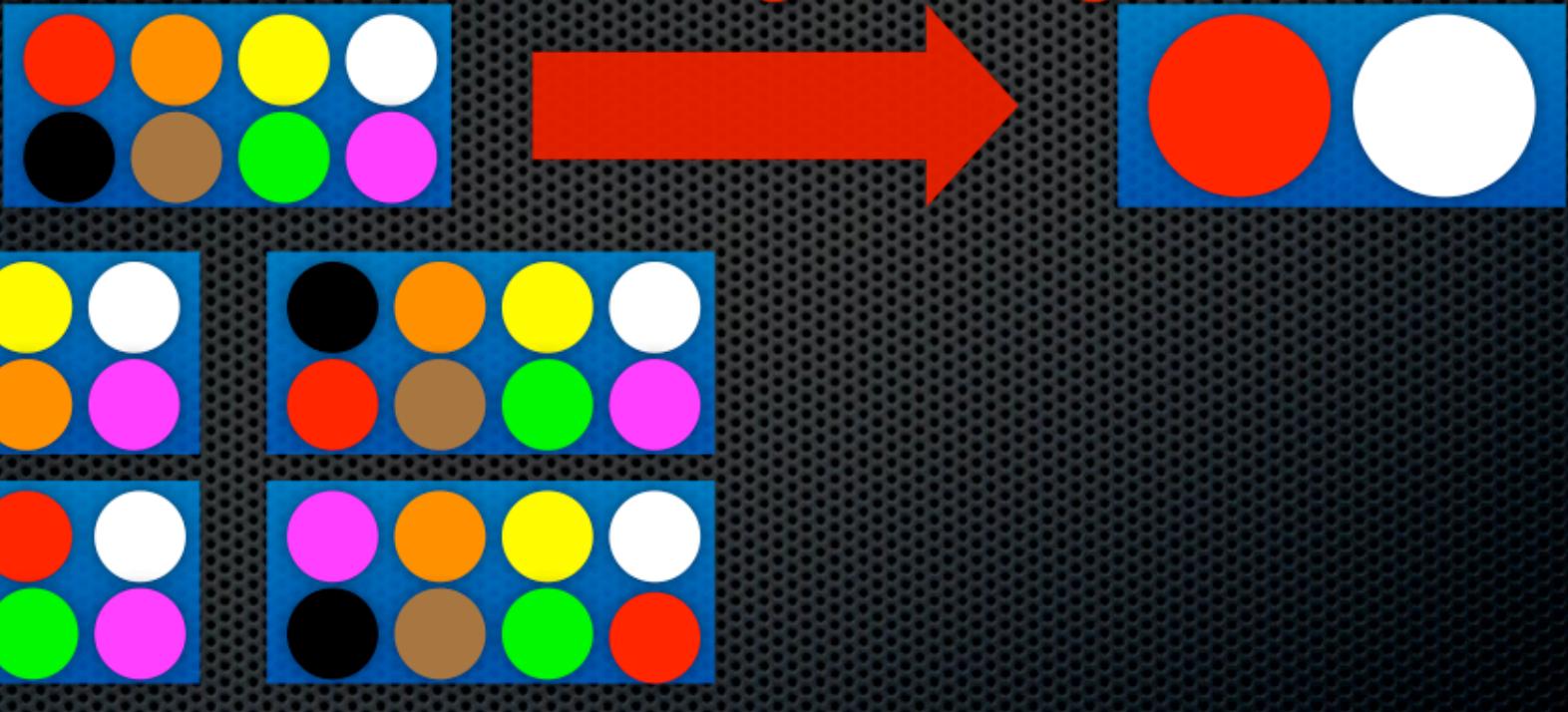




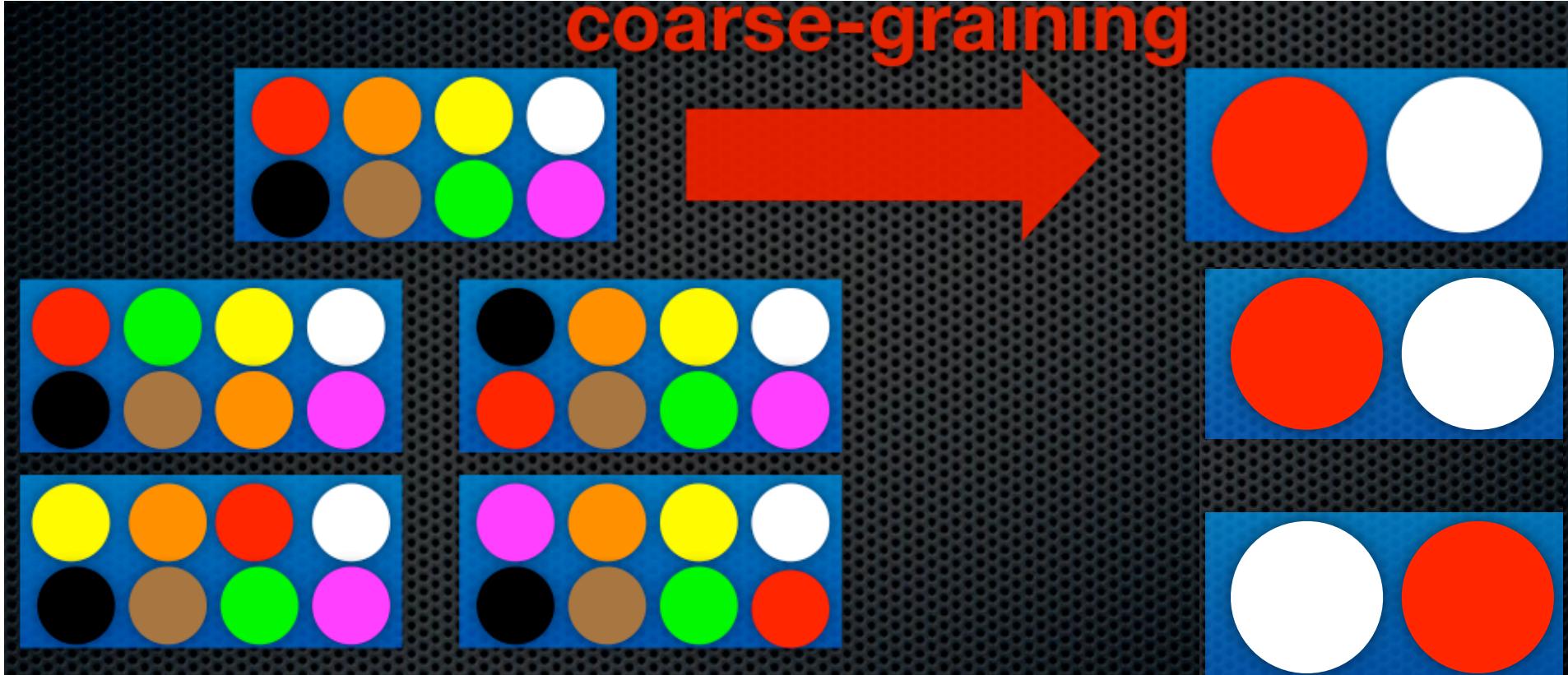
How many configurations?

Permutations = $8! = 40320$

coarse-graining



coarse-graining

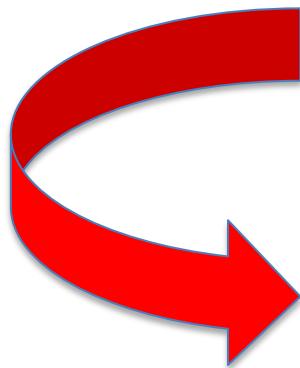


Permutations : 8!

Permutations : 2!

What we need

- Description of CG
- Interaction potential



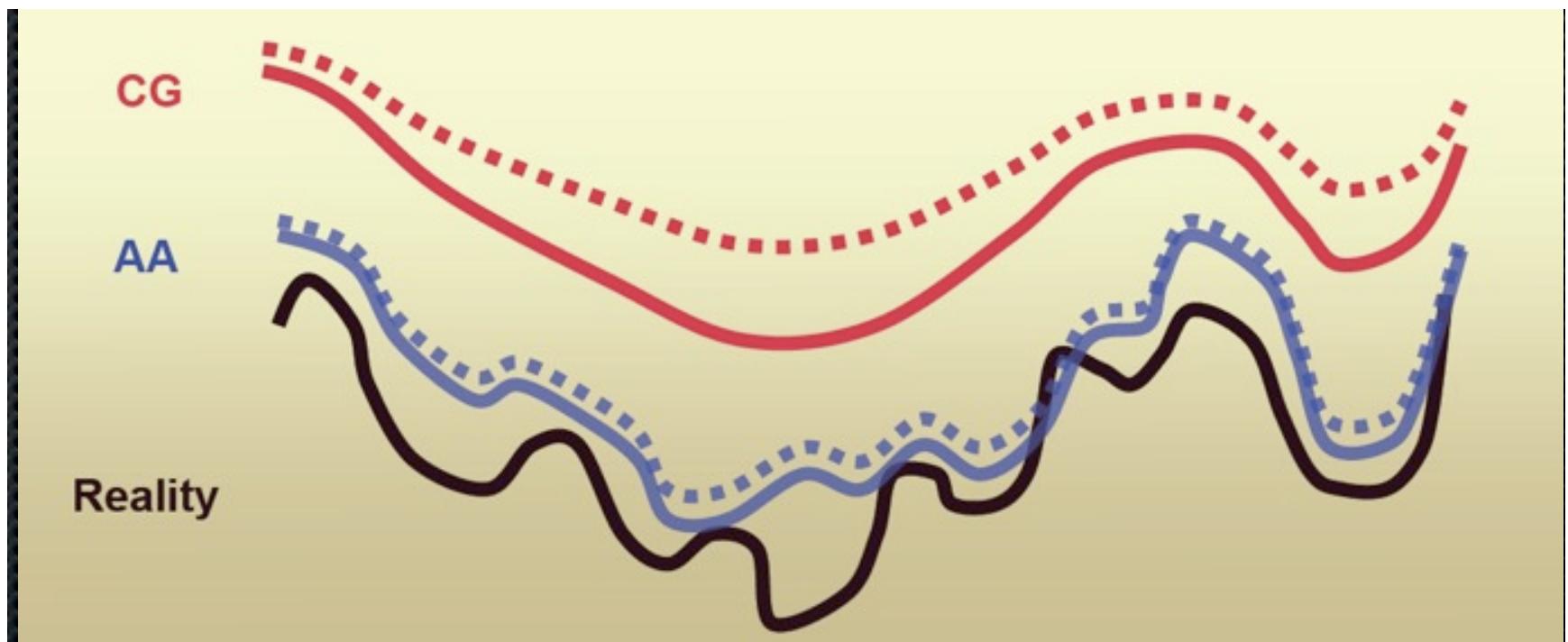
Respect Real Physics!

A goal

- Transferable potential.... ???

Drawbacks in CG

1. Poor transferability because the CG potential is actually a free energy
2. kinetics is altered because the free energy landscape is smoother



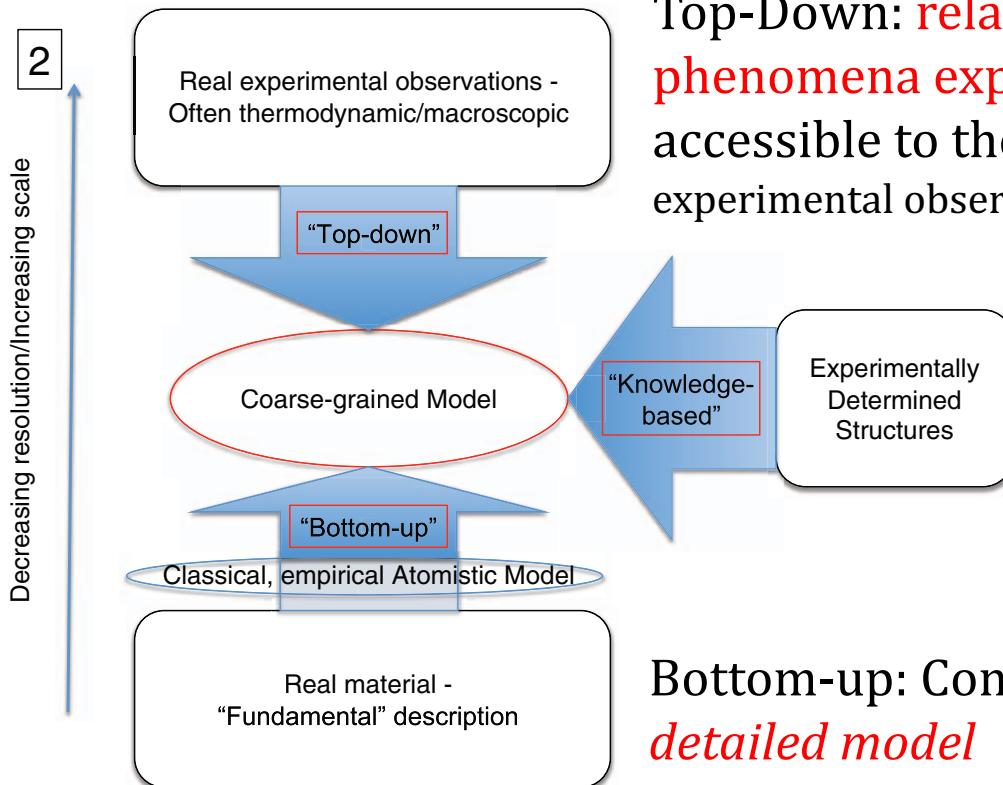
Consequences

- Entropy is reduced !
- As $\Delta G = \Delta H - T\Delta S$ if ΔG is correct
 - => ΔH is not correct
 - => T dependence is not correct

Two main categories

- **Structure-independent models :**
 - Based on physical-chemistry of interactions
 - Enable structure prediction of small proteins, disordered proteins, aggregates
 - Not effective for functional dynamics: Protein native structure is specific to its sequence => not stable enough with large fluctuations in the native state
- **Structure-based models :**
 - Use native structural information of the target molecule as input

Two main philosophies



Top-Down: related to the « real » system and phenomena experimentally observed on length scales accessible to the CG model. Fine grained model from experimental observables even more coarse

Bottom-up: Construction *from a more detailed model*

FIG. 2. Schematic illustrating the distinctions between top-down, bottom-up, and knowledge-based strategies for coarse-graining.

Consequences

- Bottom up: over-constrained models, loss of information
- Top-down: under-constrained models by the available experimental data. Many different models may lead to the same phenomena

The Potential

- Physics-based : physical forces to infer interactions between CG .
 - Thermodynamics-based models : need thermodynamic quantities (e.g. free energies) to fit on
- Knowledge-based: information extracted from experimental 3D structures. Empirical relations between statistics and interactions.

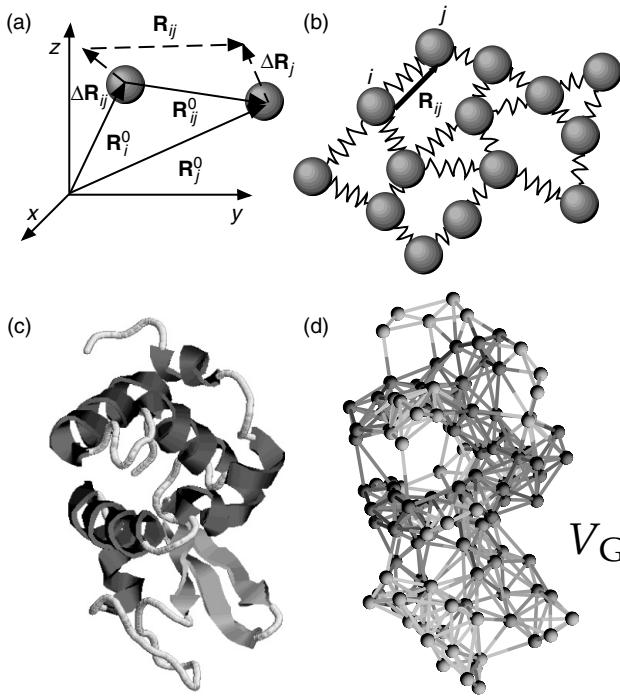
CG Models: a long (and old) story

- HP model (See Lecture 1). Very crude
 - Beads: Hydrophobic, Polar, Potential: ϵ for HH interaction and 0 for others

CG Models: a long (and old) story

- Tirion's model: Elastic Network Model
 - Beads: C α residues, Potential:
 - Two parameters: R_c and k_{enm}

$$V = \frac{1}{2} k_{enm} \sum_{d_{ij}^0 < R_c} (d_{ij} - d_{ij}^0)^2$$



Gaussian Network Model

$$\Gamma_{ij} = \begin{cases} -1, & \text{if } i \neq j \text{ and } R_{ij} \leq r_c \\ 0, & \text{if } i \neq j \text{ and } R_{ij} > r_c \\ -\sum_{j,j \neq i}^N \Gamma_{ij}, & \text{if } i = j \end{cases}$$

$$V_{GNM} = \frac{\gamma}{2} \left[\sum_{i,j}^N \Gamma_{ij} [(\Delta X_i - \Delta X_j)^2 + (\Delta Y_i - \Delta Y_j)^2 + (\Delta Z_i - \Delta Z_j)^2] \right]$$

CG Models: a long (and old) story

- Go-like models:
 - Beads: Residues, Potential: Attractive and Repulsive potential between non-bonded interactions biased toward the native configuration (use of native contacts)
⇒Based on minimal frustration principle

Limits of Go-like Models

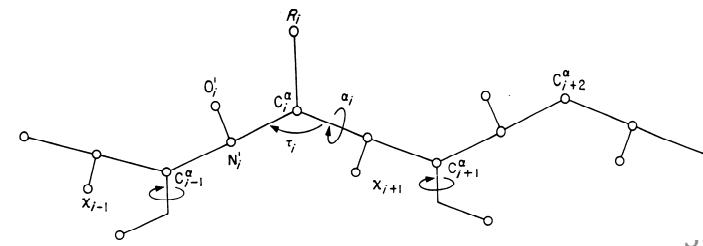
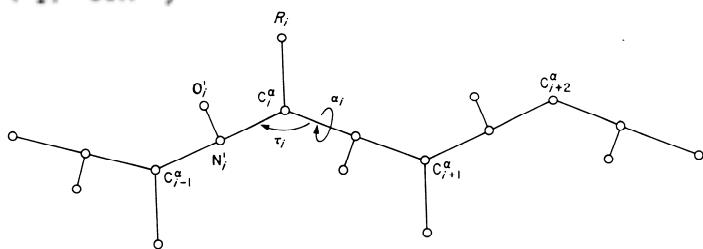
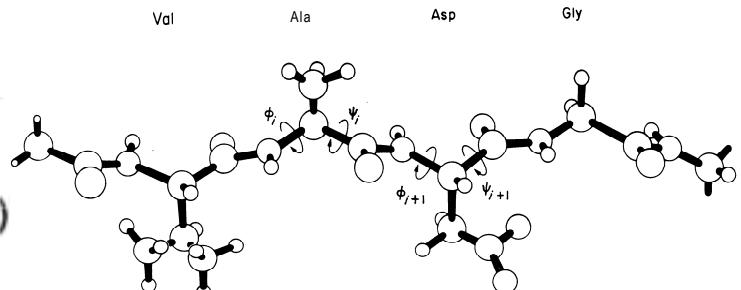
- Difficulties to reproduce ϕ analysis data
- Non-native contacts may play a role
- Lost of jigsaw puzzle properties of aa (as most CG)
- Problems to model the transition state

First attempts to overcome

Michael LEVITT, J. Mol.Biol.(1976) 104, 59-107

« A Simplified Representation of Protein Conformations for Rapid Simulation of Protein Folding »

$$\begin{aligned}
 V_{\text{tot}}(\alpha) = & \sum_{i,j} \epsilon_{ij} \left\{ 3(r_{ij}^{\circ}/r_{ij})^6 - 4(r_{ij}^{\circ}/r_{ij})^6 \right\} \\
 & + \sum_{\substack{i,j \\ r_{ij} < 9\text{\AA}}} (s_i + s_j) g(r_{ij}) + \sum_{\text{SS bonds}} K_{\text{ss}}(r_{ij}^{\text{ss}} - r_{ij}^{\circ}) \\
 & + \sum_{i,j} \epsilon_p \left\{ (r_p^{\circ}/r_{NO})^{12} - 2(r_p^{\circ}/r_{NO})^6 + (r_p^{\circ}/r_{ON})^{12} - 2(r_p^{\circ}/r_{ON})^6 \right\} \\
 & + 332 \sum_{i,j} q_p^2 \left\{ 1/r_{NN} + 1/r_{NO} - 1/r_{NO} - 1/r_{ON} \right\} \\
 & + \sum_i \left\{ 2 \sum_{k=1}^i A_k^i \cos[(k-1)\alpha_i] + B_k^i \sin[(k-1)\alpha_i] \right\}
 \end{aligned}$$



Example of improvements of Go-like Models

Das et al, PNAS July 19, 2005 vol. 102 no. 29 10141-10146

$V = V_{\text{local}} + V_{\text{nonlocal}}$, where V_{local} encompasses bond, angle, and torsional energy terms

$$V_{\text{nonlocal}} = \sum_{\substack{i,j=1 \\ i < j-3}}^N s(c_i, c_j) \left[5 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \delta(c_i, c_j) 6 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{10} \right]$$

$\sigma_{i,j}$ for a pair of residues (i, j) : extracted from a distribution $P(\sigma; c_i, c_j, |i - j|)$.
3 distributions for 3 separation ranges.

P: Occurrence of C α -C α distance over all native contacts for residue types c_i and c_j for the give $|i-j|$ range

Example of improvements of Go-like Models

Das et al, PNAS July 19, 2005 vol. 102 no. 29 10141-10146

- Determination of ε and δ through an iterative process in which the maximum energy gap criterion is applied on a set of decoys:

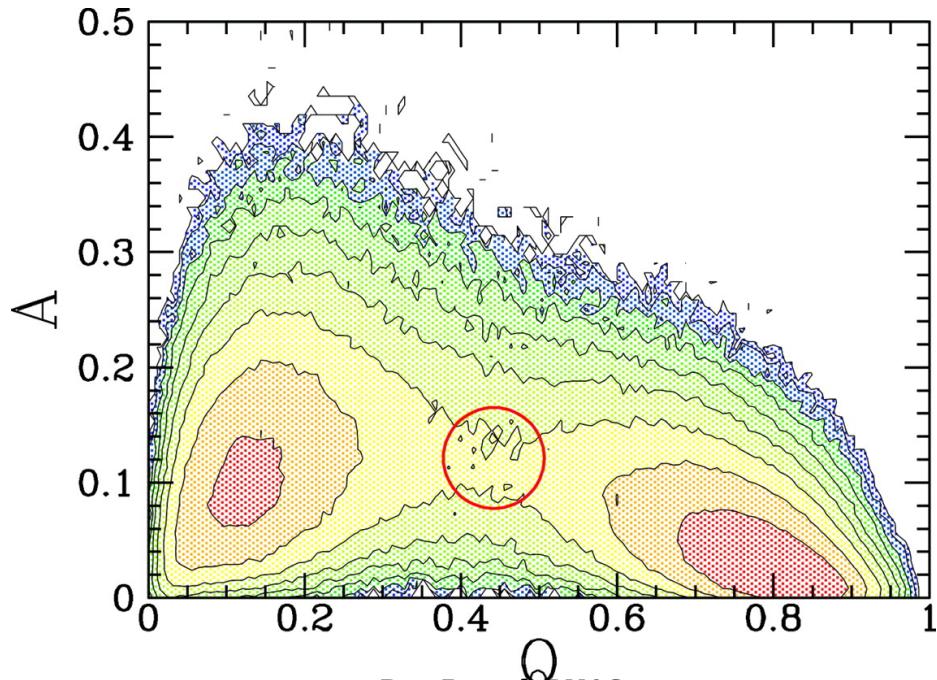
$$\Delta E(\{\varepsilon_{\text{opt}}\}, \{\delta_{\text{opt}}\}) = \max_{\{\varepsilon\}, \{\delta\}} [(\min_{i \in \{1, \dots, N_{\text{decoys}}\}} E_{\Gamma_i}) - E_{\Gamma_{\text{nat}}}],$$

where Γ_{nat} indicates the native structure (with energy $E_{\Gamma_{\text{nat}}}$), and E_{Γ_i} is the energy corresponding to the decoy structure Γ_i .

- Multiple “heat and quench” unfolding/refolding MD simulations with the obtained parameter set.
- If the native state of the considered protein is consistently recovered parameter set : effective => Further exploration
- Otherwise, the compact misfolded structures obtained in simulations are added to the decoy set and another iteration is implemented (step i).

