Modelos Coarse-Grained

Modelagem em Escala Multipla
VII EMMSB, LNCC, 19.08.2014

Hubert Stassen

gullit@iq.ufrgs.br

Grupo de Química Teórica

Instituto de Química - UFRGS

The Nobel Prize in Chemistry 2013

Photo: A. Mahmoud

Martin Karplus

Prize share: 1/3

Photo: A. Mahmoud

Michael Levitt

Prize share: 1/3

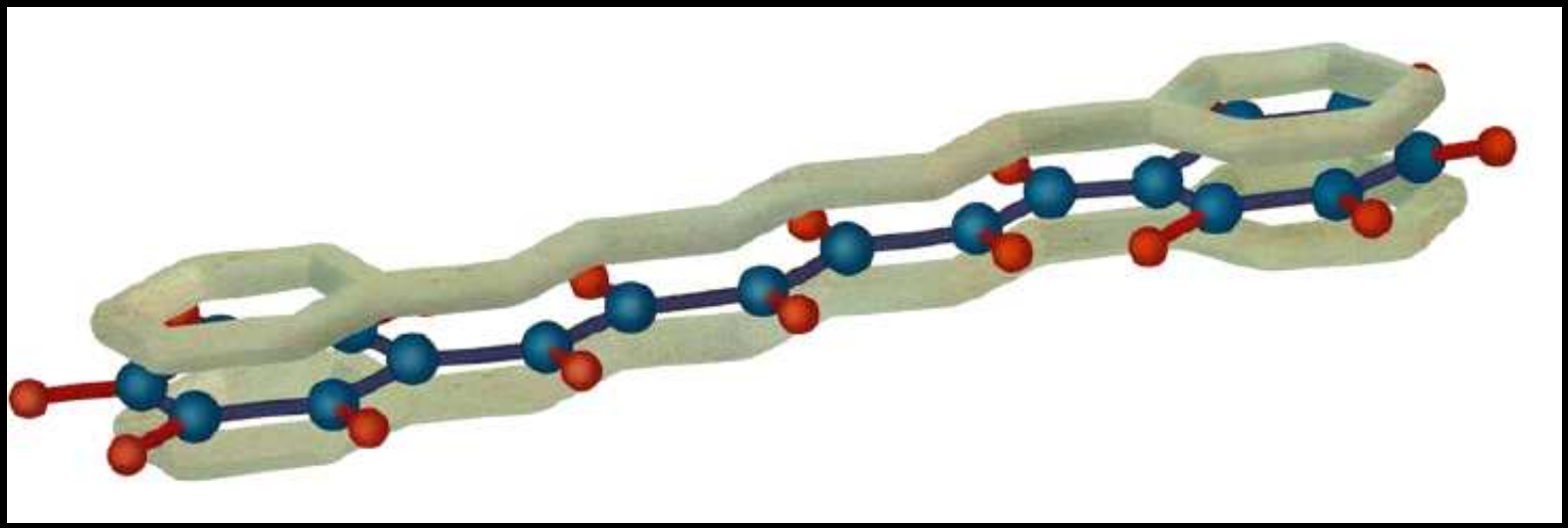
Photo: A. Mahmoud

Arieh Warshel

Prize share: 1/3

The Nobel Prize in Chemistry 2013 was awarded jointly to Martin Karplus, Michael Levitt and Arieh Warshel *"for the development of multiscale models for complex chemical systems"*.