Example of improvements of Go-like Models

- Results: Many solutions fill the criterion (16 sets of parameters highly correlated)
- Test on SH3 protein

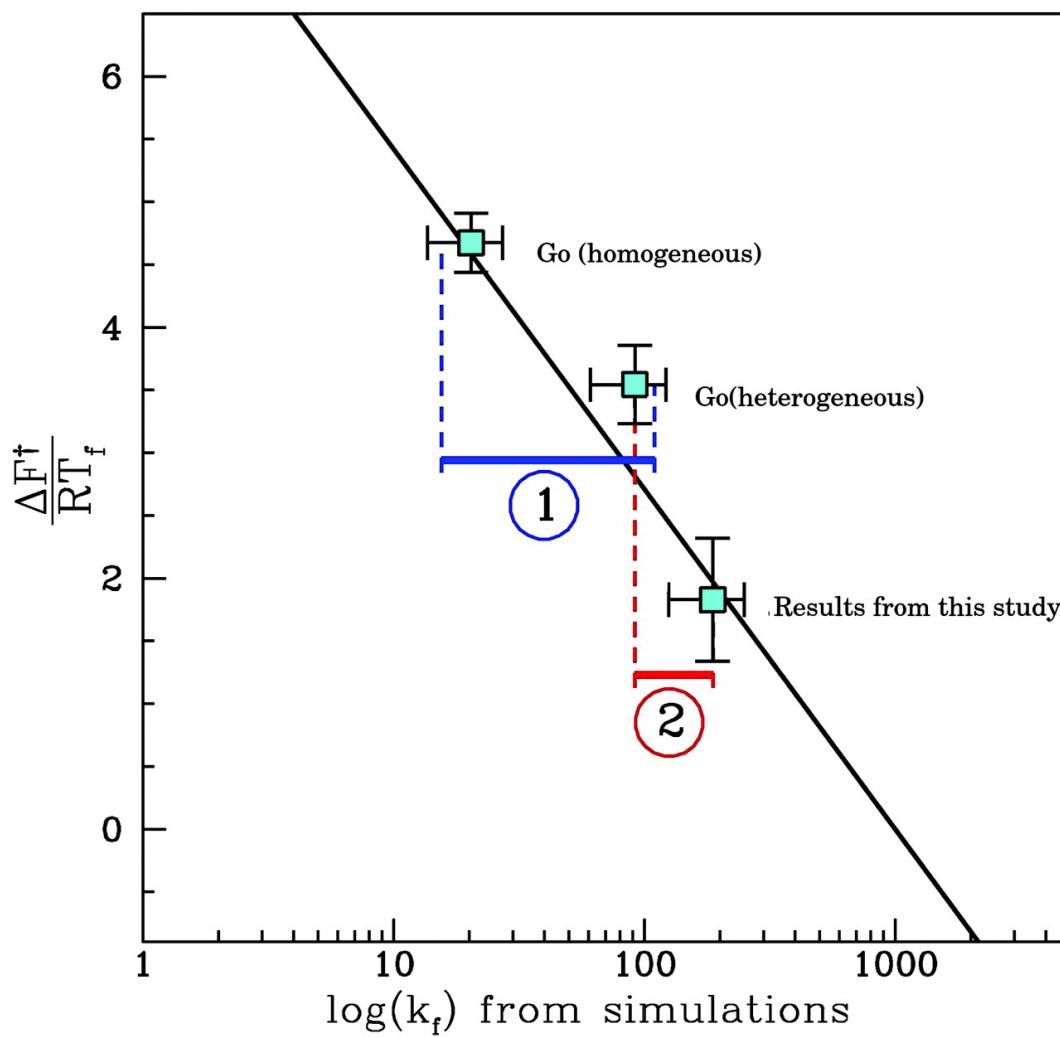


Free-energy surface as a function of the fraction of native, Q , and nonnative, A , contacts formed.

Each contour level marks a free-energy change of $1 RT_f$. The transition-state region is circled in red.

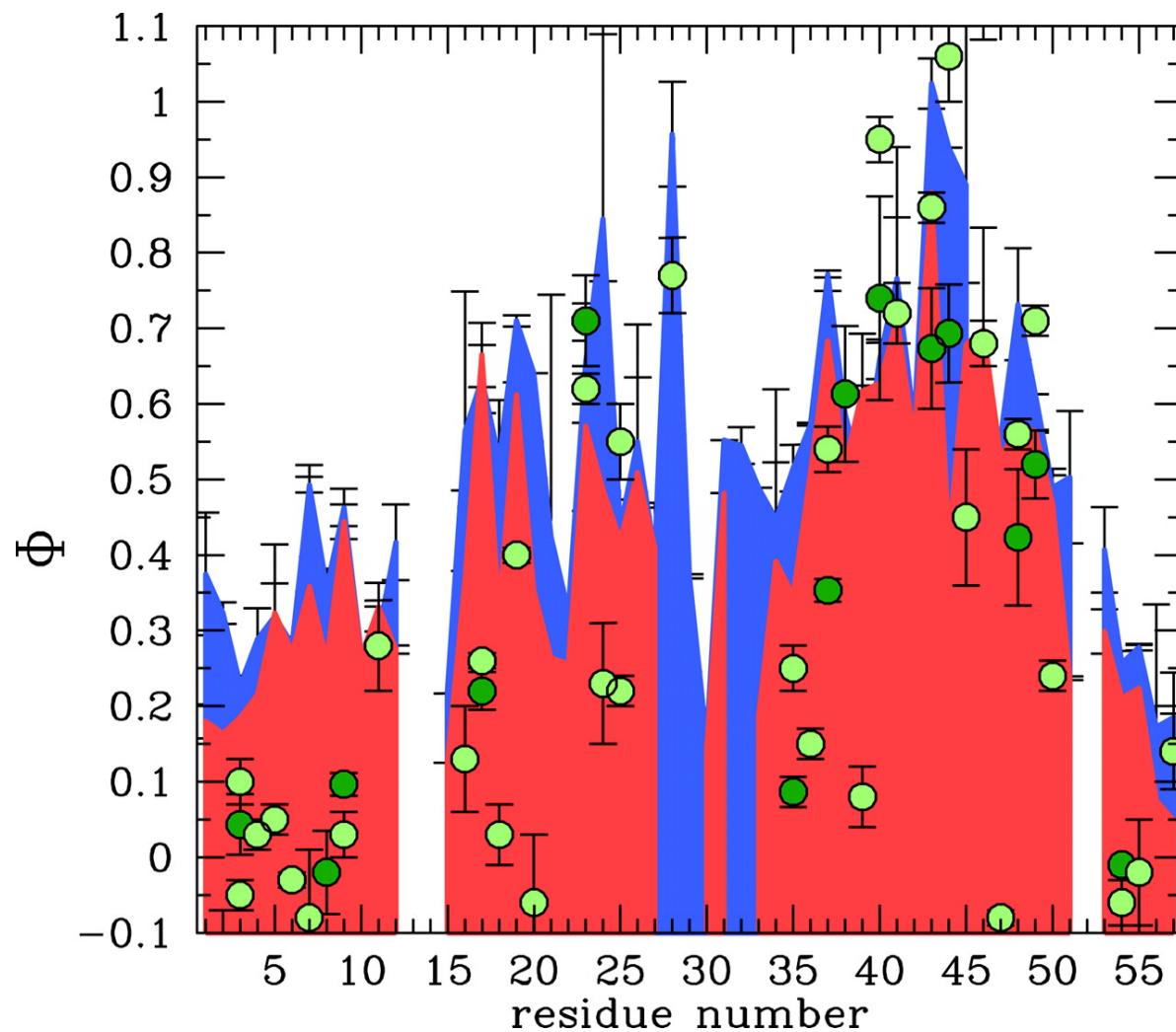
Impact of non native contacts.

The folding free-energy barrier ($\Delta F^\ddagger/RT_f$) is plotted against the logarithm of the folding rate (at the T_f), for the model defined in the text, the homogeneous G_0 model, and the heterogeneous G_0 model.



Das P et al. PNAS 2005;102:10141-10146

Φ -values obtained from simulations are compared with the experimental data.



Das P et al. PNAS 2005;102:10141-10146

Limits

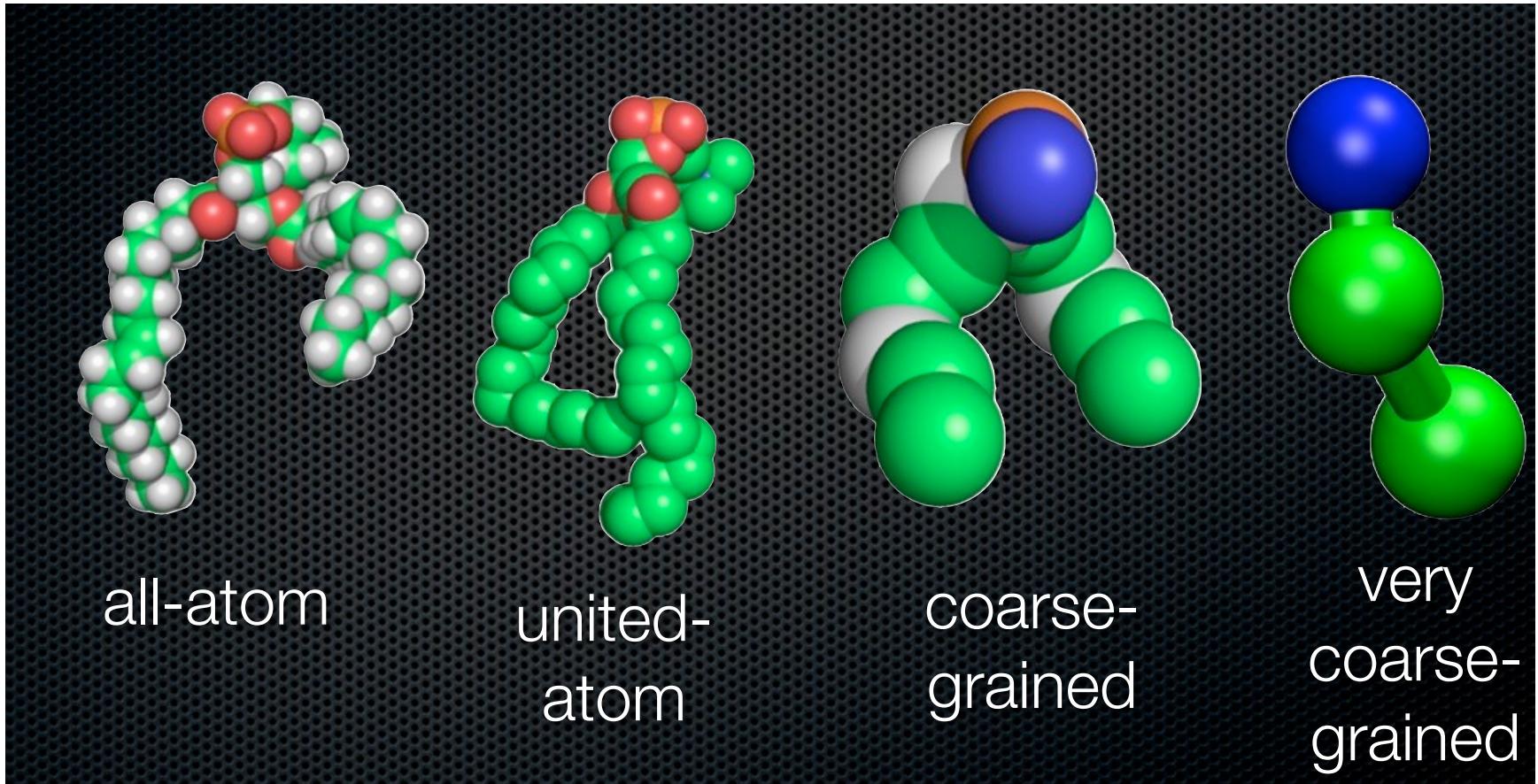
ENM and Go-like models depends on a reference configuration.

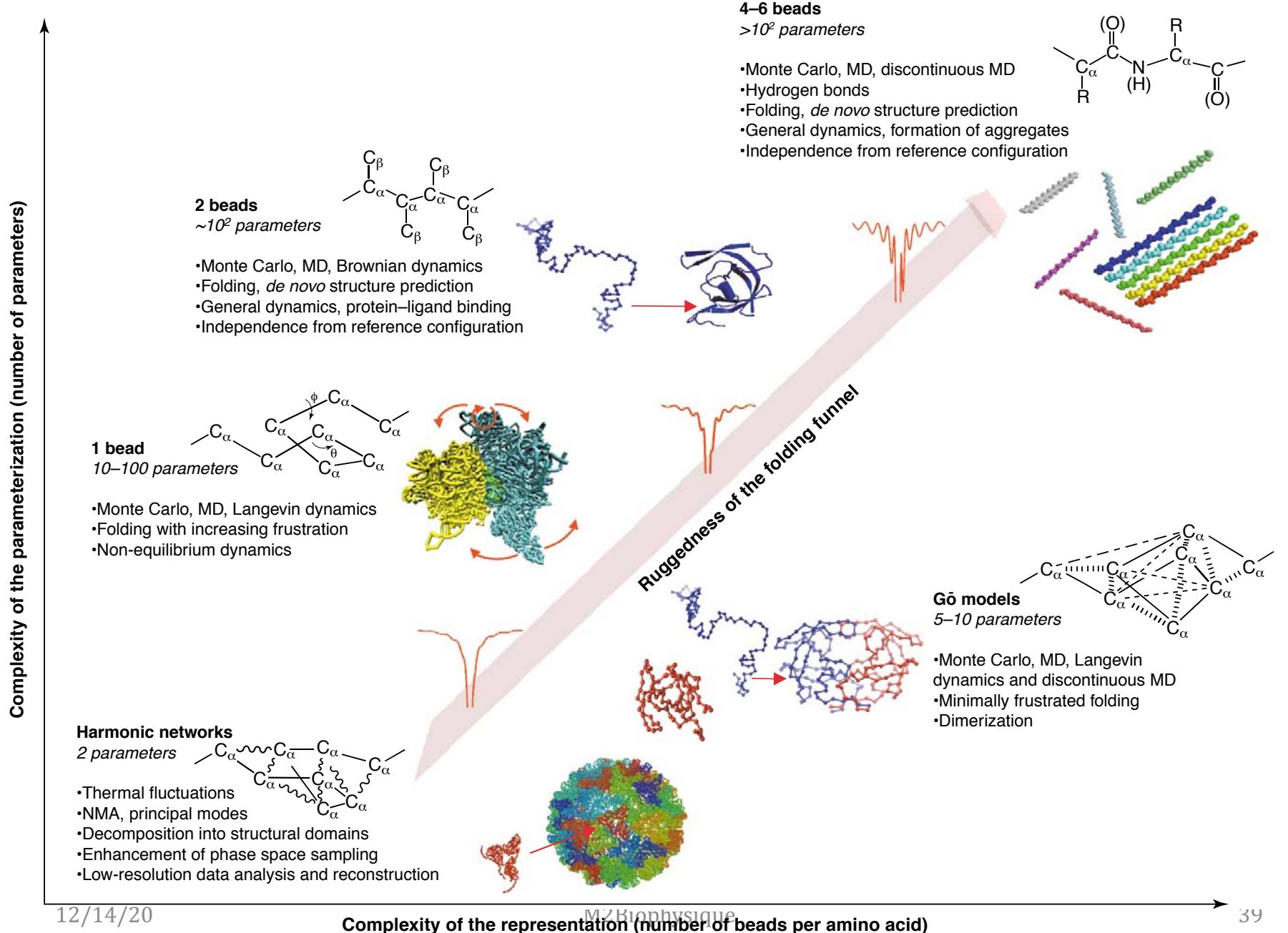
Structure-Based Model

Knowledge-Based Potential

- Use a set of PDB structures to design a transferable potential.
 - 3 classes:
 - pure PDB statistics
 - Optimized on the basis of known structures
 - Explicitly tailored for a particular protein on the basis of known structures.
- Statistical Potentials
 - Scheraga & Tanaka, Sippl, Miyazawa-Jernigan
 - Sippl: Boltzmann device=> Boltzmann inversion.

Examples





Main Equation

$$\exp(-\beta F) \propto \int d\mathbf{R} \exp[-\beta V(\mathbf{R})]$$

F: (Helmhotz) Free Energy: A thermodynamics property

$V(r)$: potential energy. Depends on coordinates of the system
of N_r atoms

$$\beta = 1/k_B T$$

For the recall

$U(S, V)$: Internal Energy p : Pressure

$H(S, p) = U + pV$: Enthalpy V : Volume

$A(T, V)$ or $F(T, V) = U - TS$: Helmholtz Free Energy T : Temperature

$G(T, p) = H - TS$: Gibbs Free Energy S : Entropy

Main Equation

$$\exp(-\beta F) \propto \int d\mathbf{R} \exp[-\beta V(\mathbf{R})]$$

Rarely solved !

BUT

The basis of numerous properties thermodynamics:
Distribution Functions
Equilibrium

Don't forget

The microstate of a system with N sites is defined by
N cartesian coordinates \mathbf{R} and...
N momenta \mathbf{P}

The Hamiltonian: with M_i Masse

$$H(\mathbf{R}, \mathbf{P}) = \sum_{i=1,N} \frac{1}{2M_i} \mathbf{P}_i^2 + V(\mathbf{R})$$

Similar Equation with CG

- Replace \mathbf{R} by \mathbf{R}_{CG} and V by V_{CG}

$$\int d\mathbf{R} \exp[-\beta V(\mathbf{R})] \equiv \int d\mathbf{R}_{CG} \exp[-\beta V_{CG}(\mathbf{R}_{CG})]$$

$$\begin{array}{ll} \mathbf{R} \leftrightarrow \mathbf{R}_{CG} & \text{Mapping?} \\ \mathbf{V} \leftrightarrow \mathbf{V}_{CG} & \mathbf{V}_{CG} ? \end{array}$$

Vectors are noted in Bold or
overlined with arrows

CG sites?

- How to define \mathbf{R}_{CG} ?

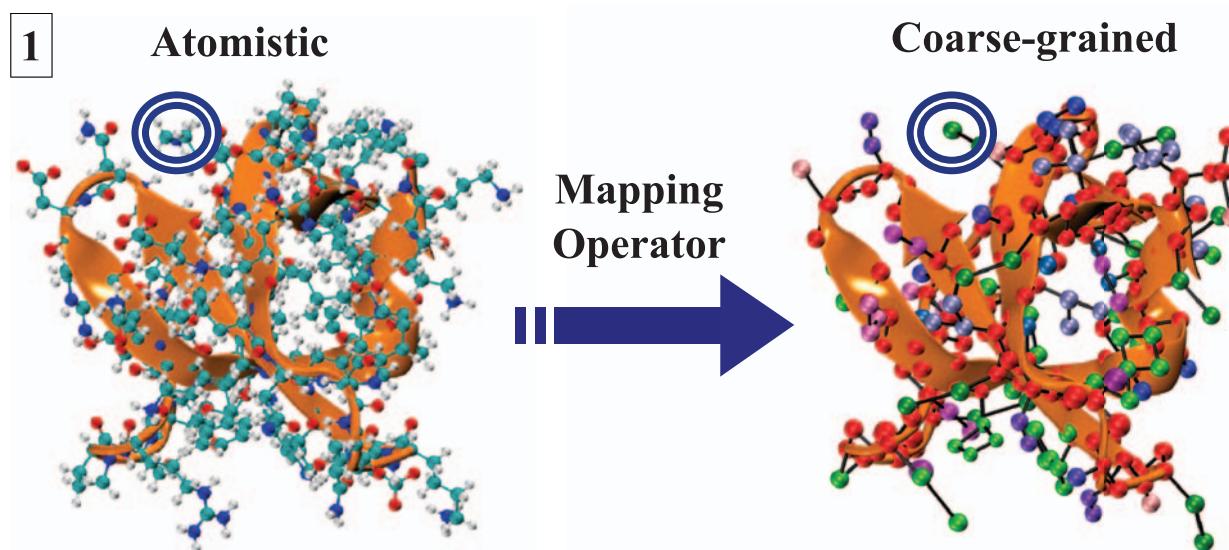
CG Sites

- How to define \mathbf{R}_{CG} ? from \mathbf{R} ?

CG Sites? Mapping

- How to define \mathbf{R}_{CG} ? from \mathbf{R} ?
- The most « natural » choice ***CG: Groups of atoms.***

=> Needs a mapping operator



Mapping

- The mapping operator?
- Set of Mapping Functions M: Linear combinations

$$M_{R_{CG}}^N(\mathbf{R}^n) = \left\{ M_{R_{CG1}}(\mathbf{R}^n), \dots, M_{R_{CGN}}(\mathbf{R}^n) \right\}$$

$$M_{R_{CGI}}(\mathbf{R}^n) = \sum_{i=1}^n C_{Ii} \mathbf{R}_i$$

$$\sum_{i=1}^n C_{Ii} = 1$$

$$M_{P_{CGI}}^N(\mathbf{P}^n) = M_I \sum_{i=1}^n C_{Ii} \mathbf{P}_i / m_i$$

For I runs N CG sites and i runs for n atoms

.... Few Equations later!

- CG model consistent with atomistic model in ***phase space*** if and only if :

Eq. 1

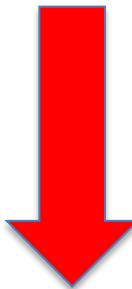
$$\exp(-\beta V_{CG}(\mathbf{R}_{CG}^N)) \propto \int d\mathbf{R} \exp(-\beta V(\mathbf{R}^n)) x \delta(M_{R_{CG}}^N(\mathbf{R}^n) - \mathbf{R}_{CG}^N)$$

$$\exp(-\beta \sum_{I=1}^N \frac{\mathbf{P}_{ICG}^2}{2M_I}) \propto \int d\mathbf{P}^n \exp(-\beta \sum_{i=1}^n \frac{\mathbf{P}_i^2}{2m_i}) x \delta(M_{P_{CG}}^N(\mathbf{P}^n) - \mathbf{P}_{CG}^N) \quad \text{Eq. 2}$$

- CG model consistent with atomistic model in ***configuration space*** if and only if Eq.1 holds

Consequences

$$\exp(-\beta V_{CG}(\mathbf{R}_{CG}^N)) \propto \int d\mathbf{R} \exp(-\beta V(\mathbf{R}^n)) x \delta(M_{R_{CG}}^N(\mathbf{R}^n) - \mathbf{R}_{CG}^N)$$

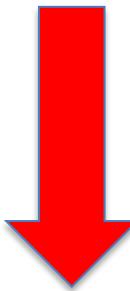


$$V_{CG}^0(\mathbf{R}_{CG}^N) = \beta \ln z(\mathbf{R}_{CG}^N) + Cste$$

$$z((\mathbf{R}_{CG}^N)) = \int d\mathbf{R} \exp(-\beta V(\mathbf{R}^n)) x \delta(M_{R_{CG}}^N(\mathbf{R}^n) - \mathbf{R}_{CG}^N)$$

Consequences

Z depends on the CG configuration R and on the value of the atomistic potential $V(R)$ through the entire configurational space



$V^0(R_{CG})$: a many-body potential of mean force

PMF : the potential yielding to the same force as that resulting from all the average over all the relevant configurations of a given system

Consequences

- $V^0(R_{CG})$: depends upon the system **and** the thermodynamic state point, i.e, *includes entropic and enthalpic contributions*
 - Entropic because certain atomistic degrees of freedom are included in CG

But...

- It remains impractical to evaluate the integral

$$V_{CG}^0(\mathbf{R}_{CG}^N) = \beta \ln z(\mathbf{R}_{CG}^N) + Cste$$

$$z((\mathbf{R}_{CG}^N)) = \int d\mathbf{R} \exp(-\beta V(\mathbf{R}^n)) x \delta(M_{R_{CG}}^N(\mathbf{R}^n) - \mathbf{R}_{CG}^N)$$



V_{CG}^0 will be approximated

Note that :

- Two situations:
 - Involved atoms :means that some atoms could be involved in **many** sites
 - Specific atoms : only involved in one site.

The coarse-grained force field

- Calculation of the forces as linear combination of spatial derivate with regard to atomistic coordinates.

$$\rightarrow F_I(R_{CG}^N) = \langle \mathcal{F}_I(R^n) \rangle_{R_{CG}^N} \quad \text{Eq A}$$

$$\mathcal{F}_I(R^n) = \sum_{Specific} \frac{d_{Ij}}{c_{Ij}} f_j(R^n) \quad \sum_{Specific} d_{Ij} = 1$$

Summary

- The force on CG site I : in terms of an equilibrium average, for the atomistic model, of $\mathcal{F}_I(R^n)$, which is a linear combination of the atomistic forces acting on the atoms that are specific to CG site I.
- => A CG model will be consistent in configuration space with a given atomistic model if there is at least one atom specific to each site and if the force on each CG site I in a given CG configuration R_{CG}^N is given by Eq A

What about momenta?

- Same reasoning
- Assumption: no atom i be involved in the definition of more than one CG site

\Rightarrow CG Masses satisfy:

$$M_I = \left(\sum_{i \text{ specific}} \frac{c_{li}^2}{m_i} \right) \text{ for all } I$$

Mapping coefficient
Atomic mass



Consistency in the phase space

Algorithms for Consistency

- **Direct Boltzmann inversion (DBI)**: as for Sippl's potential BUT from a *many body* PMF describing the average force on N-sites and which is determined by an N body distribution.
- **Iterative Boltzmann inversion (IBI)**: extension of DBI, addresses the question of correlation between different interactions by iteratively modifying the CG potentials to quantitatively reproduce a given set of structural distribution functions.
- **Inverse Monte-Carlo**: similar to IBI but employs linear response theory to estimate how variations in the potential impact all of the distributions functions.

CG Models for Membrane

Some Examples

Structure of a water/oil interface in the presence of micelles: a computer simulation study.

- Model: 2 types of particles, o and w.
- Surfactant : combination of the 2 or more o and w particles, connected with springs.

$$U_{ij} = \frac{1}{2}k(|\mathbf{r}_i - \mathbf{r}_j| - \sigma)^2$$

- Potential between particles

and

Smit, B., Hilbers, P.A.J., Esselink, K., Rupert, L.A.M.,
van Os, N.M., Schlijper, A.G. (1991)
J. Phys. Chem. 95:6361 – 6368.
12/14/20

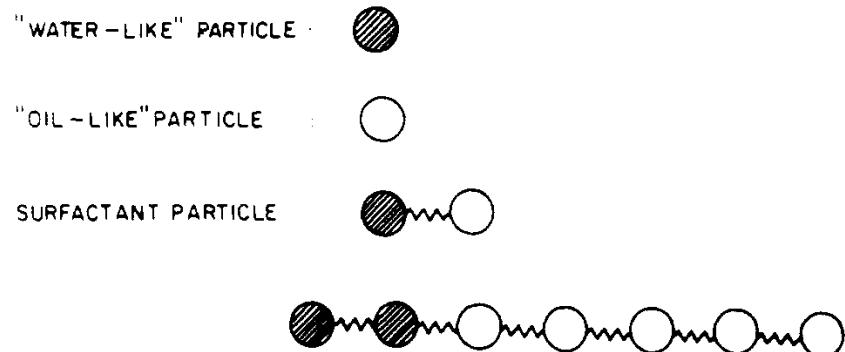


Figure 1. Schematic drawing of the oil/water/surfactant model.

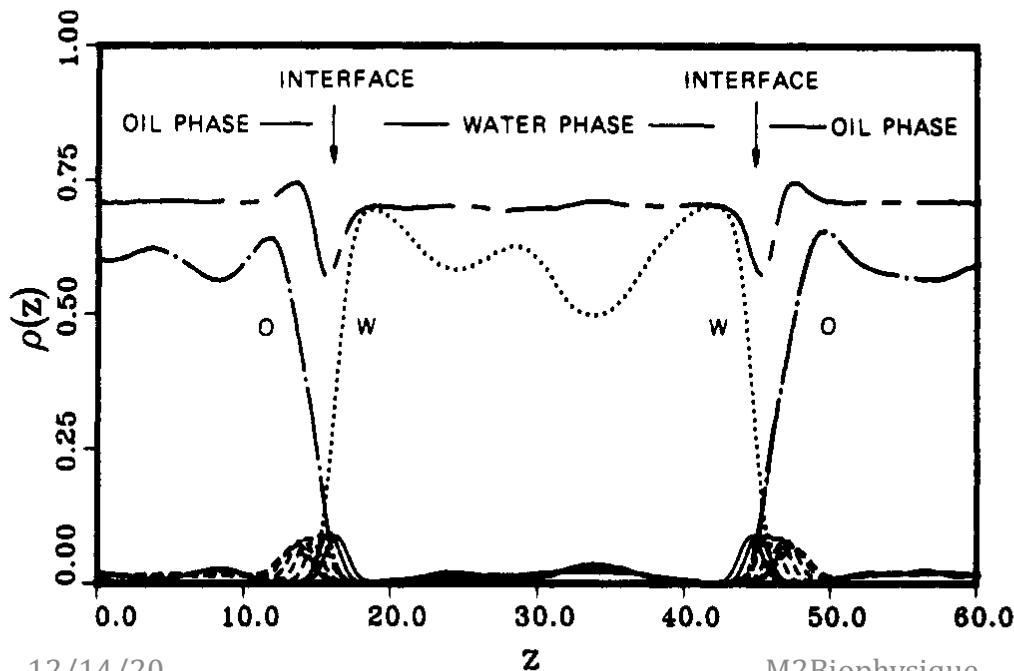
$$\Phi_{ij} = \begin{cases} \phi_{ij}(r) - \phi_{ij}(R_{ij}^c) & r \leq R_{ij}^c \\ 0 & r > R_{ij}^c \end{cases}$$

$$\phi_{ij}(r) = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r} \right)^{12} - \left(\frac{\sigma_{ij}}{r} \right)^6 \right]$$

Results

Tests on different surfactant concentration & different types of surfactants (mostly 5-o, and 2-w).

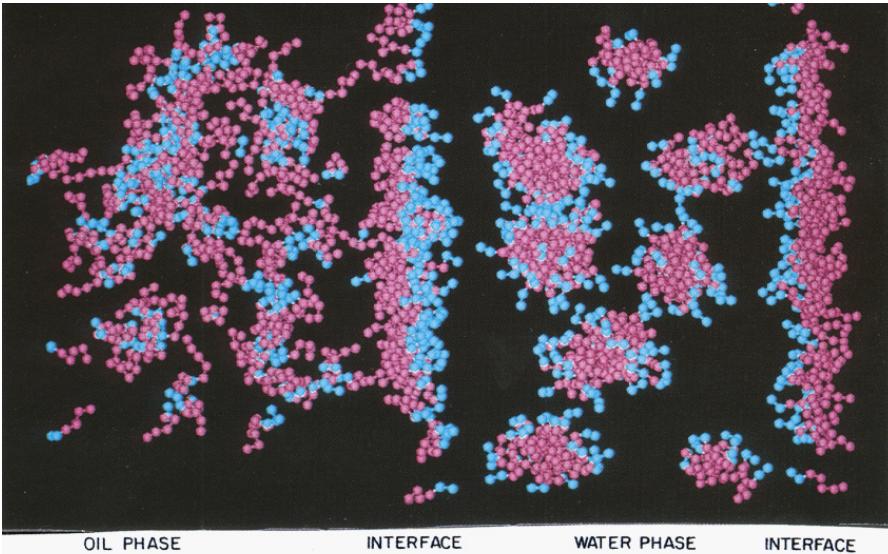
- Simulations at T, for which the oil and water do not mix and form a stable liquid-liquid interface



Density profiles of the oil, water, and surfactant particles for a 3% surfactant solution.

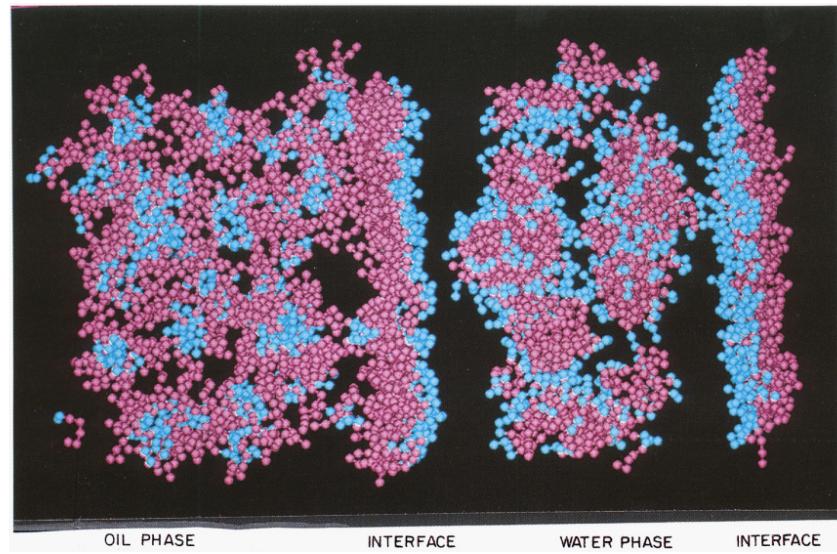
Snapshots

1.5% surfactants



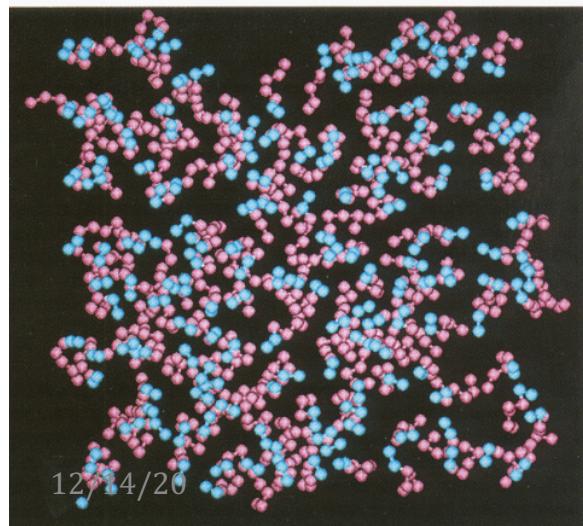
x,z
plane

3%



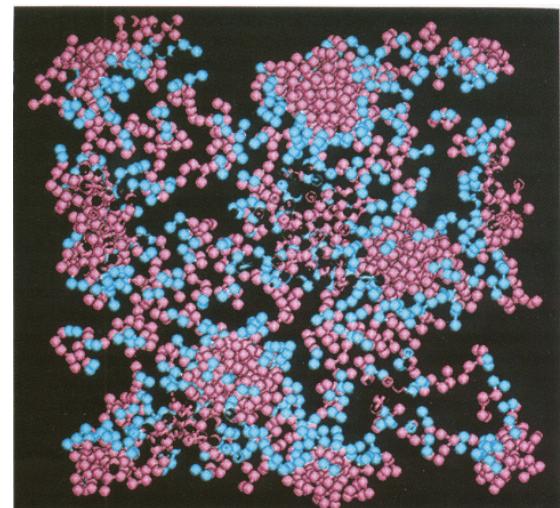
x,y
plane

Figure 4. left only the surfactants in the monolayer at the interface are drawn. In right the surfactants in the water phase



12/34/20

M2Biophysique



About Micelles?

- Clear micelles formations
 - Debate: Role of hydrogen bonding???
- Micelles can occur in fluids that do not form hydrogen bonds!

Conclusions: Model (qualitatively) reproduced experimental observations such as micelle formation, density profiles and order parameters.

Des modèles CG avec DPD

- DPD: Dissipative Particle Dynamics.
 - Briefly: The total force f_i acting on a DPD particle i is given by a sum of three pairwise-additive forces

$$f_i = \sum_{j \neq i} (F_{ij}^C + F_{ij}^D + F_{ij}^R)$$

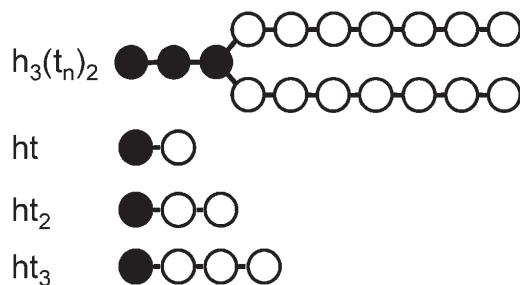
Where

$$F_{ij}^C : \text{conservative force} \quad F_{ij}^C = \begin{cases} a_{ij}(1 - r_{ij}) & r_{ij} < r_c \\ 0 & r_{ij} \geq r_c \end{cases},$$

F_{ij}^D : dissipative acts as a friction

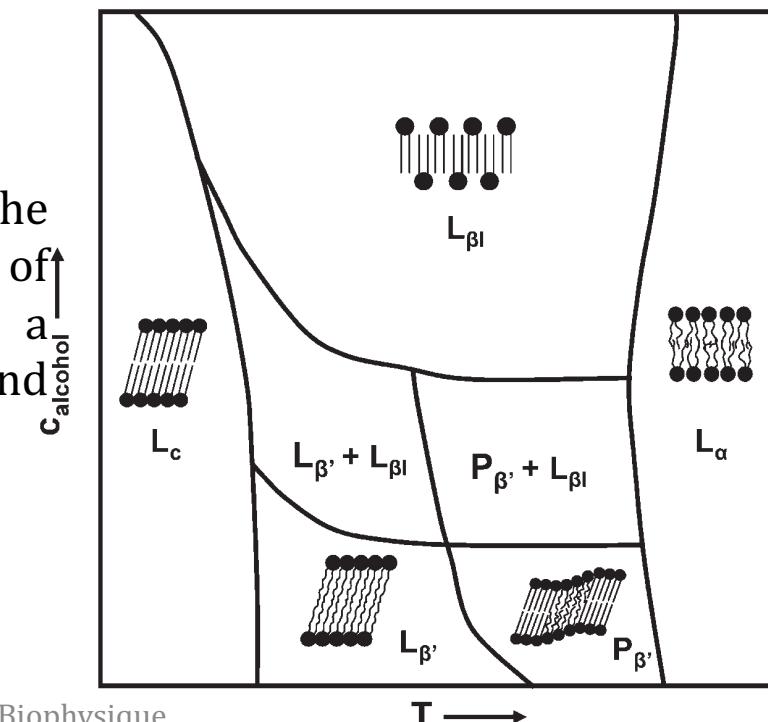
F_{ij}^R : random forces act as a heat pump

Simulating induced interdigitation in membranes. lipid-water-alcohol model,



3 types of particles w, h, and t, to mimic
water and the head- and tail-atoms of a lipid,
respectively
DPD Dynamics

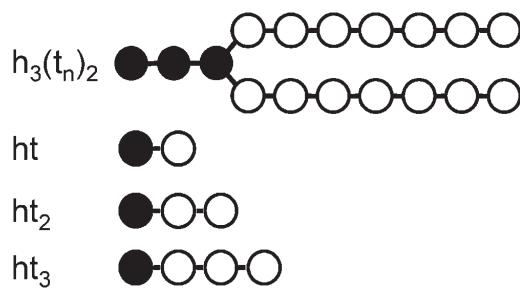
FIGURE 1 Schematic representation of the phase diagram of phosphatidylcholine/alcohol mixture as a function of alcohol concentration and temperature



Kranenburg, M., Vlaar, M., Smit, B. (2004)
Biophys. J. 87:1596 – 1605.

Simulating induced interdigitation in membranes.

lipid-water-alcohol model,



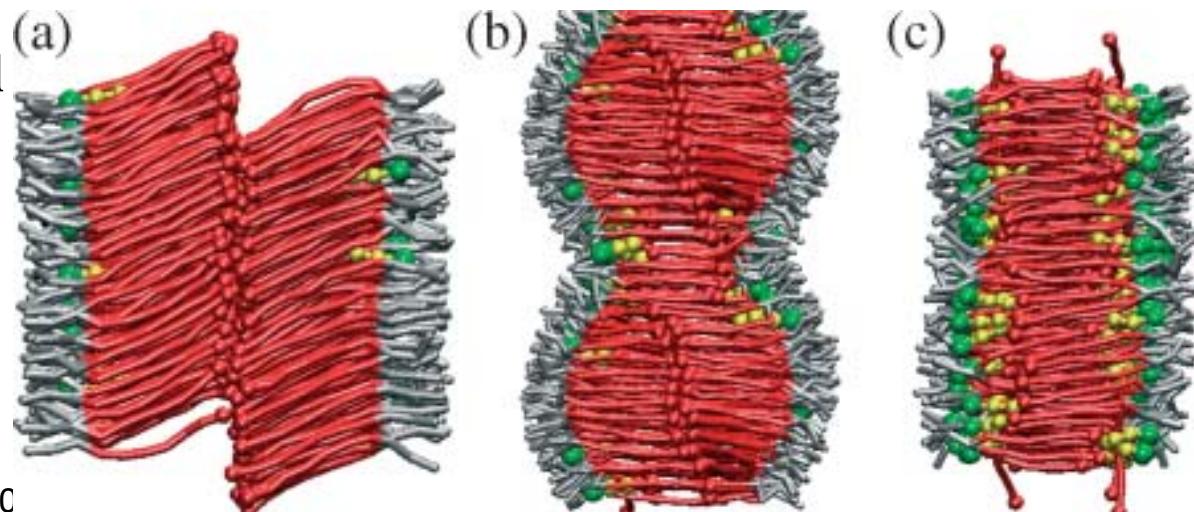
3 types of particles w, h, and t, to mimic water and the head- and tail-atoms of a lipid, respectively
DPD Dynamics

Snapshots of a bilayer consisting of the lipid $h_3(t7)_2$ with various concentrations of the model alcohol $ht2$

(a): low concentration:
noninterdigitated (sub)gel
phase Lc

(c) High concentrations:
interdigitated gel LbI

(b) Between : coexistence

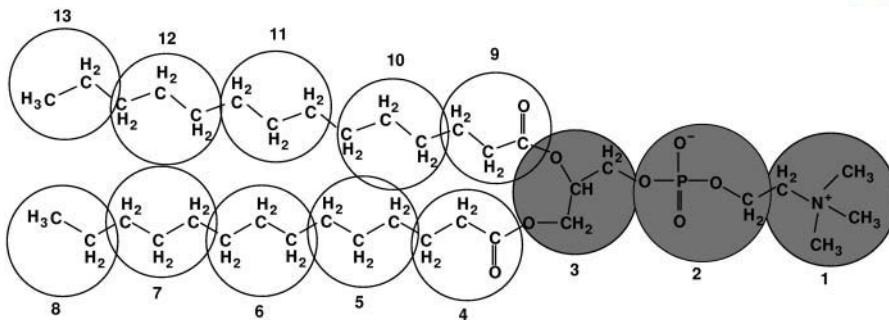


Simulation Studies of Protein-Induced Bilayer Deformations, and Lipid-Induced Protein Tilting, on a Mesoscopic Model for Lipid Bilayers with Embedded Proteins

Mesoscopic model

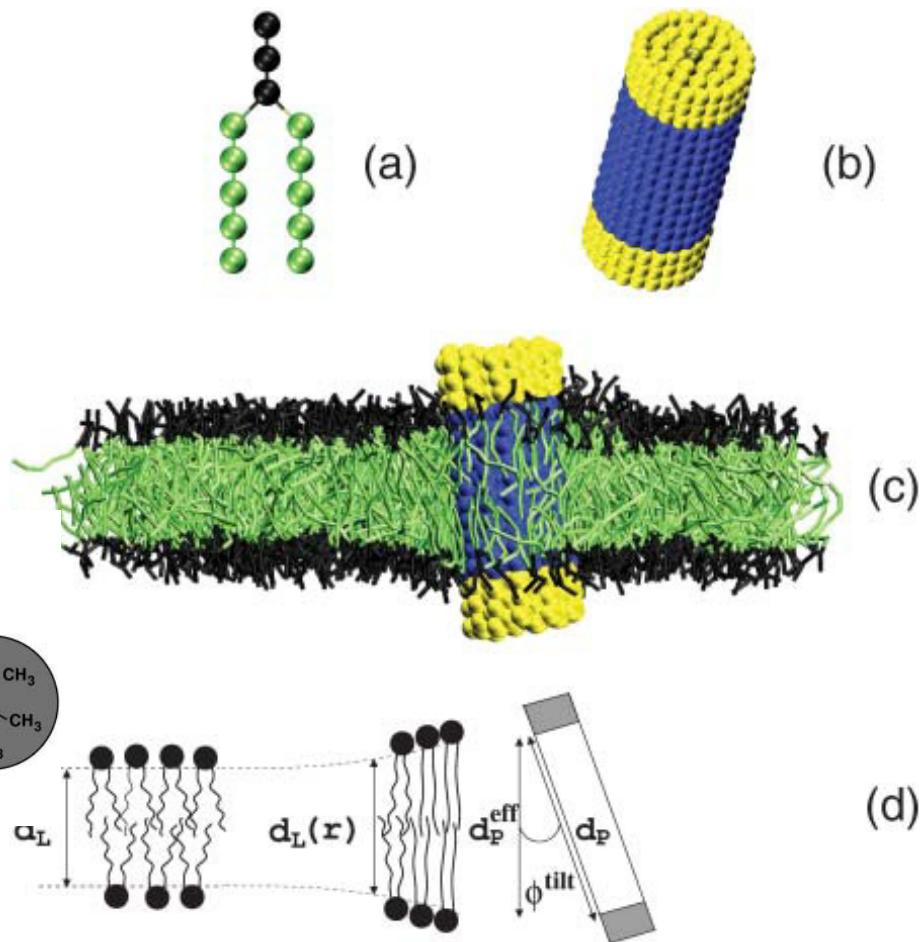
3 types of beads

- W: water-like
- H: hydrophilic
- t_L or t_P: hydrophobic
- L for lipids, P for protein



Venturoli, M., Smit, B., Sperotto, M.M. (2005)

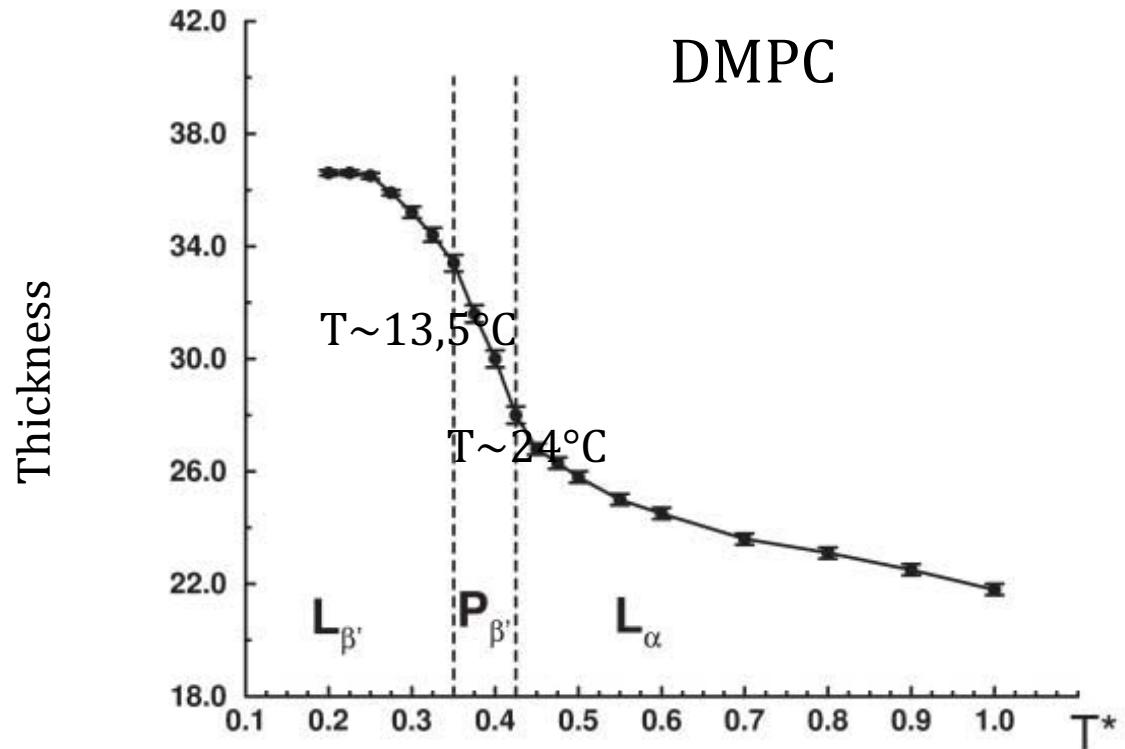
Biophys. J. 88:1778 – 1798.



Protocol

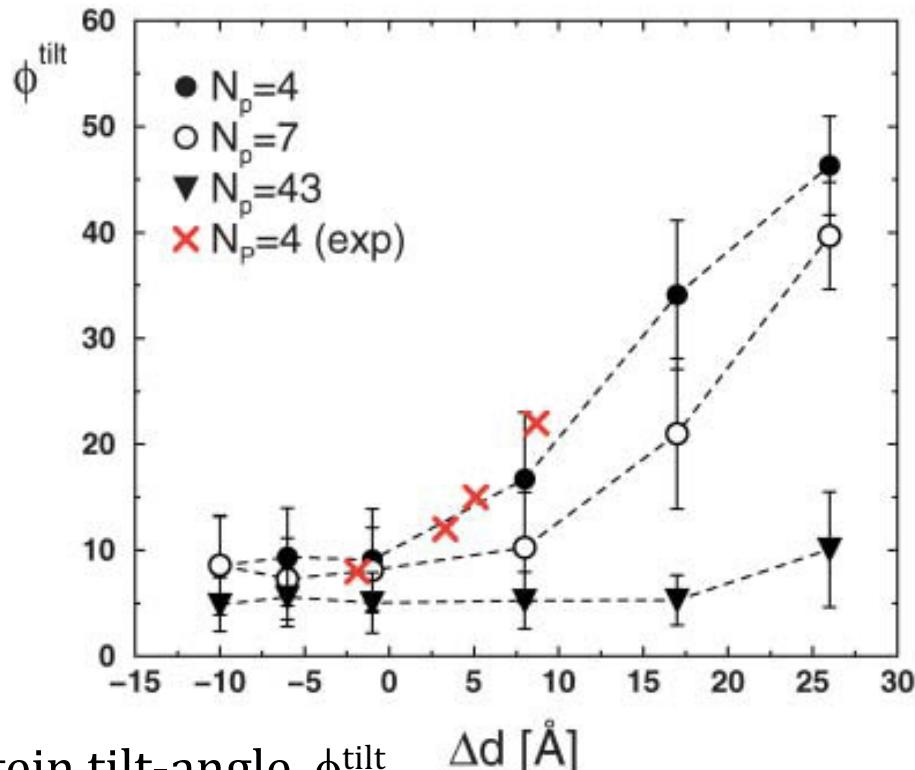
- DPD
- Calibration of all the parameters (see pub. For details), i.e interactions
- Simulation of pure lipid membranes and with embedded protein at different concentrations and temperature

Some tests and results

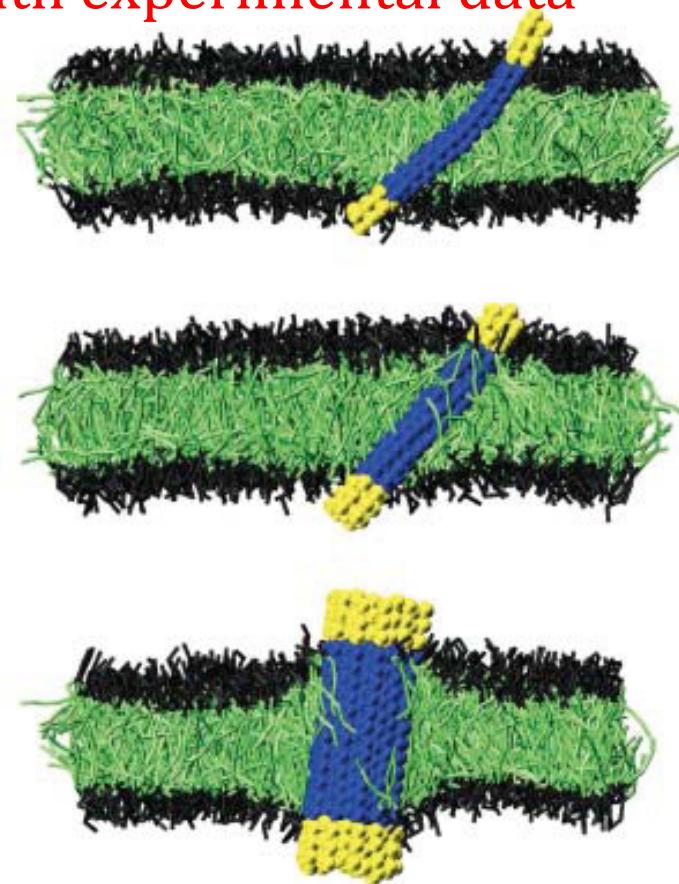


Hydrophobic Mismatch & Protein Size

Dependence in good agreement with experimental data



Protein tilt-angle, ϕ^{tilt} ,
dependence on mismatch,
 $\Delta d = d - d_L^0$



Tilt occurs

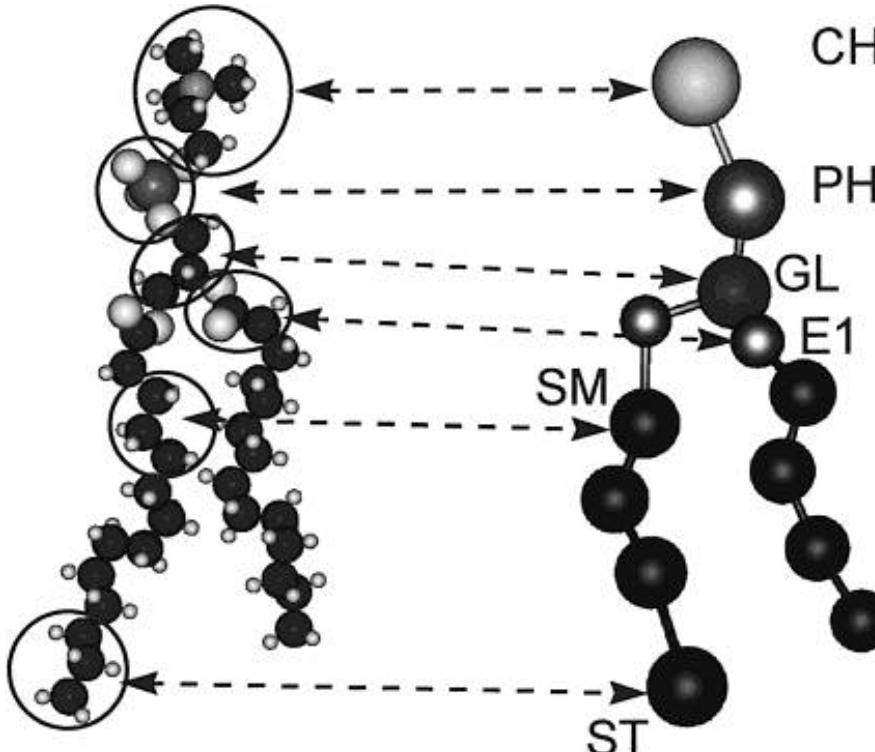
Klein Model

- Explicit (but) incomplete treatment of electrostatic interactions
- Water model: Requirements
 - momentum to be consistent with hydrodynamics
 - correct density at the desired temperature,
 - to maintain a liquid/vapor interface over the desired temperature range (i.e., have a liquid phase for temperatures between 0 and 100 °C).

→ W Bead= 3 water molecules

Shelley, J.C., Shelley, M.Y., Reeder, R.C., Bandyopadhyay, S., Klein, M.L. (2001) A coarse grain model for phospholipid simulations. *J. Phys. Chem. B.* 105:4464 –

Klein's DMPC model



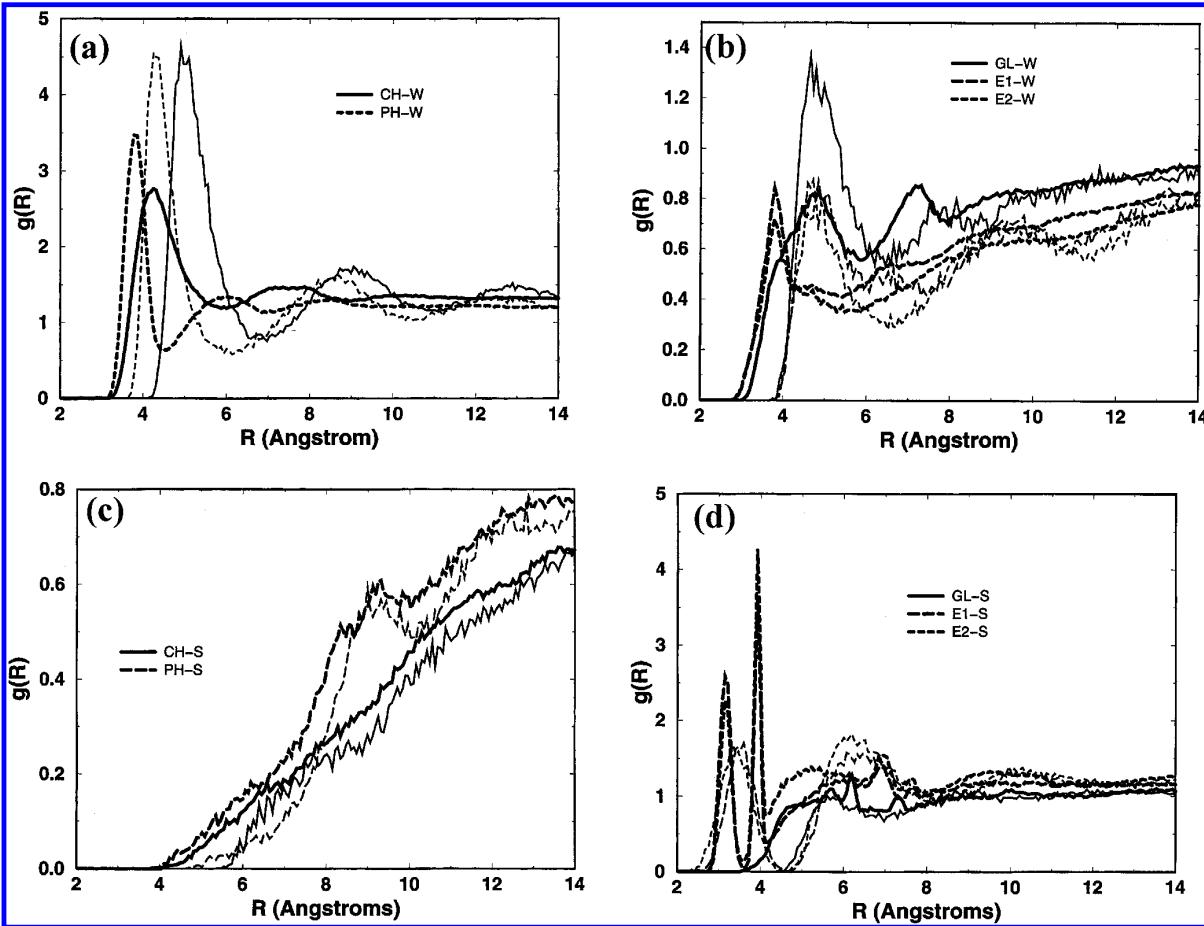
CH : Choline
PH: Phosphate
GL: Glycerol
E1 and E2: ester groups
S(M) or S(T): alkyl chains

The choline and phosphate groups were assigned charges of +e and -e, respectively, and a dielectric constant of 78 was used for all electrostatic interactions

Calibration of the model

- W : on a crude estimate of crystal properties, melting temperature, and vapor pressure
- DMPC: Radidal Distribution Functions obtained from Atomistic distributions

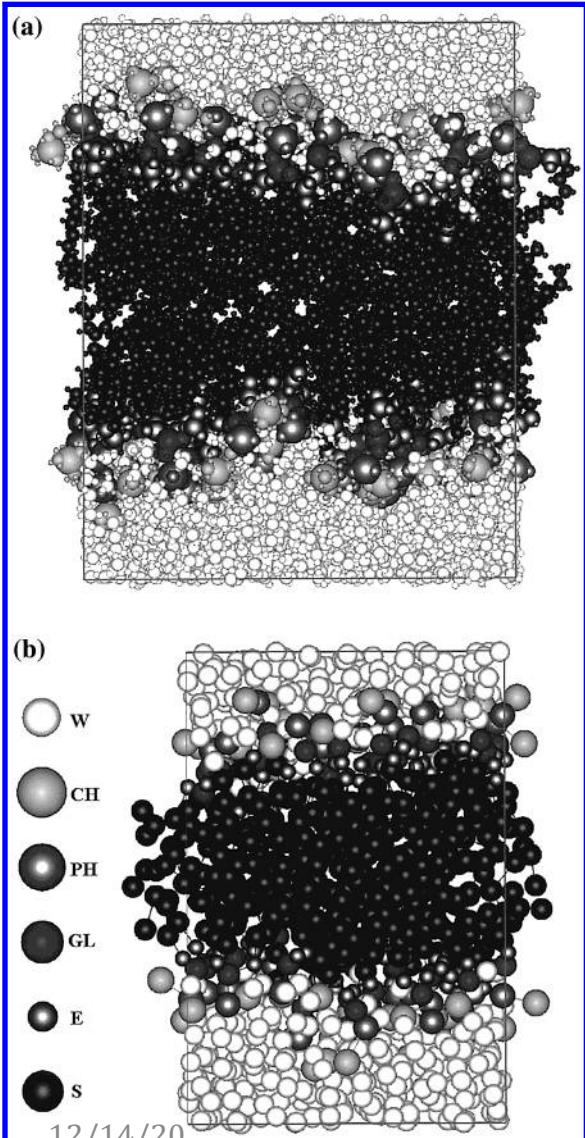
Example



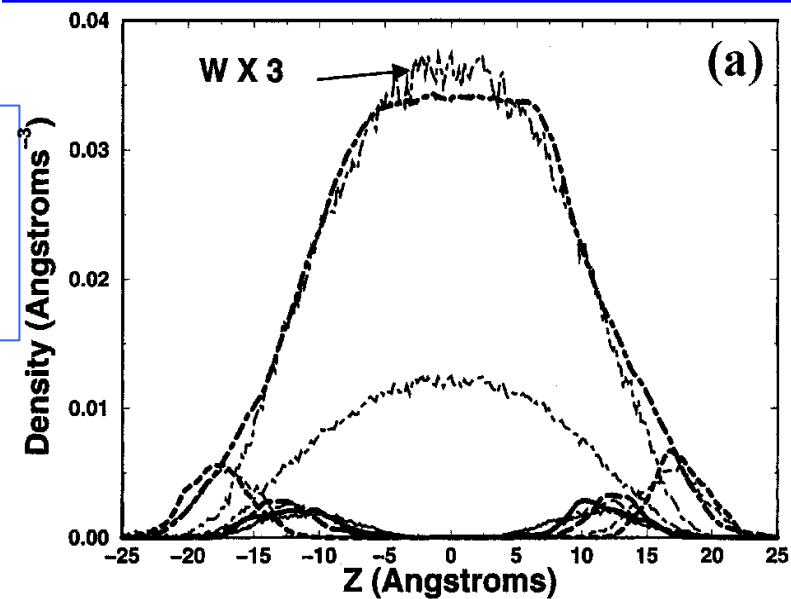
Radial distribution functions for choline, phosphate, glycerol , and ester groups with water and hydrocarbons for the atomistic(thick) and coarse grain (thin) models.

Results

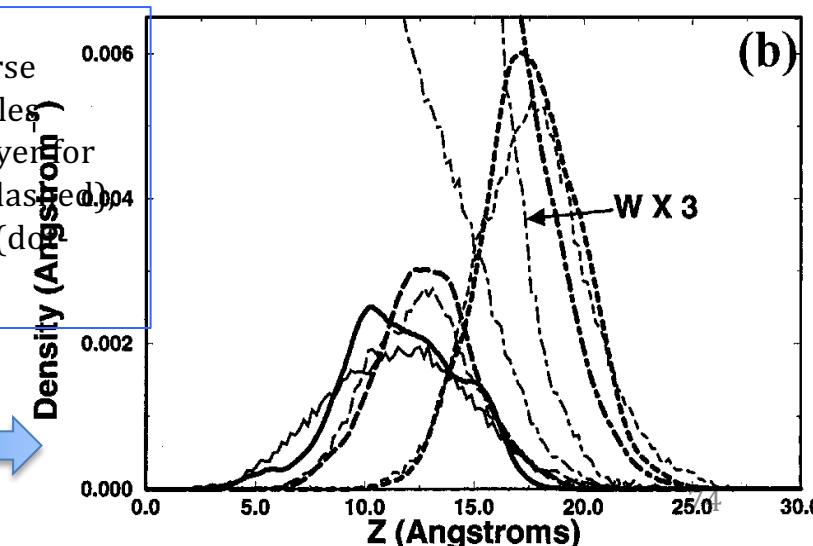
- Starting from a pre-assembled bilayer



Snapshot images of a phospholipid bilayer obtained from atomistic (a) and coarse grain (b) models.

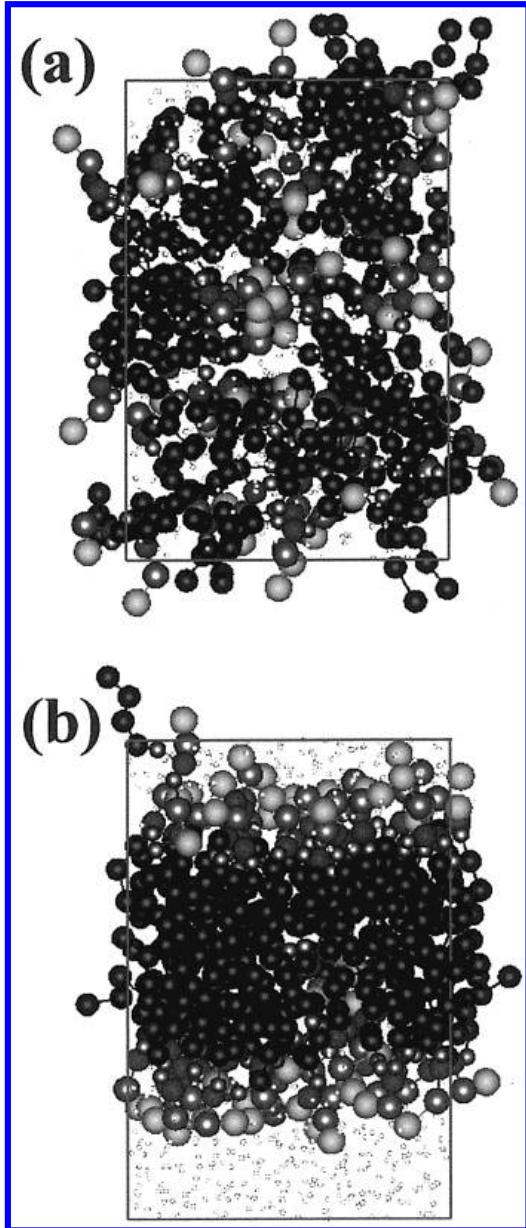


Comparison of the atomistic(thick) and coarse grain(thin) density profiles, perpendicular to the bilayer for the CH (solid), PH (long dashed), E (short dashed), and W (double dashed).



Results

Simulation of a bilayer assembly starting from (a) random configuration (b) to a final configuration

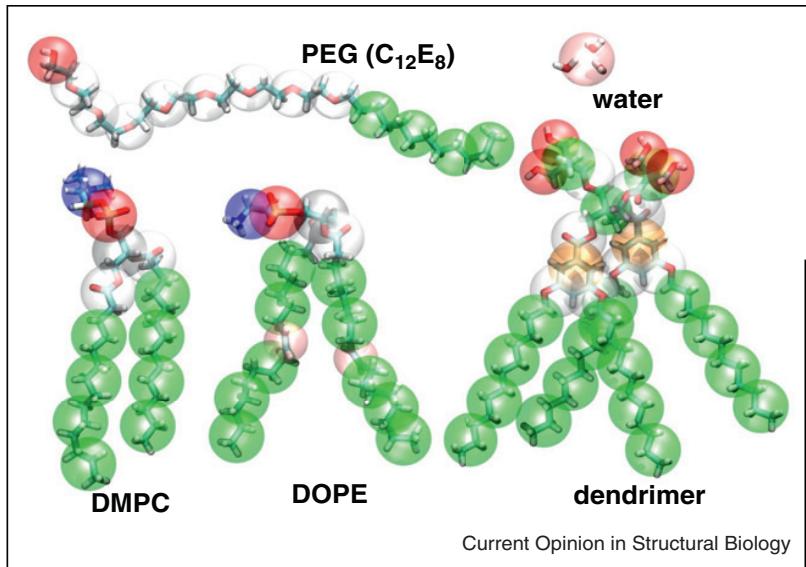


Limits of the model

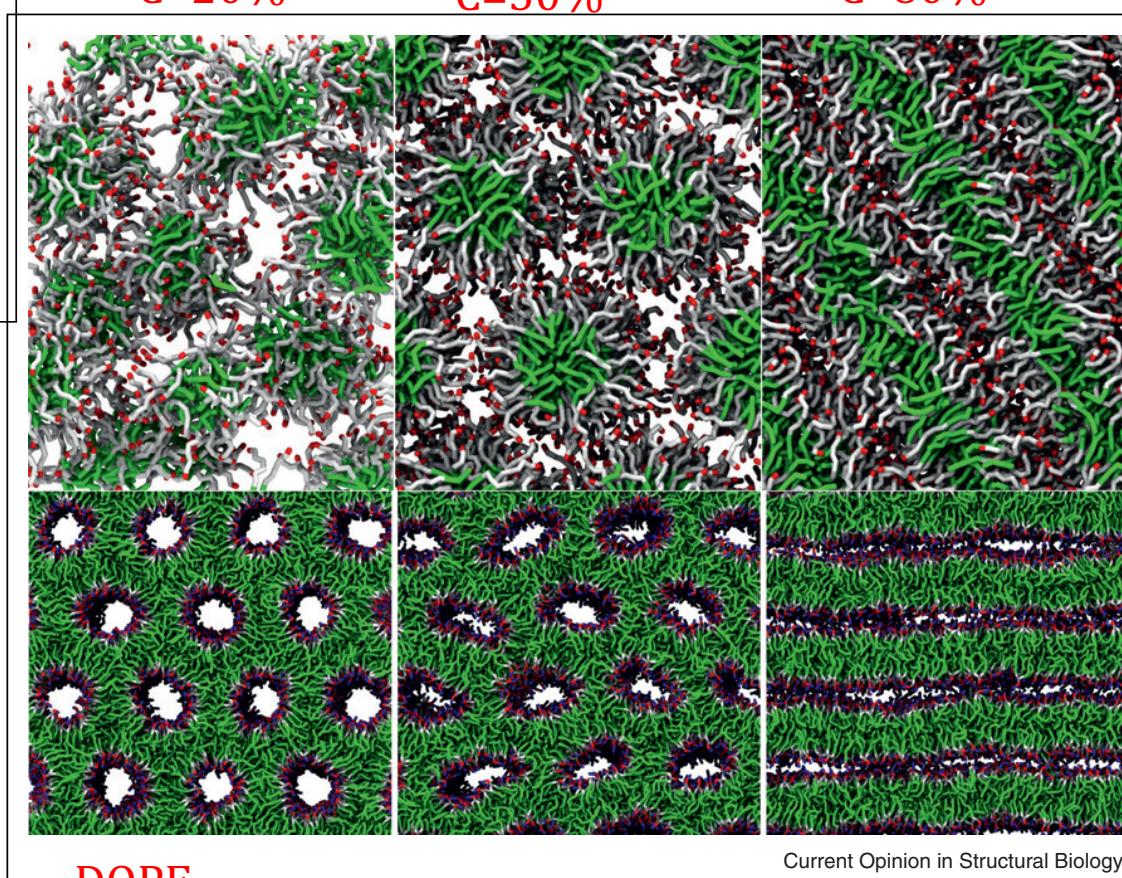
- « CG model was developed to mimic a phospholipid bilayer. The compromises involved in creating this model are not unique and may be specific **for that phase**. Hence, this model may not accurately describe other states, including the nonequilibrium states encountered during self-assembly. In addition, the smoothing out of the bumpiness of the molecules in the current CG models may preclude the reproduction of various solid-solid and solid-liquid transitions which are sensitive to the shape of the hydrocarbon chains. »

Even More

PEG Surfactant



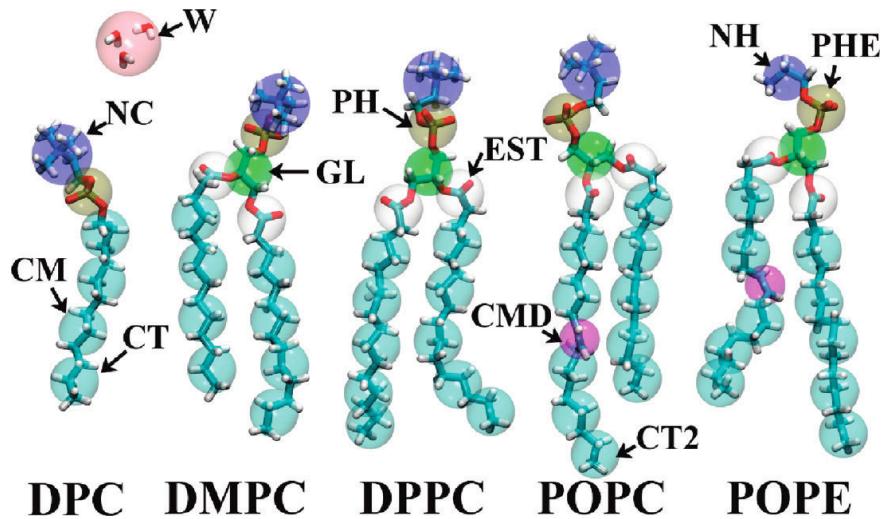
Current Opinion in Structural Biology 2012, 22:175–186



DOPE
Inverse
hexagonal
phase
M2Biophysique

DOPC
Lamellar
phase

CG for Zwitterionic lipids



$$U_{\text{nb}}(r_{ij}) = \begin{cases} \frac{3\sqrt{3}}{2}\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^4 \right] & \text{for the pairs involving W} \\ \frac{27}{4}\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^9 - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^6 \right] & \text{for any other pairs} \end{cases}$$

Coarse-Grained Model for Lipid Membranes J.
Phys. Chem. B, Vol. 114, No. 20, 2010 6847

12/14/20

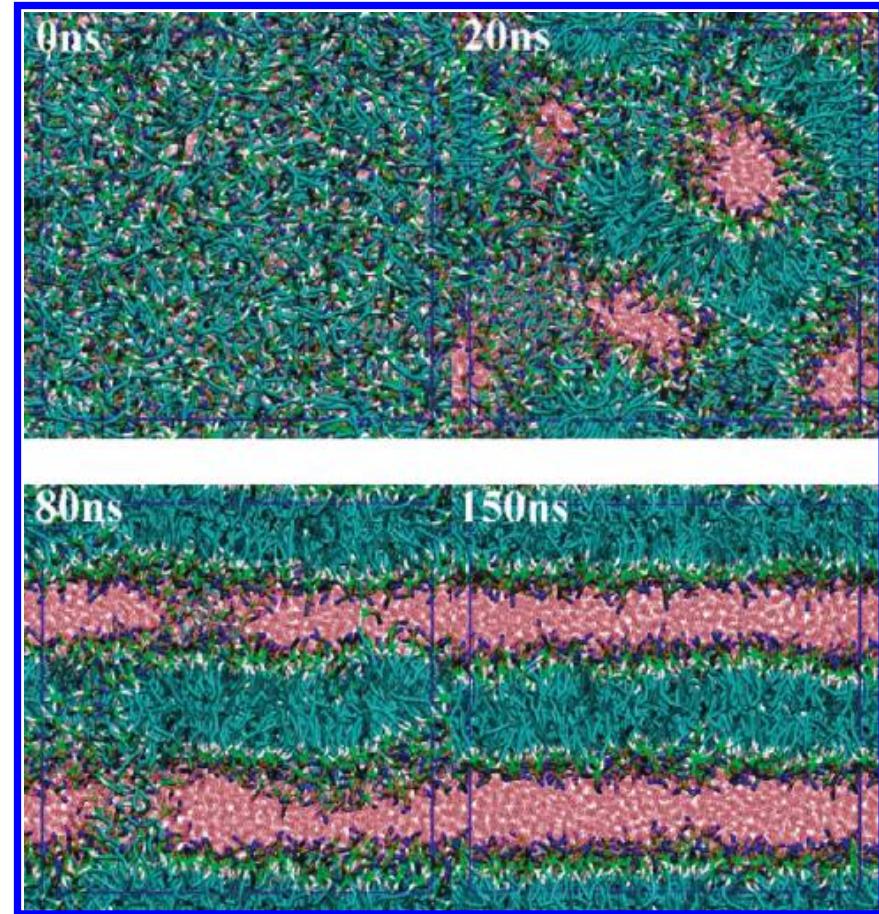


Figure 5. Formation of lamellar structure from a random initial configuration of DMPC/water. Color codes are the same as in Figure 1. Blue lines denote the simulation box.

buckling of the monolayer

Buds start to grow

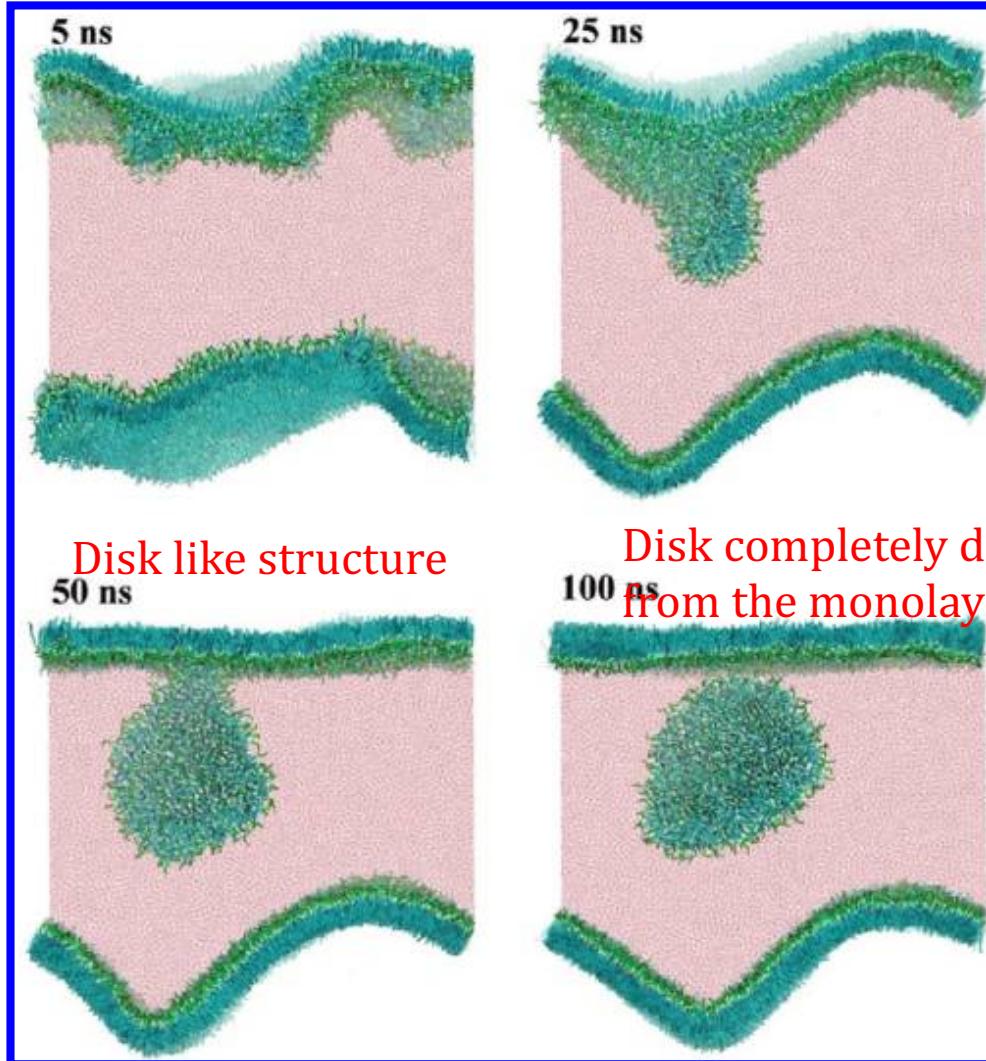


Figure 8. Budding and fission from an oversaturated DPPC monolayer at the air/water interface. The initial area per DPPC molecule on the flat monolayer at the air/water interface was set to 0.40 nm^2 . Color codes are the same as in Figure 1.

Larger and larger assemblies

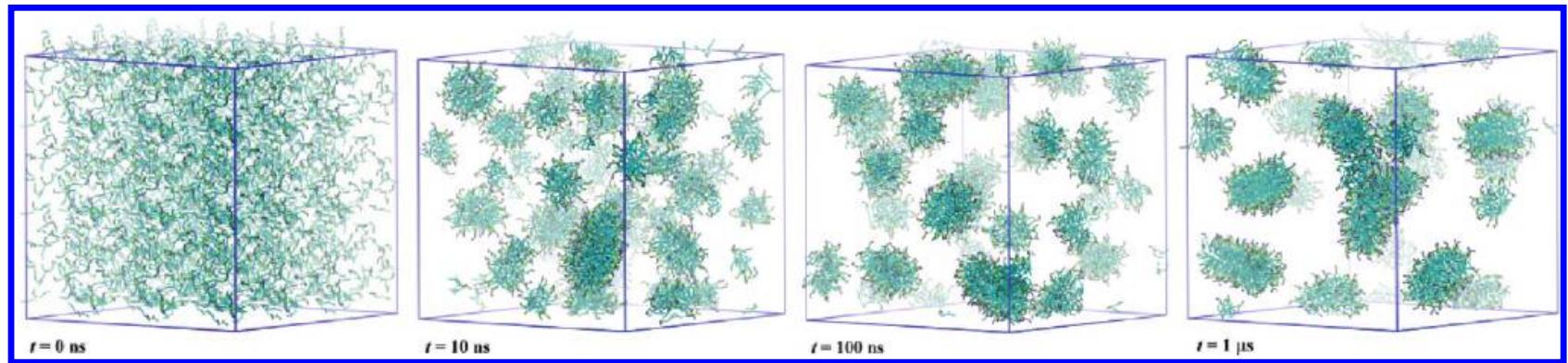


Figure 10. Liposome formation from a DMPC bilayer fragment. Water is not shown for clarity. Color codes are the same as in Figure 1.

Transformation of self-assembled lipid aggregates in a large water box. The five rows represent lipid aggregates with 1000, 1512, 2500, 3500, and 5000 DMPC molecules, respectively.

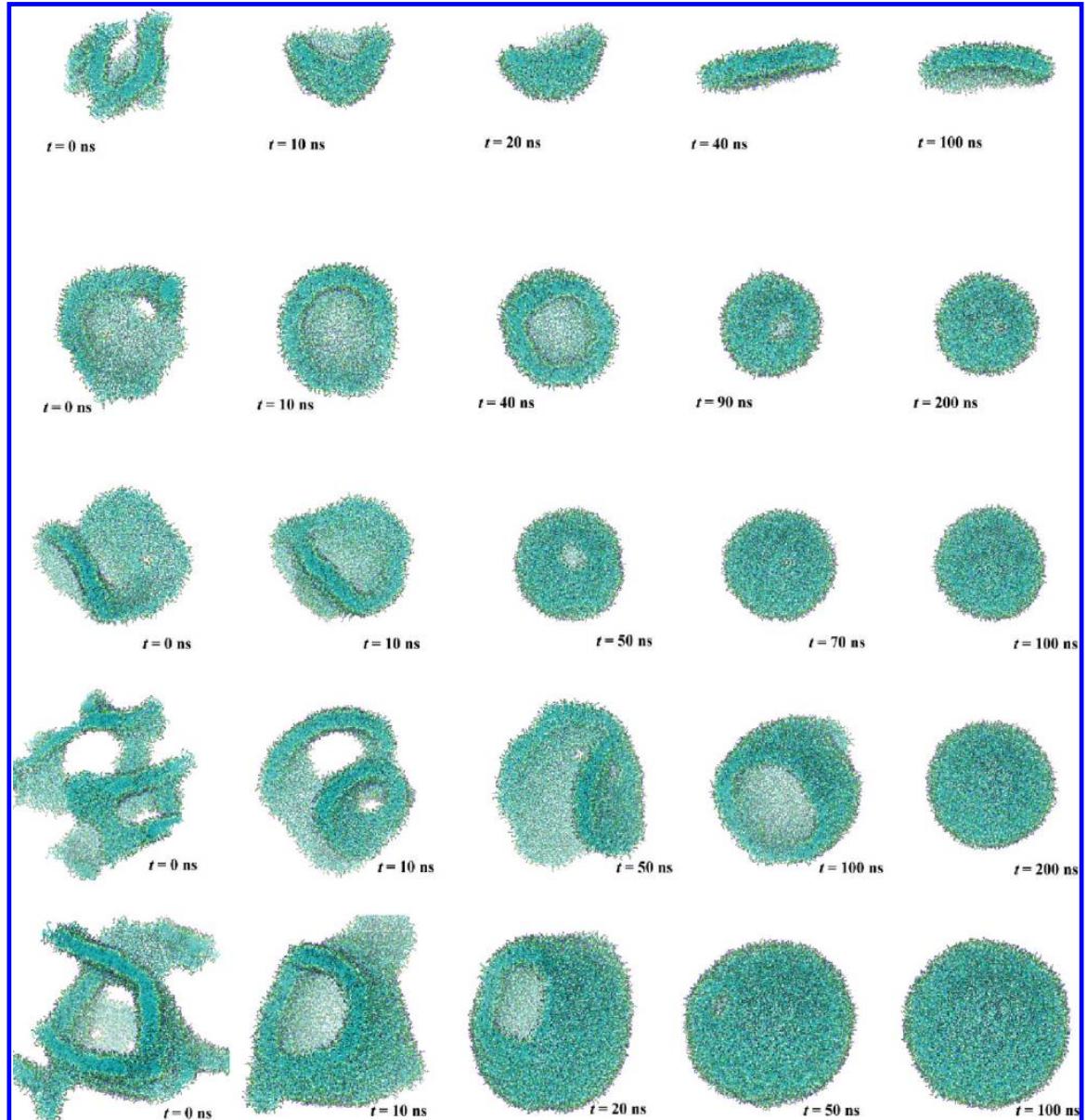
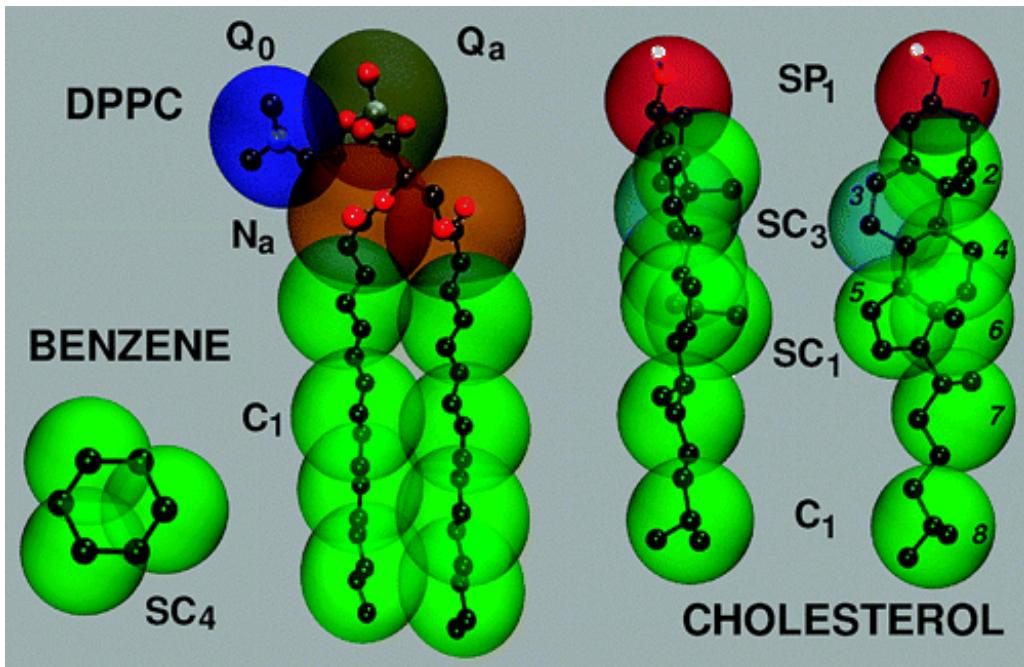


Figure 11. Transformation of self-assembled lipid aggregates in a large water box. The five rows represent lipid aggregates with 1000, 1512, 2500, 3500, and 5000 DMPC molecules, respectively. Color codes are the same as in Figure 1.

One of the most popular model Martini CG for lipids and small organic molecules

Marrink *et al.*, 2007

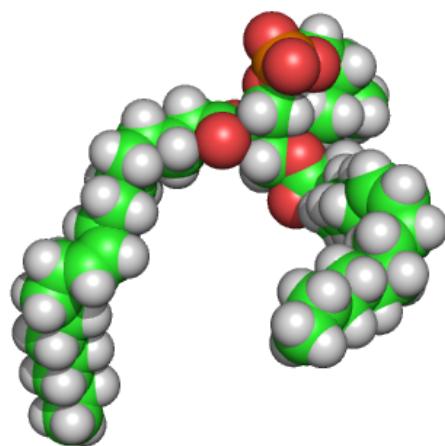


- On average 4 heavy atoms \leftrightarrow 1 site
- 4 types of interaction sites: polar (P), nonpolar (N), apolar (C), and charged (Q)
- Subtypes of interaction sites (18):
 - HB: d; a,da, none
 - Polarity: 1 (low) to 5 (high)

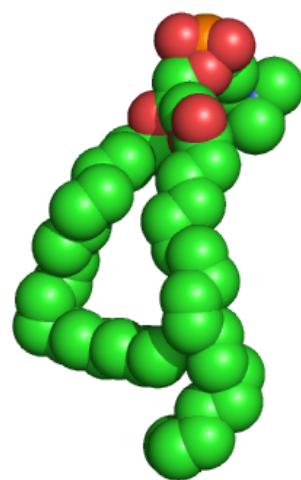
– allow time steps of 40 fs ! (*but faster kinetics*)

For comparison

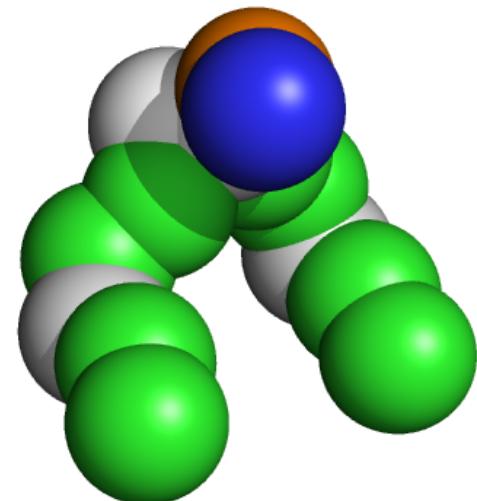
(DOPC = dioleoylphosphatidylcholine)



all-atom
138 particles
(CHARMM)



united-atom
54 particles
(Berger)



coarse-grained
14 particles
(MARTINI)

Martini Model

- Bonded interactions: bond and angles.
- Rings: More sites are required 2-3 hvy atoms/site
 - New particle : S with some reduced parameters
- Water: 4 molecules. Introduction of antifreeze particle BP4. And BP4-P4 interactions.

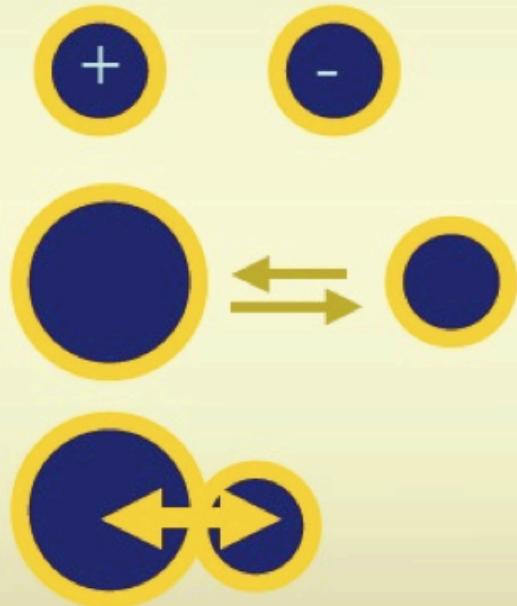


Bonded interactions:
harmonic bonds and angles

Coulomb: only for ions!

$\epsilon_{\text{rel}} = 15$, $r_{\text{cut}} = 1.2 \text{ nm}$, $r_{\text{shift}} = 0.0 \text{ nm}$

New development: PME



} Lennard-Jones (6-12)
 $r_{\text{cut}} = 1.2 \text{ nm}$, $r_{\text{shift}} = 0.9 \text{ nm}$

The interaction Matrix

TABLE 1: Interaction Matrix^a

		Q				P					N				C				
	sub	da	d	a	0	5	4	3	2	1	da	d	a	0	5	4	3	2	1
Q	da	O	O	O	II	O	O	O	I	I	I	I	I	IV	V	VI	VII	IX	IX
	d	O	I	O	II	O	O	O	I	I	I	III	I	IV	V	VI	VII	IX	IX
	a	O	O	I	II	O	O	O	I	I	I	I	III	IV	V	VI	VII	IX	IX
	0	II	II	II	IV	I	O	I	II	III	III	III	III	IV	V	VI	VII	IX	IX
P	5	O	O	O	I	O	O	O	O	O	I	I	I	IV	V	VI	VI	VII	VIII
	4	O	O	O	O	O	I	I	II	II	III	III	III	IV	V	VI	VI	VII	VIII
	3	O	O	O	I	O	I	I	II	II	II	II	II	IV	IV	V	V	VI	VII
	2	I	I	I	II	O	II	II	II	II	II	II	II	III	IV	IV	V	VI	VII
N	1	I	I	I	III	O	II	II	II	II	II	II	II	III	IV	IV	IV	V	VI
	da	I	I	I	III	I	III	II	II	II	II	II	II	IV	IV	V	VI	VI	VI
	d	I	III	I	III	I	III	II	II	II	II	III	II	IV	IV	V	VI	VI	VI
	a	I	I	III	III	I	III	II	II	II	II	II	III	IV	IV	V	VI	VI	VI
C	0	IV	IV	IV	IV	IV	IV	IV	III	III	IV	IV	IV	IV	IV	IV	IV	V	VI
	5	V	V	V	V	V	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	V	V
	4	VI	VI	VI	VI	VI	VI	V	IV	IV	V	V	V	IV	IV	IV	IV	V	V
	3	VII	VII	VII	VII	VI	VI	V	V	IV	VI	VI	VI	IV	IV	IV	IV	IV	IV
2	2	IX	IX	IX	IX	VII	VII	VI	VI	V	VI	VI	VI	V	V	V	IV	IV	IV
	1	IX	IX	IX	IX	VIII	VIII	VII	VII	VI	VI	VI	VI	V	V	V	IV	IV	IV

Roman Numbers represent the 10 levels of interaction: well depth in the LJ potential net charges for Q sites

Calibration

- On thermodynamic Properties:
 - Free Energies of Vaporization, Hydration, and Partitioning.

TABLE 3: Thermodynamic Properties of the CG Particle Types^a

type	building block	examples	ΔG^{vap}		ΔG^{hyd}		$\Delta G_{\text{HW}}^{\text{part}}$		$\Delta G_{\text{CW}}^{\text{part}}$		$\Delta G_{\text{EW}}^{\text{part}}$		$\Delta G_{\text{OW}}^{\text{part}}$		
			exp	CG	exp	CG	exp	CG	exp	CG	exp	CG	exp	CG	
Q _{da}	H ₃ N ⁺ —C ₂ —OH	ethanolamine (protonated)					-25		< -30		-18		-13	-18	
Q _d	H ₃ N ⁺ —C ₃	1-propylamine (protonated)					-25		< -30		-18		-13	-18	
	NA ⁺ OH	sodium (hydrated)					-25		< -30		-18		-13	-18	
Q _a	PO ₄ ⁻	phosphate					-25		< -30		-18		-13	-18	
	CL ⁻ HO	chloride (hydrated)					-25		< -30		-18		-13	-18	
Q ₀	C ₃ N ⁺	choline					-25		< -30		-18		-13	-18	
P ₅	H ₂ N—C ₂ =O	acetamide	sol	sol	-40	-25	-27	-28	(-20)	-18	-15	-13	-8	-10	
P ₄	HOH (× 4)	water	-27	-18	-27	-18	-25	-23		-14	-10	-7	-8	-9	
	HO—C ₂ —OH	ethanediol	-35	-18	-33	-18	-21	-23		-14		-7	-8	-9	
P ₃	HO—C ₂ =O	acetic acid	-31	-18	-29	-18	-19	-21	-9	-10	-2	-6	-1	-7	
	C—NH—C=O	methylformamide	-35	-18		-18			-21		-10		-6	-5	-7
P ₂	C ₂ —OH	ethanol	-22	-16	-21	-14	-13	-17	-5	-2	-3	1	-2	-2	
P ₁	C ₃ —OH	1-propanol	-23	-16	-21	-14	-9	-11	-2	-2	0	1	1	-1	
		2-propanol	-22	-16	-20	-14	-10	-11	-2	-2	-1	1	0	-1	
N _{da}	C ₄ —OH	1-butanol	-25	-16	-20	-9	-5	-7	2	0	4	2	4	3	
N _d	H ₂ N—C ₃	1-propylamine	-17	-13	-18	-9	(-6)	-7	(1)	0	(-3)	2	(3)	3	
N _a	C ₃ =O	2-propanone	-17	-13	-16	-9	-6	-7	1	0	-1	2	-1	3	
	C—NO ₂	nitromethane	-23	-13	-17	-9	-6	-7				2	-2	3	
	C ₃ =N	propionitrile	-22	-13	-17	-9	-5	-7				2	1	3	
	C—O—C=O	methylformate	-16	-13	-12	-9	(-6)	-7	(4)	0	(-1)	2	(0)	3	
	C ₂ HC=O	propanal		-13	-15	-9	-4	-7		0	2	2	3	3	
N ₀	C—O—C ₂	methoxyethane	-13	-10	(-8)	-2	(1)	-2		6	(3)	6	(3)	5	
C ₅	C ₃ —SH	1-propanethiol	-17	-10		1		5		10		10		6	
	C—S—C ₂	methyl ethyl sulfide	-17	-10	-6	1	(7)	5		10		10	(9)	6	
C ₄	C ₂ =C ₂	2-butyne	-15	-10	-1	5		9		13		13	9	9	
	C=C—C=C	1,3-butadiene		-10	2	5	11	9		13		13	11	9	
	C—X ₄	chloroform	-18	-10	-4	5	(7)	9	14	13		13	11	9	
C ₃	C ₂ =C ₂	2-butene		-10		5		13		13		13	13	14	
	C ₃ —X	1-chloropropane	-16	-10	-1	5	12	13		13		13	12	14	
		2-bromopropane	-16	-10	-2	5		13		13		13	12	14	
C ₂	C ₃	propane	gas	-10	8	10		16		15		14	14	16	
C ₁	C ₄	butane	-11 ^b	-10	9	14	18	18		18		14	16	17	
		isopropane	gas	-10	10	14		18		18		14	16	17	

Marrink's Recipe

- Step I, Mapping onto CG Representation.
 - dividing the molecule into small chemical building blocks, ideally of four heavy atoms each.
- Step II, Selecting Bonded Interactions.
 - For most molecules: standard bond length (0.47 nm) and force constant of Kbond) 1250 kJ mol⁻¹ nm⁻² work well.
- Step III, Optimization:
 - not have to lead to a unique assignment of particle types and bonded interactions. A powerful way to improve the model is by comparison to AA level simulations,

Martini For Amino Acids

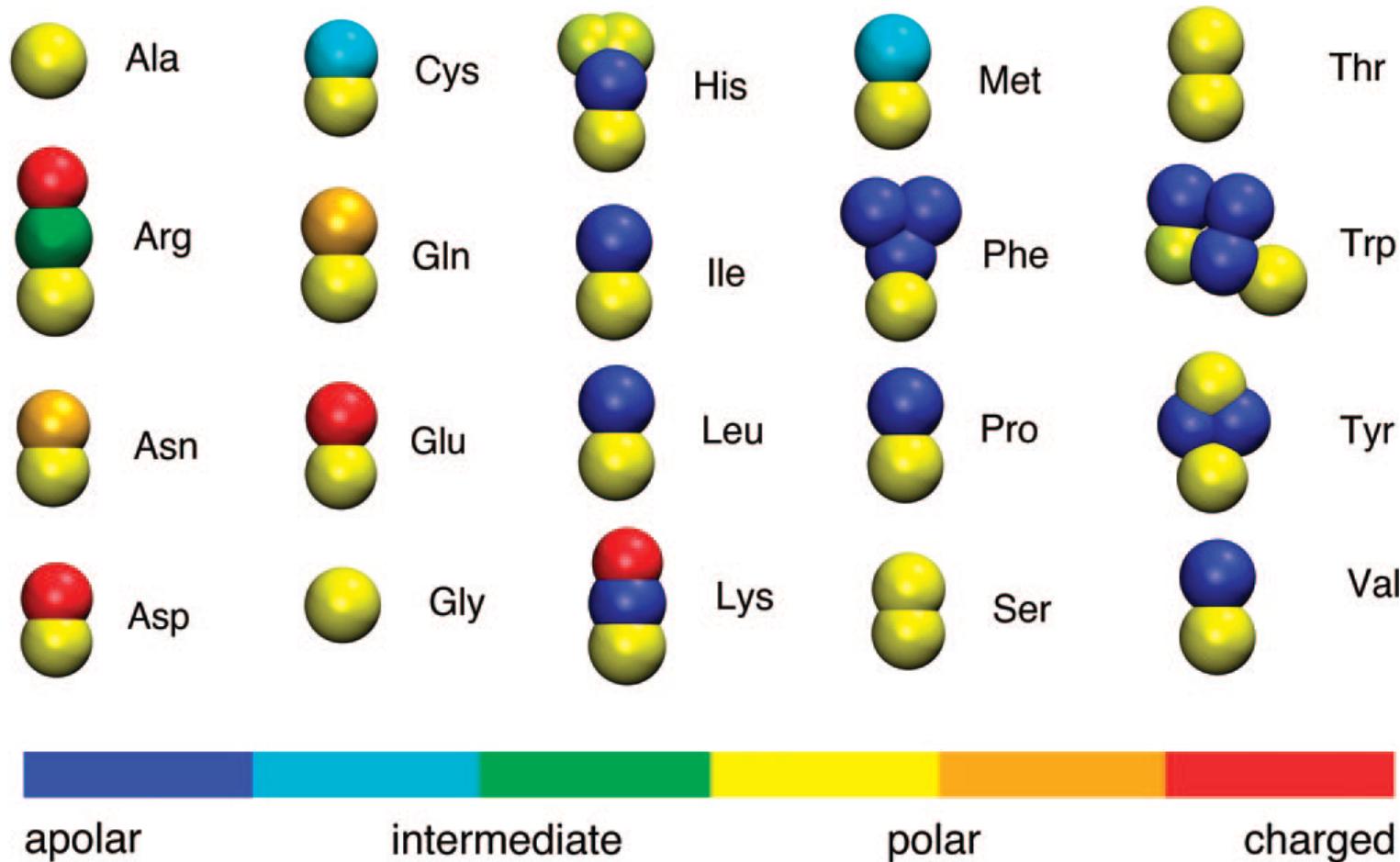


Figure 1. Coarse-grained representation of all amino acids.

Different colors represent different particle types.

Parametrization: non bonded interactions

* Lennard-Jones

$$V_{\text{Lennard-Jones}}(r_{ij}) = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$$

ϵ_{ij} = 5.6 kJ/mol (polar-polar) \Leftrightarrow ϵ_{ij} = 2.0 kJ/mol (polar-apolar) \rightarrow
mimicks the hydrophobic effect

σ = 0.47 nm (normal particle types)

\rightarrow effective size of the particle

* Electrostatics

$$V_{\text{el}} = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_{\text{rel}} r_{ij}}$$

ϵ_{rel} = 15 for explicit screening
 $q_i = \pm 1$ only for net charge

Parametrization: partitioning of the Amino-Acids

$$\Delta G^{\text{oil/aq}} = kT \ln\left(\frac{\rho_{\text{oil}}}{\rho_{\text{aq}}}\right)$$

Table 1. Mapping of the Amino Acids and Free Energy of Partitioning between Water and Butane (Calculated) or Water and Cyclohexane (Experimental Measure^{40,41})

side chain	CG representation	mapping scheme ^a	free energy (kJ/mol)	
			CG	exptl.
Leu	C1 ^b		22	22
Ile	C1 ^b		22	22
Val	C2 ^b		20	17
Pro	C2 ^b		20	
Met	C5		9	10
Cys	C5		9	5
Ser	P1		-11	-14
etc...				

Parametrization: bonded interactions

Bond

$$V_b = \frac{1}{2} K_b (d_{ij} - d_b)^2$$

→ Flexible molecule

Angle

$$V_a = \frac{1}{2} K_a [\cos(\varphi_{ijk}) - \cos(\varphi_a)]^2$$

Proper
Dihedral

$$V_d = K_d [1 + \cos(n\psi_{ijkl} - \psi_d)]$$

→ Impose Secondary structure

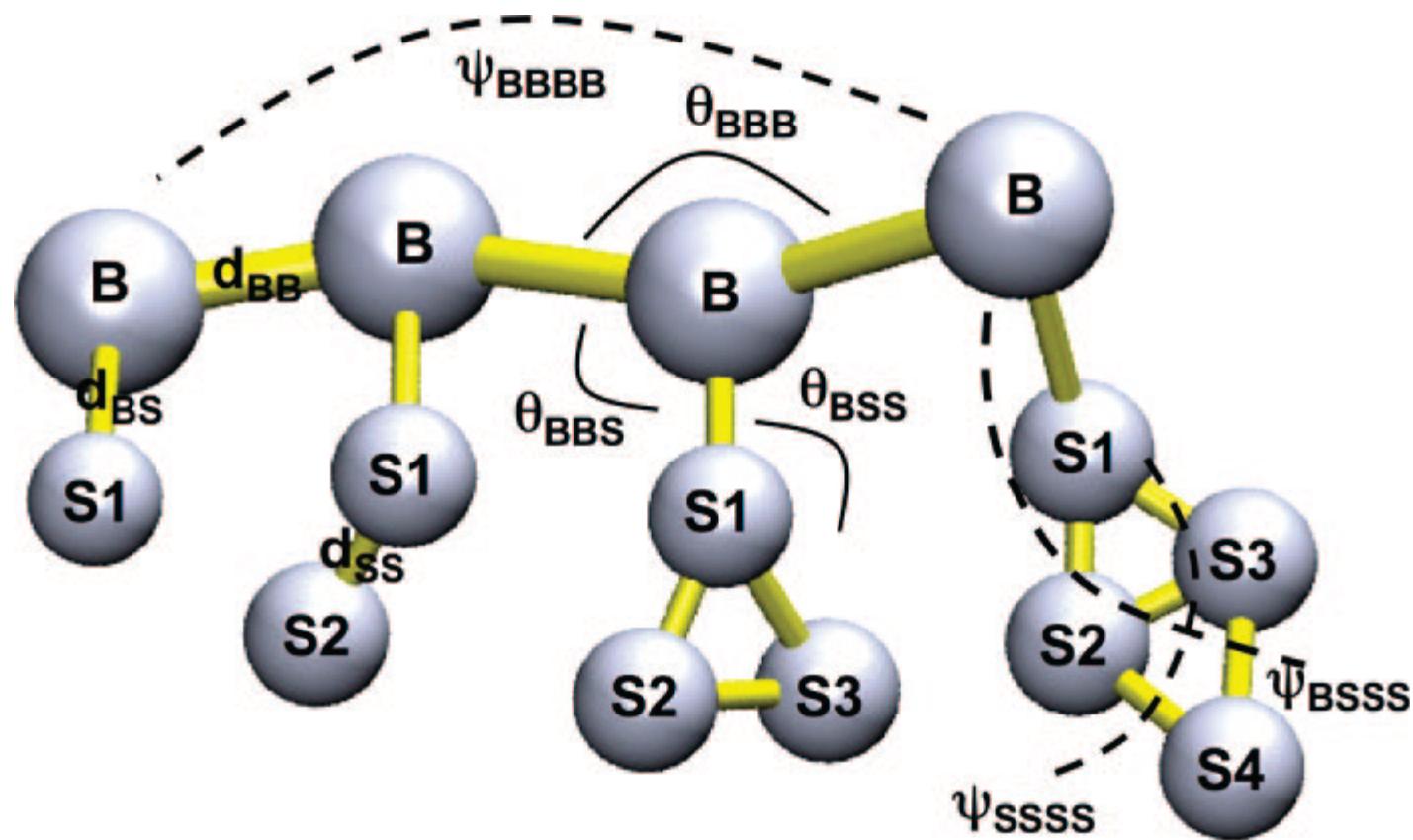
Improper
Dihedral

$$V_{id} = K_{id} (\psi_{ijkl} - \psi_{id})^2$$

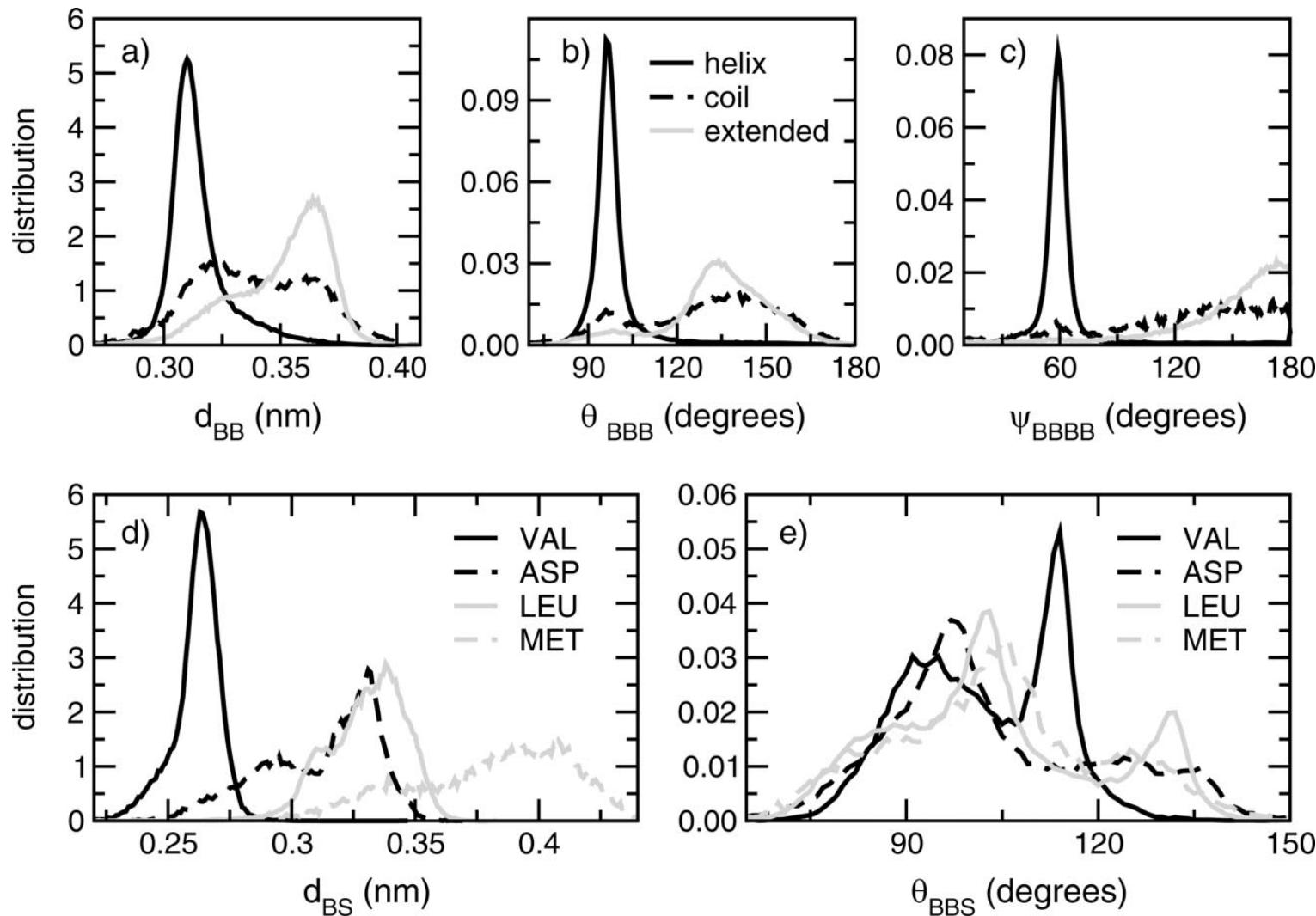
→ For rings only

→ Limitation : conformational change not adequately modelled

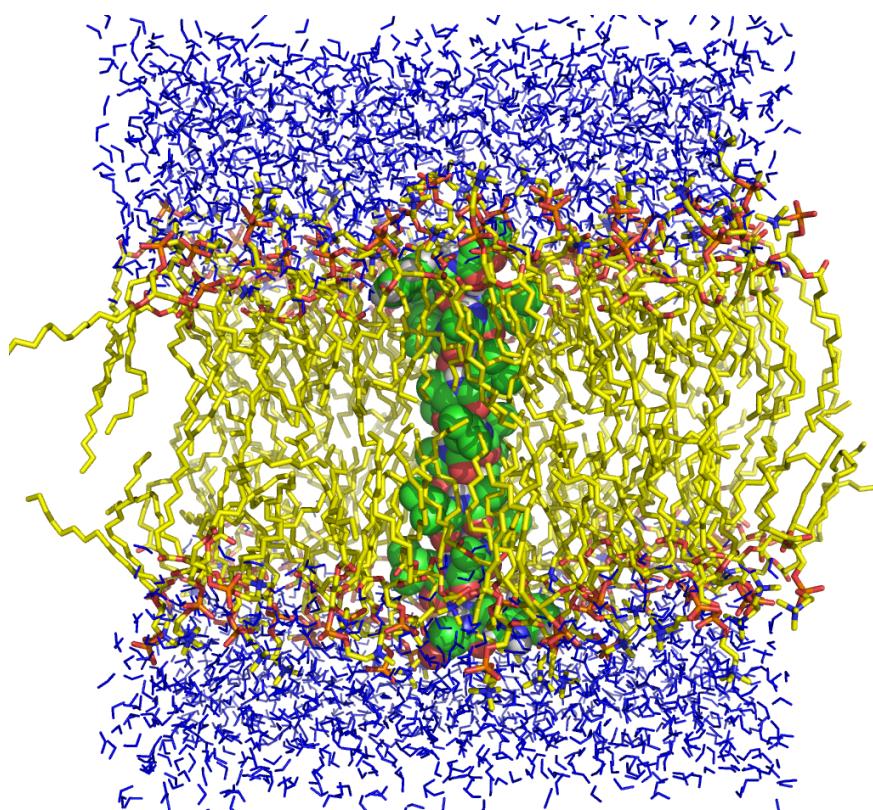
Parametrization: bonded interactions (2)



Parametrization: representative distributions from the PDB

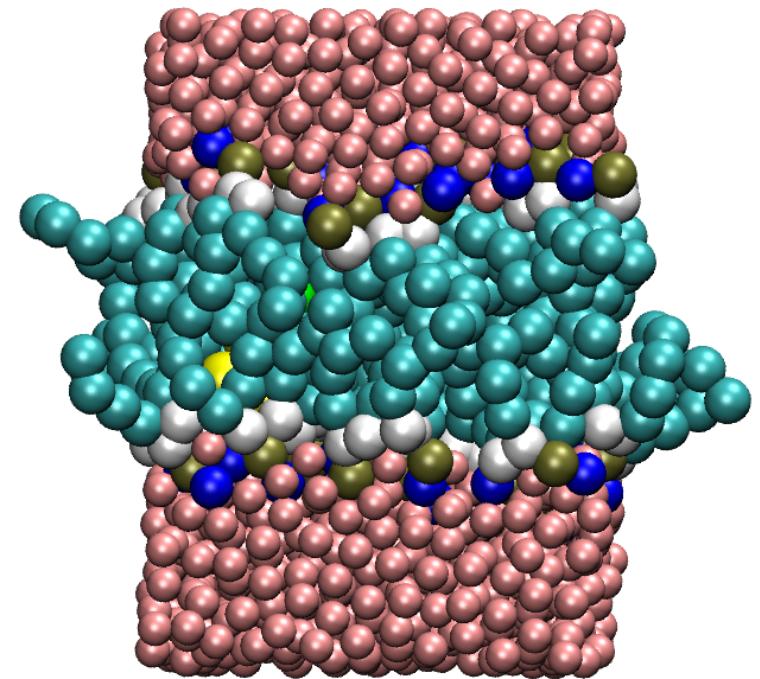


Example



Atomistic (WALP23 in DPMC)
14,652 atoms

*~ 200 CPU days for 1 μ s of
simulation*

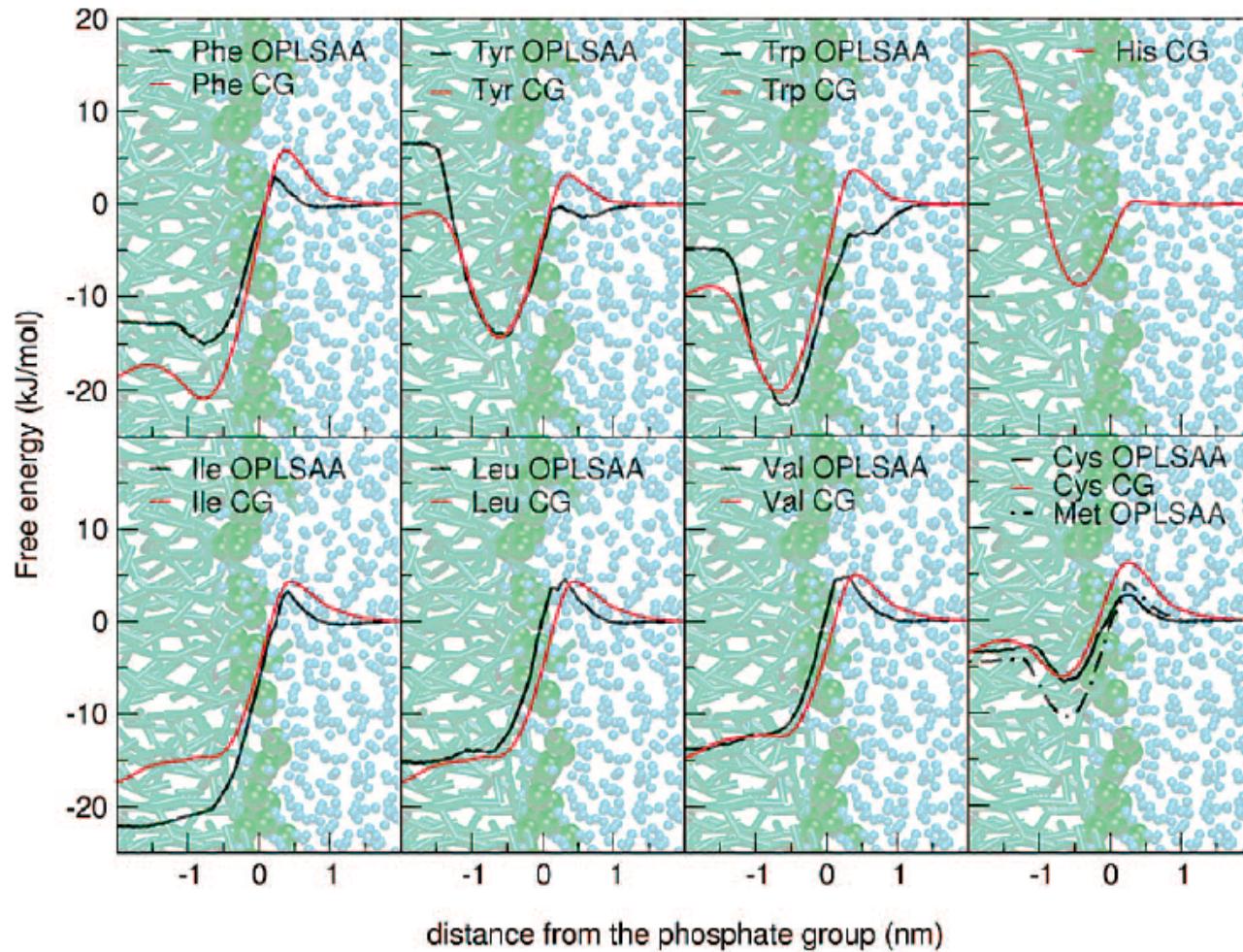


Coarse-Grained (WALP23 in DLPC)
1,986 particles

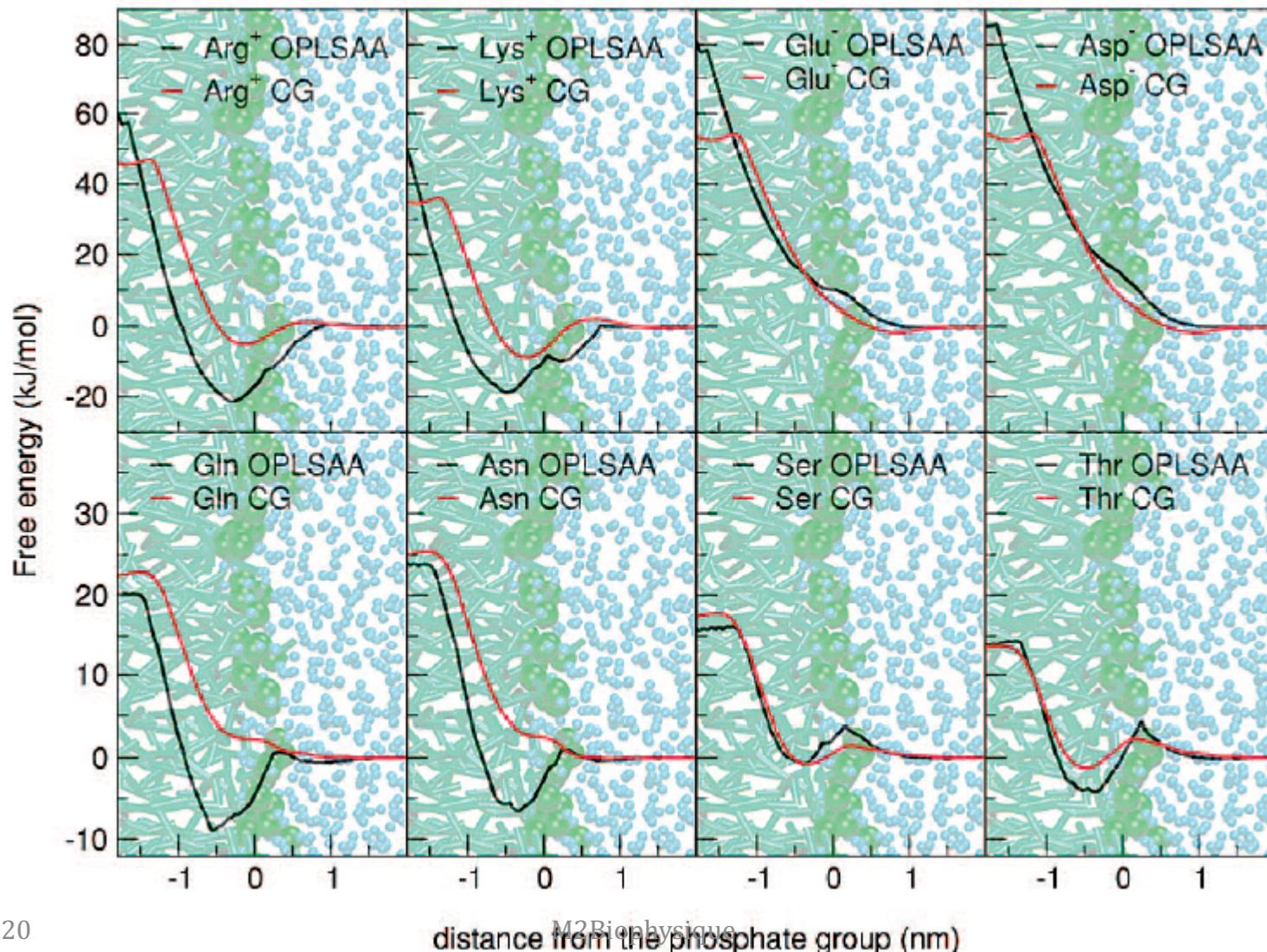
*1 μ s simulated in one night on a
dual CPU computer !!!*

Results

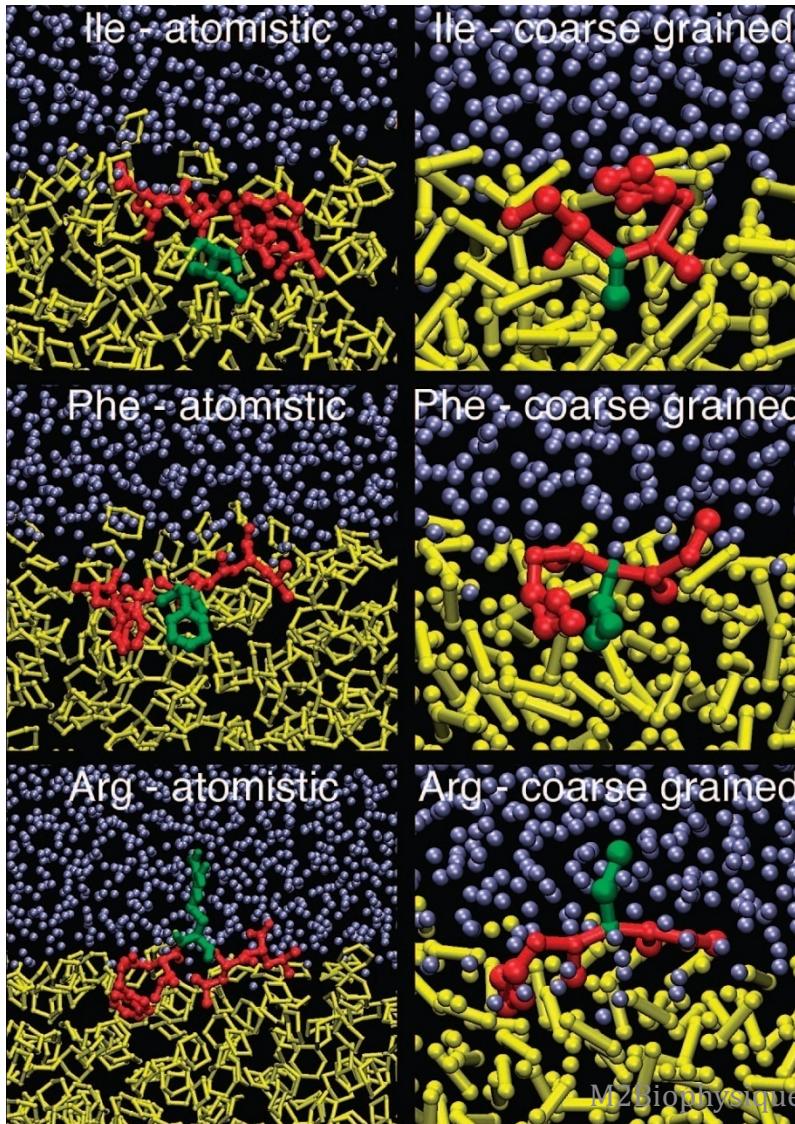
Partitioning of amino-acids in a bilayer



Partitioning of amino-acids in a bilayer (2)



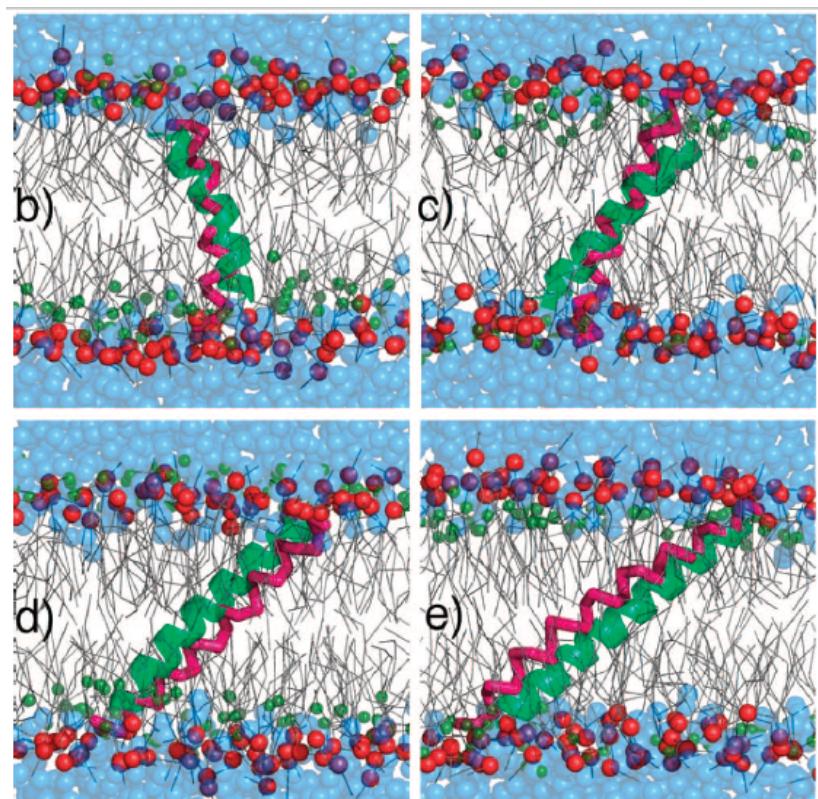
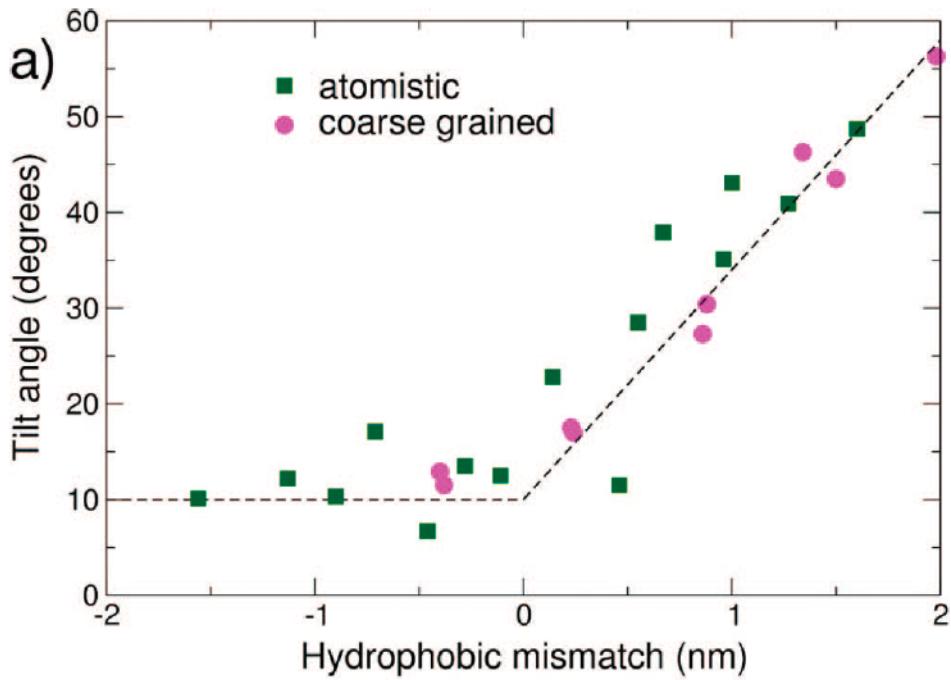
Partitioning of Ace-WLXLL at water/cyclohexane interface



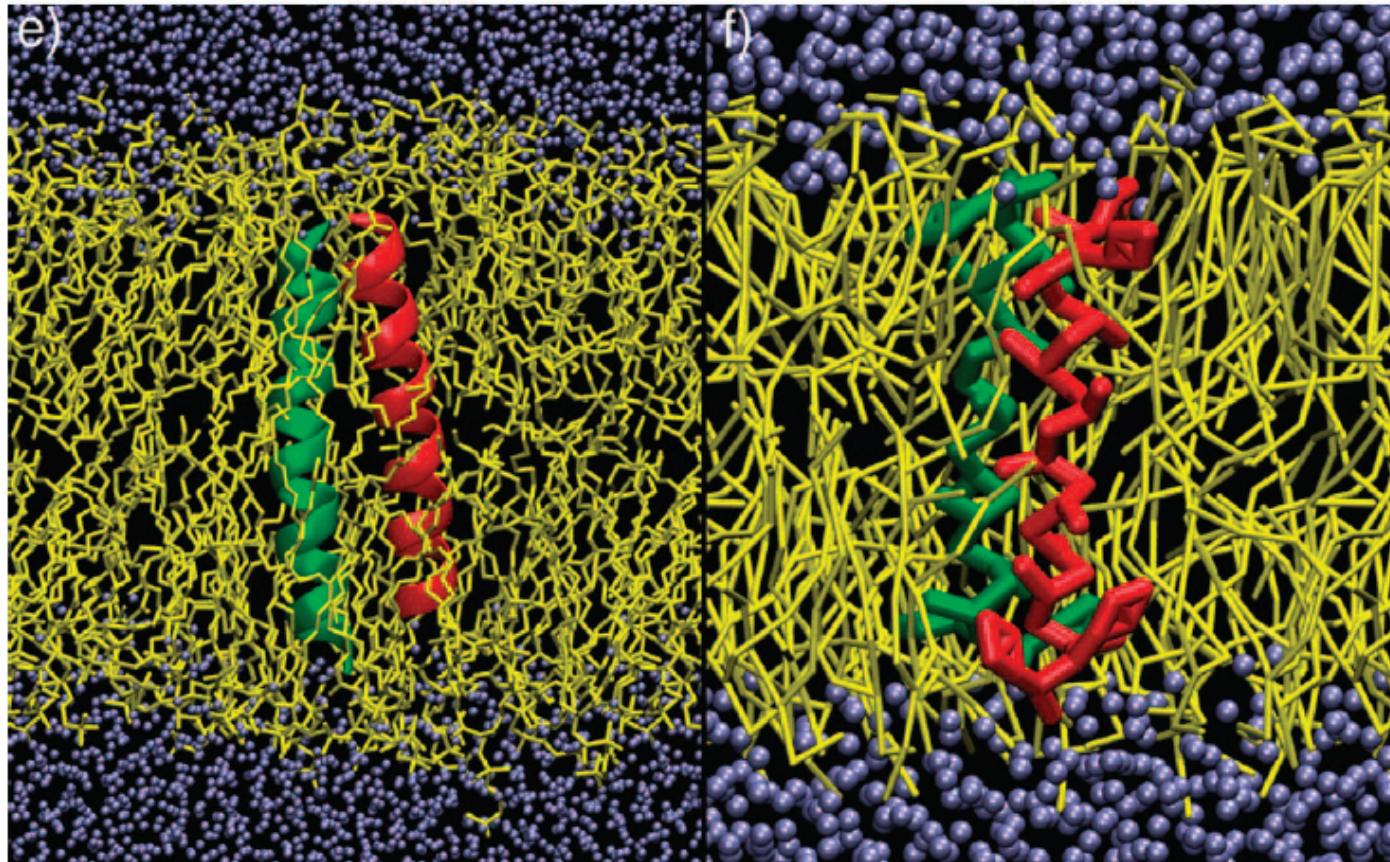
Good Comparison with:

- experiments
- atomistic simulations

Orientation of transmembrane peptides (KALP)



Helix-Helix Association



WALP dimer

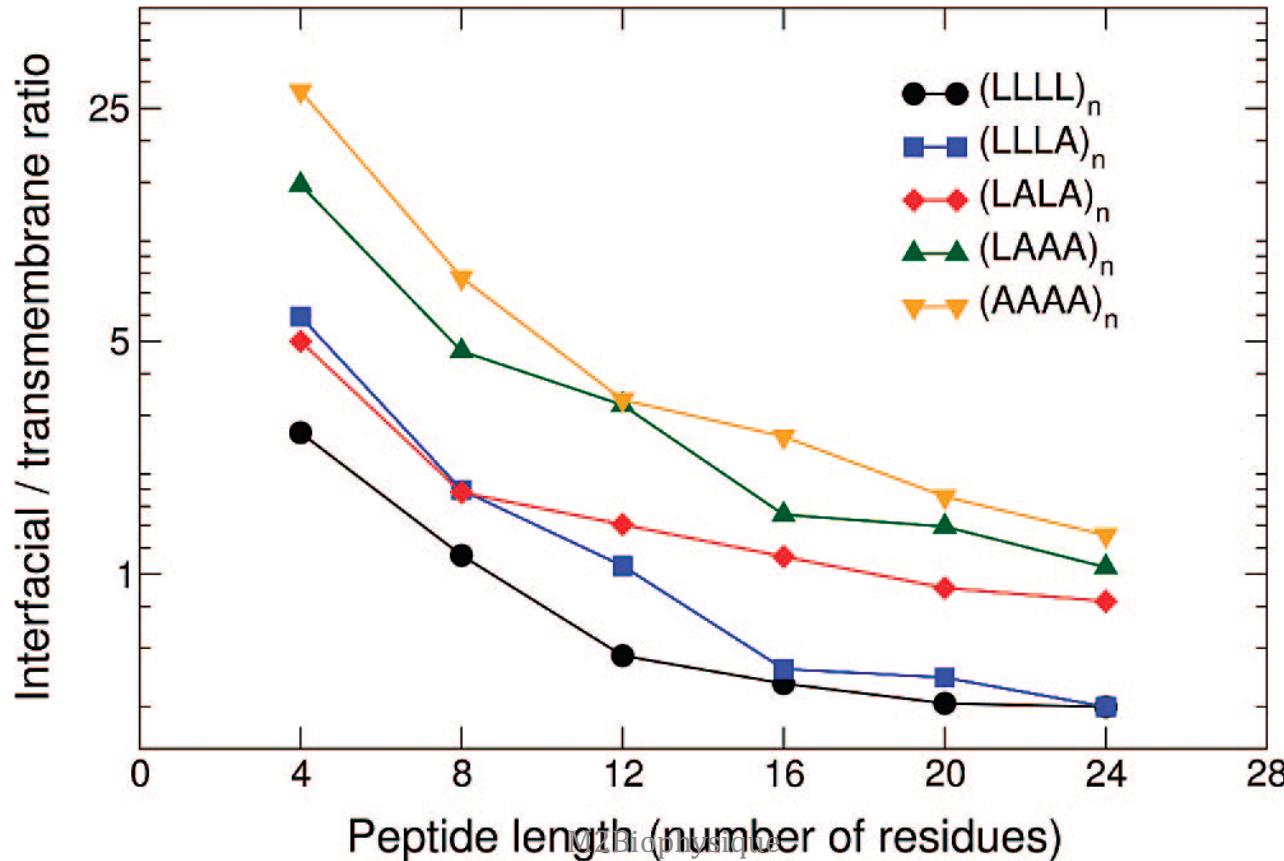
Helix-Helix Association (2)

Table 7. Interhelical Distance and Tilt Angles in the Simulations of the WALP23 Peptide Dimer in DOPC, in Both Atomistic and CG Simulations

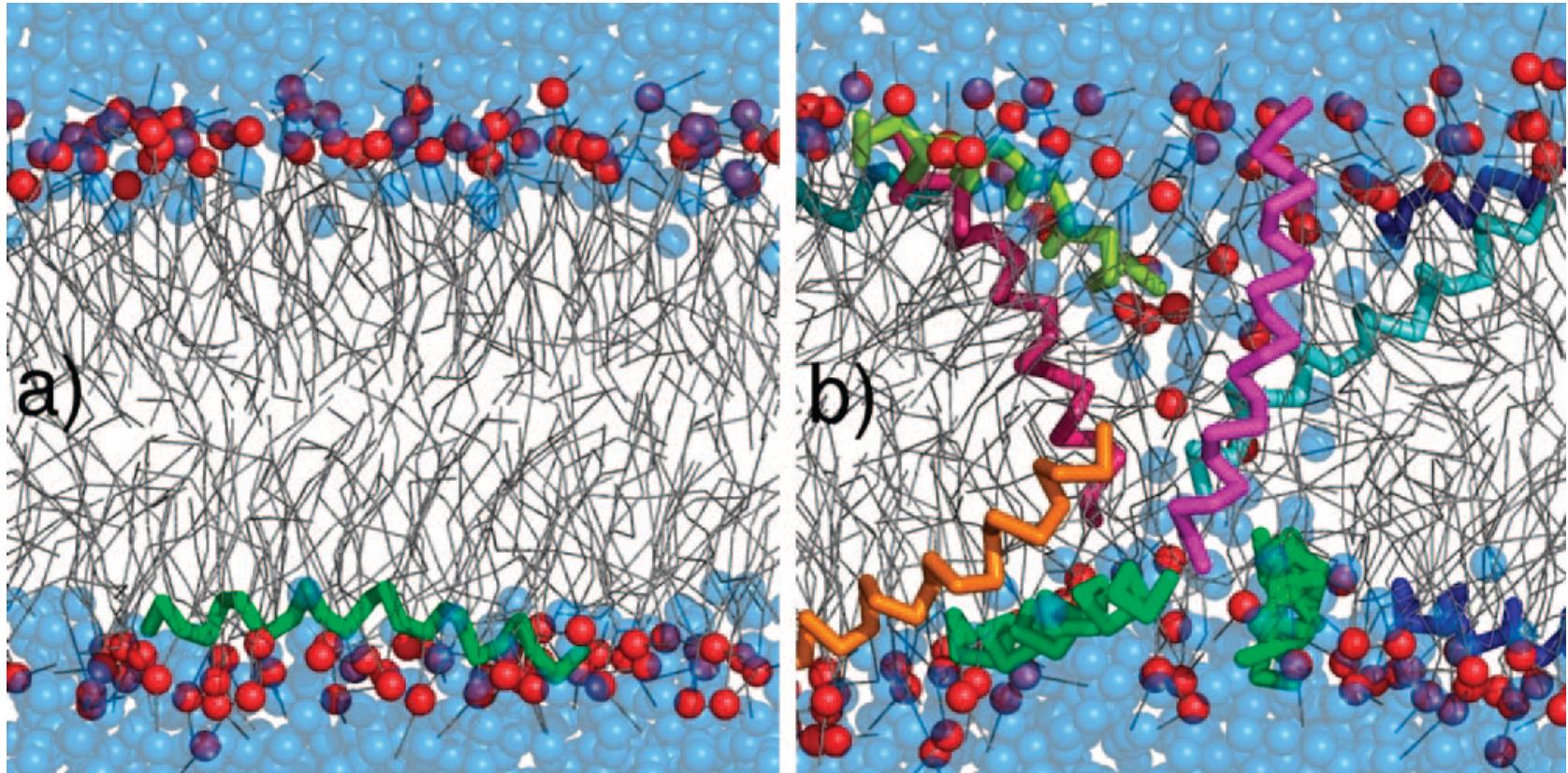
simulation	helix–helix distance ^a	tilt angle ^b (helix1)	tilt angle ^b (helix2)
GMX (1)	$0.84 \pm 0.03^\circ$	$13.5 \pm 4.5^\circ$	$25.2 \pm 4.5^\circ$
GMX (2)	$0.83 \pm 0.04^\circ$	$9.0 \pm 3.5^\circ$	$24.0 \pm 3.5^\circ$
OPLS-AA (1)	$0.86 \pm 0.02^\circ$	$5.8 \pm 3.5^\circ$	$16.1 \pm 4.1^\circ$
OPLS-AA (2)	$0.78 \pm 0.02^\circ$	$18.9 \pm 3.1^\circ$	$32.3 \pm 3.0^\circ$
CG (1)	$0.71 \pm 0.06^\circ$	$13.6 \pm 6.3^\circ$	$16.0 \pm 6.4^\circ$
CG (2)	$0.70 \pm 0.05^\circ$	$13.7 \pm 5.8^\circ$	$15.4 \pm 5.6^\circ$

Partitioning of hydrophobic peptides in a bilayer

- 30 different peptides → 200 (self-assembly) simulations per peptide → total of 6000 simulations: **1.2 ms !**



Transmembrane pores: magainin antimicrobial peptide can form toroidal pore



Advantages / Limitations

- Excellent agreement with experimental / atomistic data
- Resolution → amino-acid
- Works for the fluid phase
- Modified enthalpy / entropy balance
- Partitioning of polar groups in a low dielectric medium

Some applications already published

- Rhodopsin
- Potassium channel
- MscL (mechanosensitive channel)
- Etc...

Extension to Carbon Nanoparticles

- See L. Monticelli Web site

Voth's Model: The multiscale graining method.

- Calibration on AA simulations.
- Definition of a consistent schema for obtaining the parameters: « A rigorous bridge between atomistic and coarse-grained models »
- AND: MS-CG method employs force information sampled from AA simulations to directly approximate the many-body mean force field:
 - => The force field derived from gradients of the many body PMF = FORCE MAPPING Method

An example

The entire HIV-1 virion within a hybrid multiscale simulation approach.

HIV-1 begins assembly with multimerization of the Gag polyprotein at the plasma membrane of infected cells => System= Gag polyprotein+Membrane

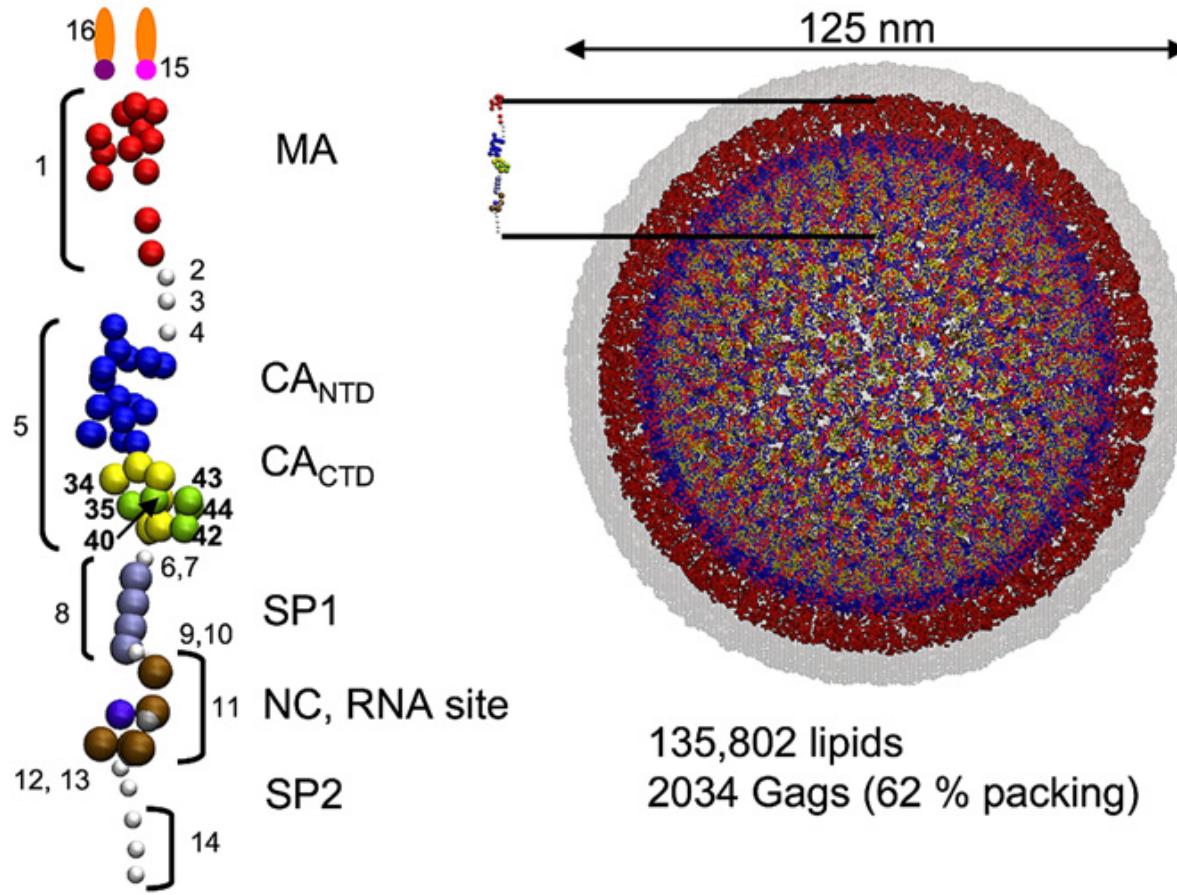
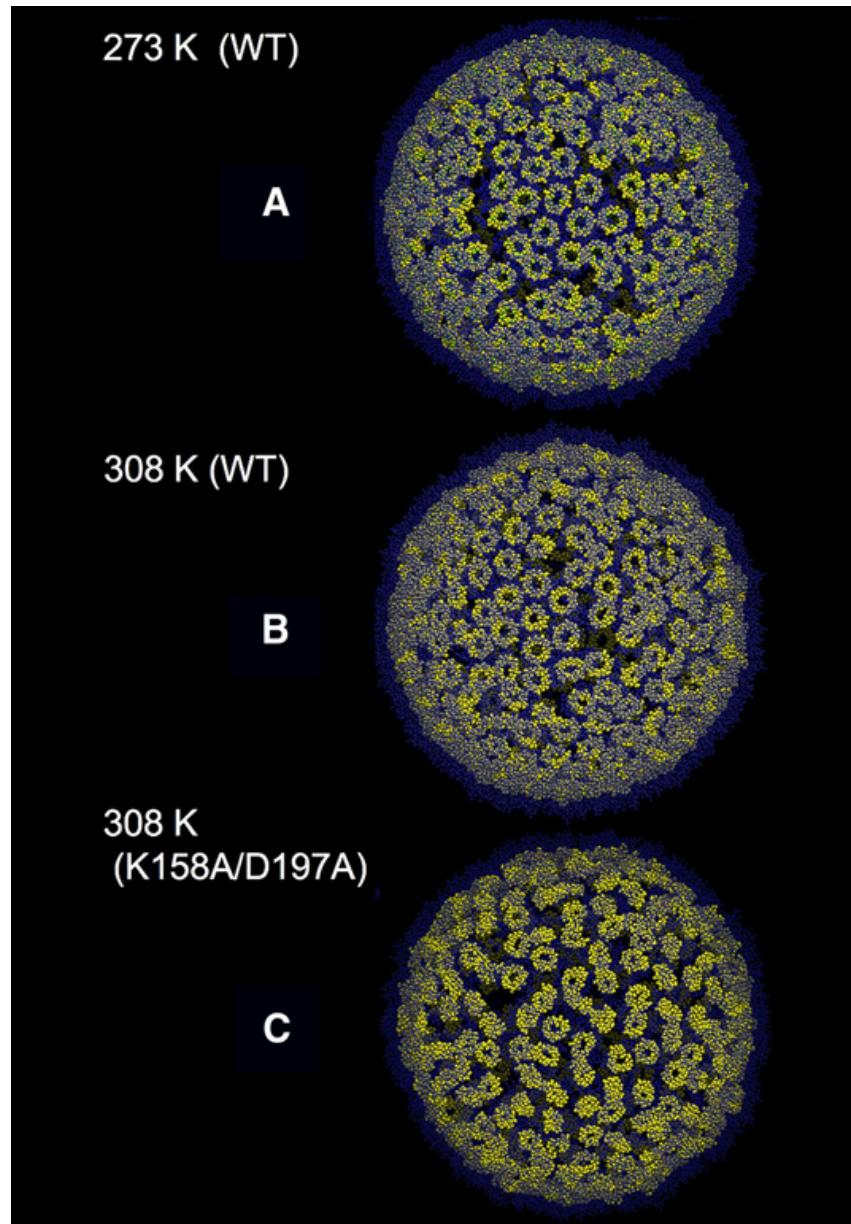


FIGURE 1 A schematic of the Gag/membrane CG model. (*Left panel*) Domain decomposition of the Gag and PS/PC membrane. (*Small spheres*) Linker sites. Detailed atomic-level information was used to parameterize the MA, CA, and NC/RNA domains. Close-contact CG sites 34, 35, 40, 42, 43, and 44 in the CA_{CTD} domain are highlighted. (*Right panel*) Initial virion structure; 2034 Gags, each in bundles of six, were replicated in a 125-nm diameter virion. The initial p6 hexagonal lattice spacing was 10 nm, with the Gags relatively uniformly spaced.



CG simulation snapshots of the CA domain p6 hexagonal ordre

12/14/20

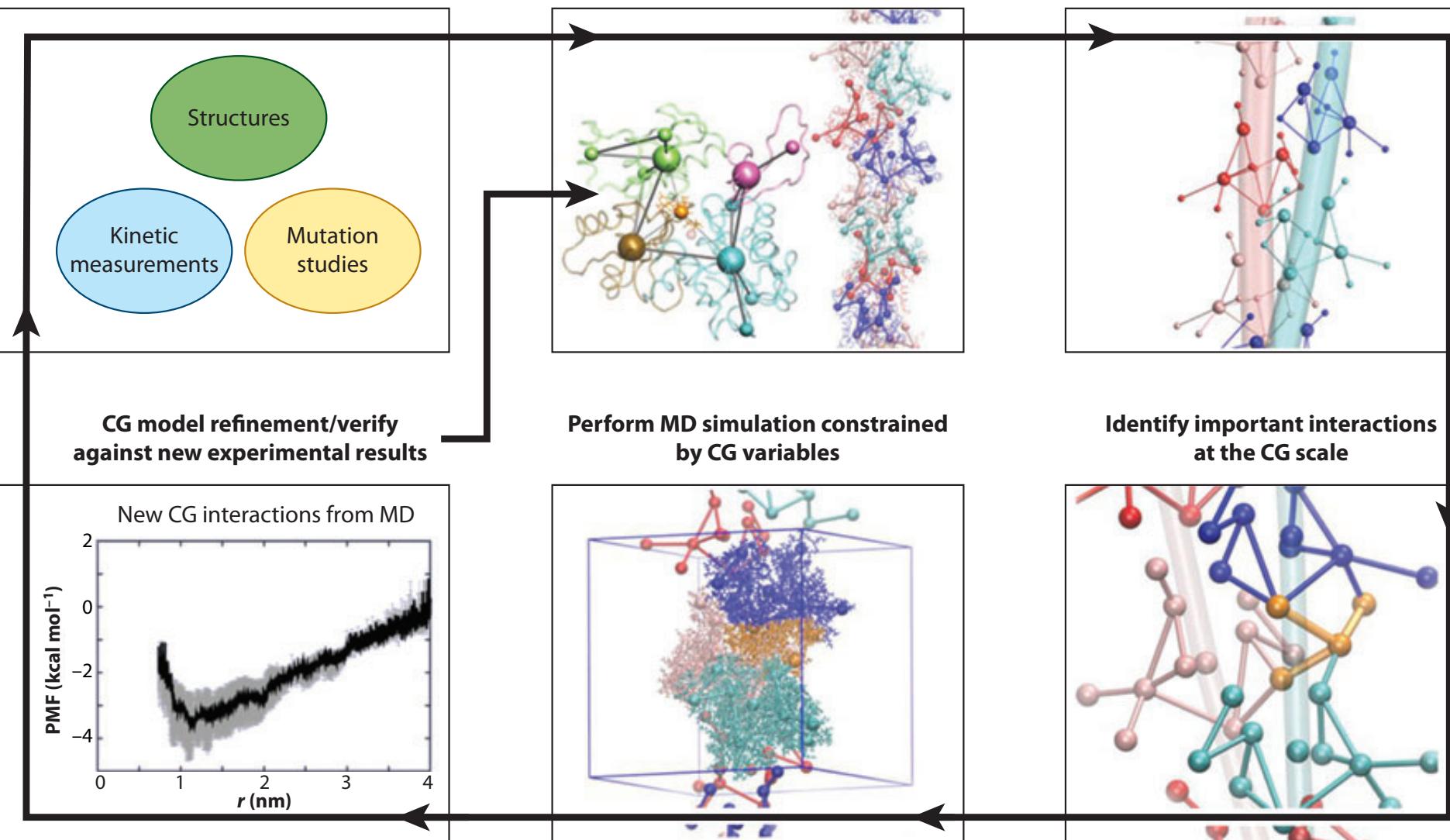
M2Biophysique

The interactions in the residues corresponding to the close-contact CG sites in the CACTD domain are largely responsible for forming and maintaining the p6 hexameric lattice structure in the immature HIV-1 virion

Identify key experimental results

Highly CG model development

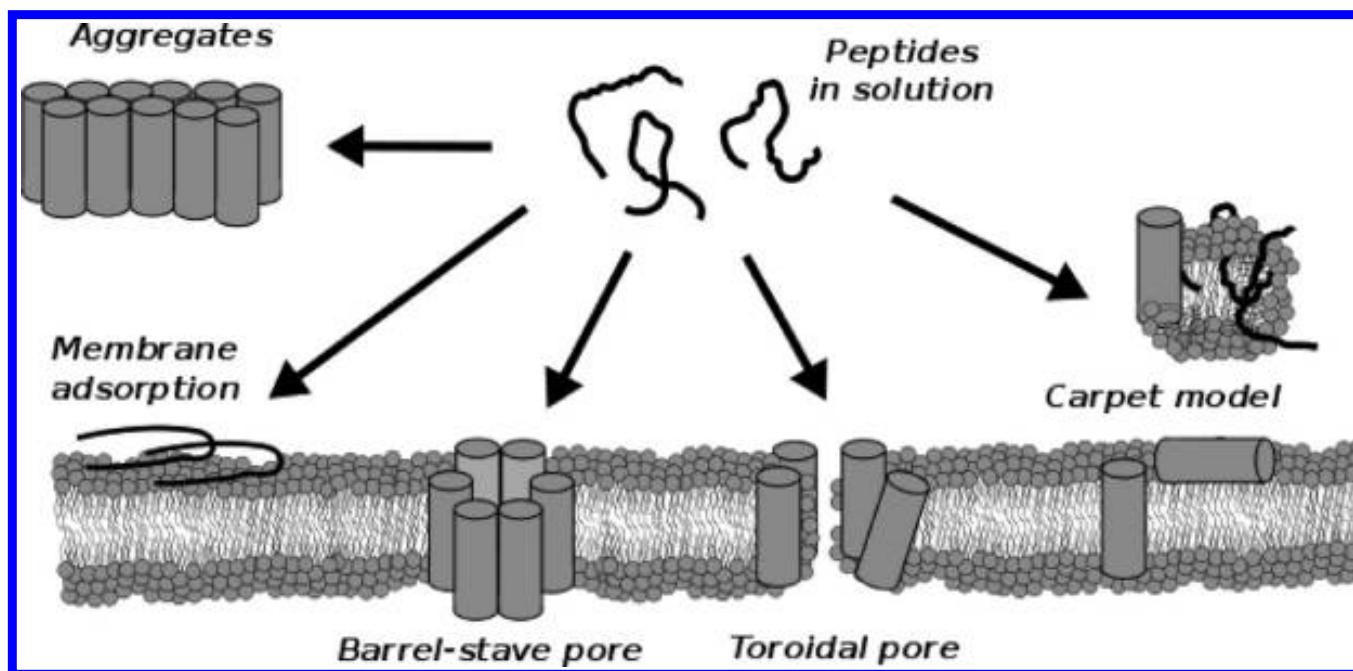
Carry out numerous large-scale CG simulations



A paradigm for multiscale problems that provides a conceptual framework for organizing coarse-graining methodologies as well as expanding their use. CG, coarse-grained; MD, molecular dynamics.

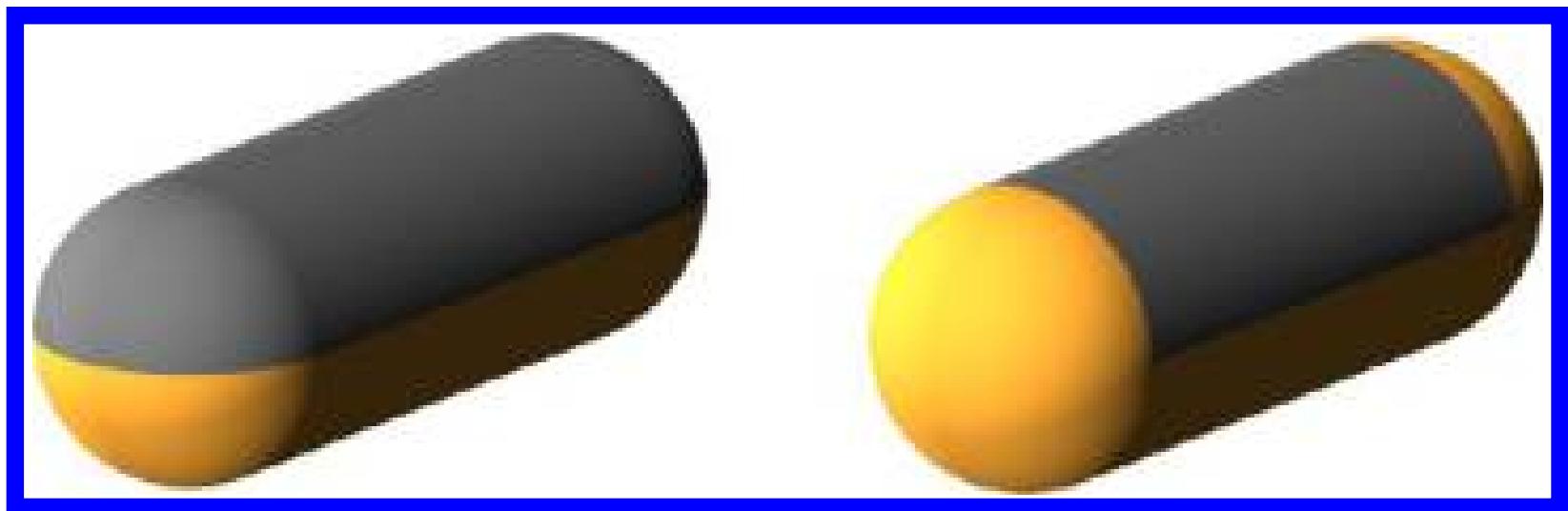
Other Models

- Formation of Pores:



A coarse-coarse Model

prolate, amphiphilic peptides interacting with a membrane.



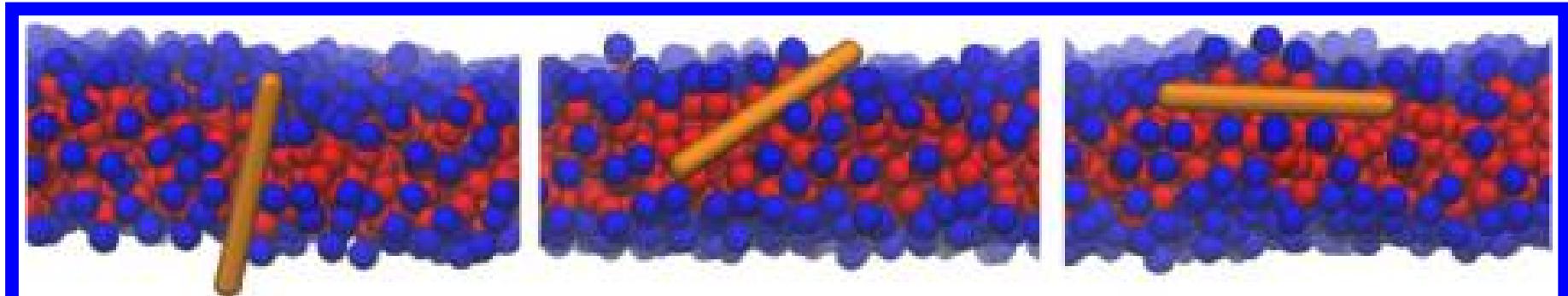
Snapshots of the employed coarse-grained model of amphiphilic peptides. Spherocylinders, shown in orange, have attractive stripe on its side (colored gray). On the left side is the model has attractive ends (AE), while on the right side is the model with no-attractive ends (NE).

Model

- Phospholipid: 3 beads
 - The bead representing the hydrophilic headgroup is purely repulsive. The other two beads represent the hydrophobic tail and have a relatively long-ranged attraction.
- Implicit Solvent Model

Results

Representative snapshots of the model amphiphilic peptide at the membrane edge (in the pore)

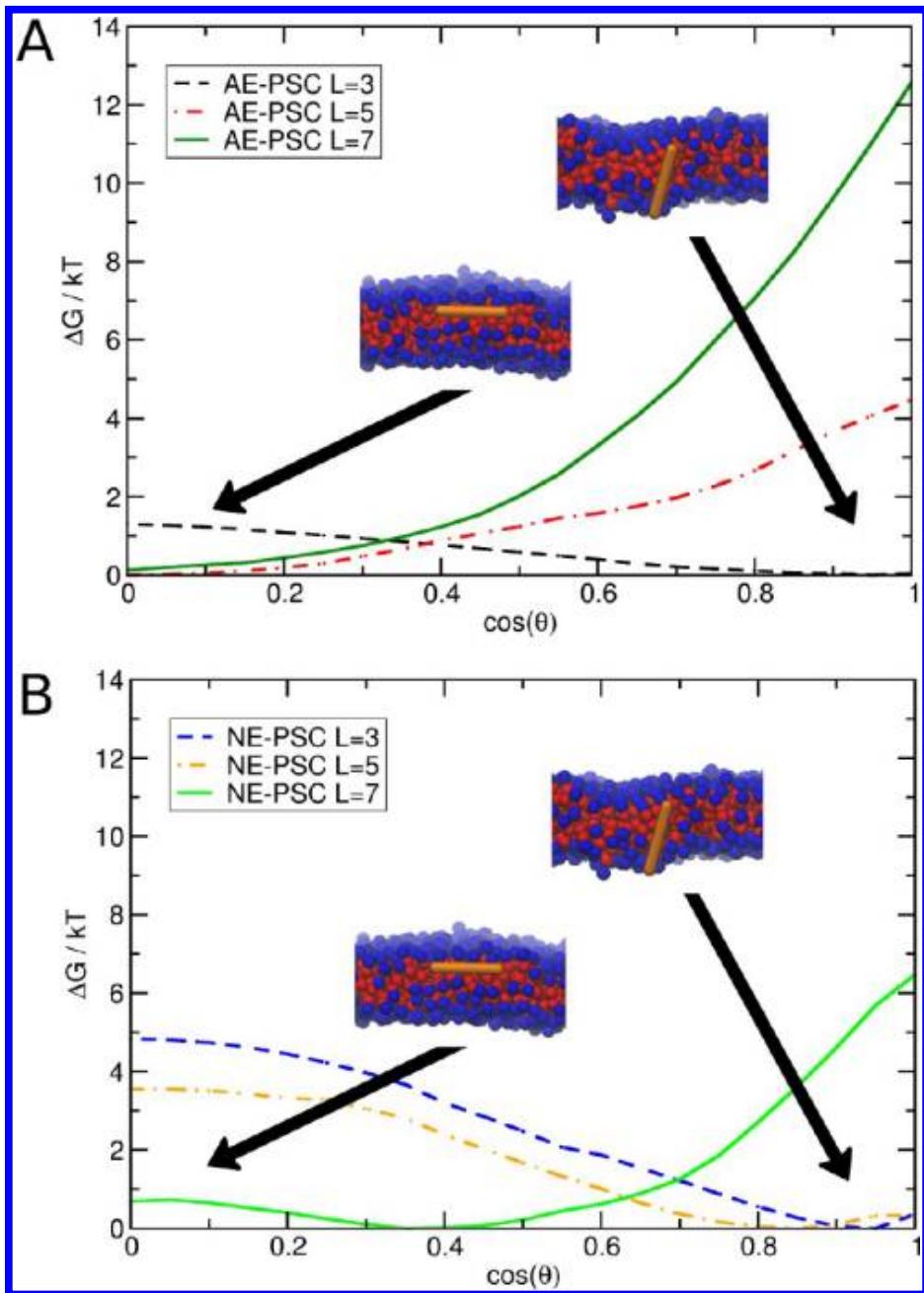


Results

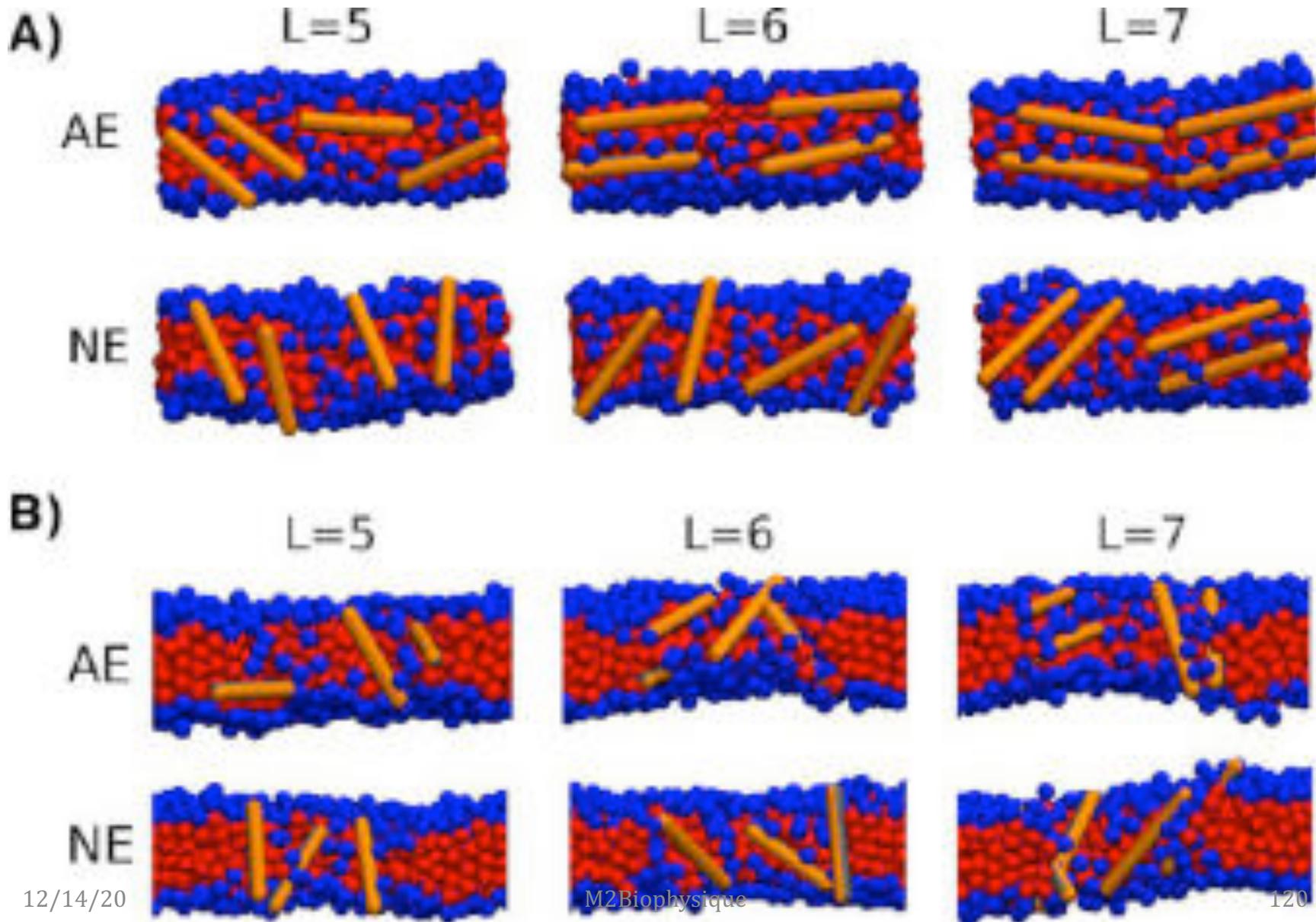
Free energy profiles of peptide orientation with both attractive ends AE-PSC (A) and no-attractive ends NE-PSC (B) models and with the peptide length from 3 to 7 nm

L=5nm: orientation depends on attractiveness of end-caps.
// for Attractive ends and Perp. for non attractive ends

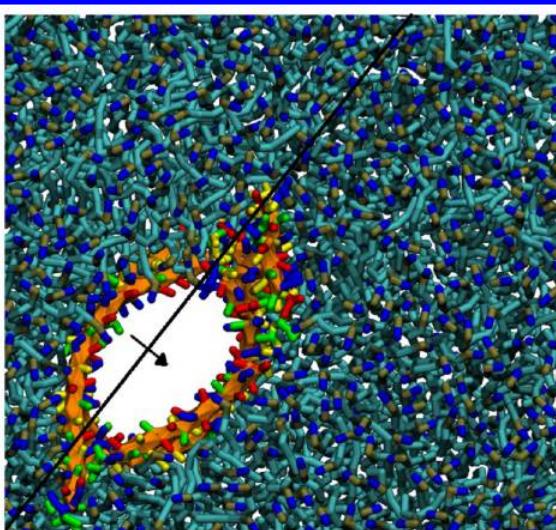
12/14/20



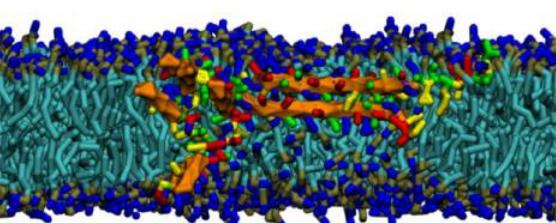
More peptides



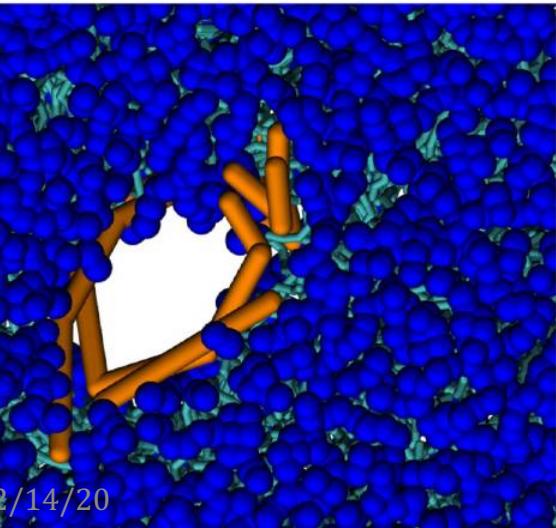
A)



B)



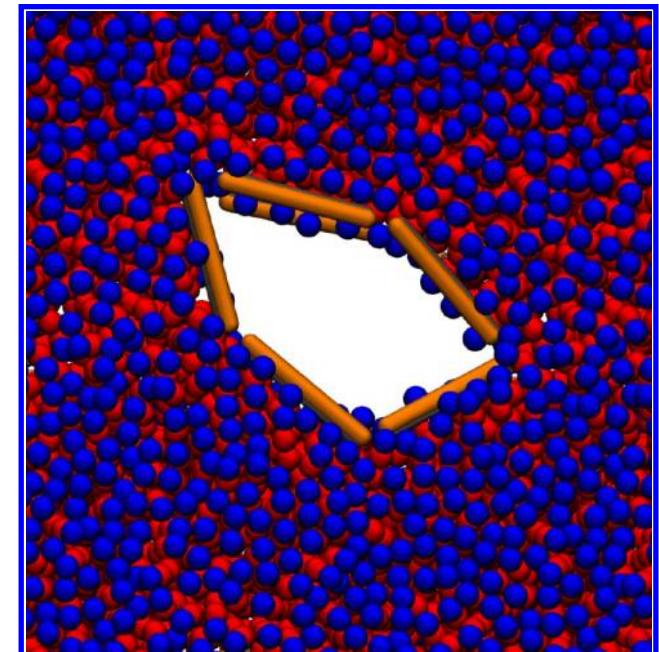
C)



A protein? Apo-A1 with Martini

No spontaneous breakup or spontaneous formation of double-belt pores.

With a ApoA1-like protein



Spontaneous double-belt pore

Conclusions

Table 1

Publicly available tools and packages for coarse-grained molecular dynamics.

Name	Major targets	Briefly	Co-authors	Ref
<i>CG-specific software</i>				
CafeMol	Proteins, NA	Go, AICG, ratchet	Takada	[13*]
ESPResSo	Lipids and DNA	Generic CG MD	Holm	[73]
LAMMPS	Generic	Generic	Plimpton	[74]
RedMD	Proteins, NA	ENM + α , specific	Trylska	[75]
REACH	Proteins	ENM + α	Smith, Moritsugu	[20*]
Yup	Nucleic acids	Generic	Harvey	[76]
<i>Force-fields/scripts for MD packages</i>				
Go-model-builder(mmtsb)	Proteins	Go	Brooks	[77]
MARTINI	Lipids, proteins	PChem + α , w/GROMACS	Marrink	[23]
NAMD-CG builder	Proteins, lipids	w/GROMACS	Schulten	[78]
SMOG	Proteins, RNA	Go, w GROMACS	Onuchic	[32*]
<i>Web service</i>				
anm/oGNM/iGNM	Proteins	ENM	Bahar, Yang	[16]
<i>CG to AA reconstruction</i>				
BBQ	Protein backbones	Reconstruct backbone from C α	Kolinski	[79]
SCWRL	Protein sidechains	Reconstruct side-chain	Dunbrack	[80]

S. Takada , Current Opinion in Structural Biology 2012, 22:130–137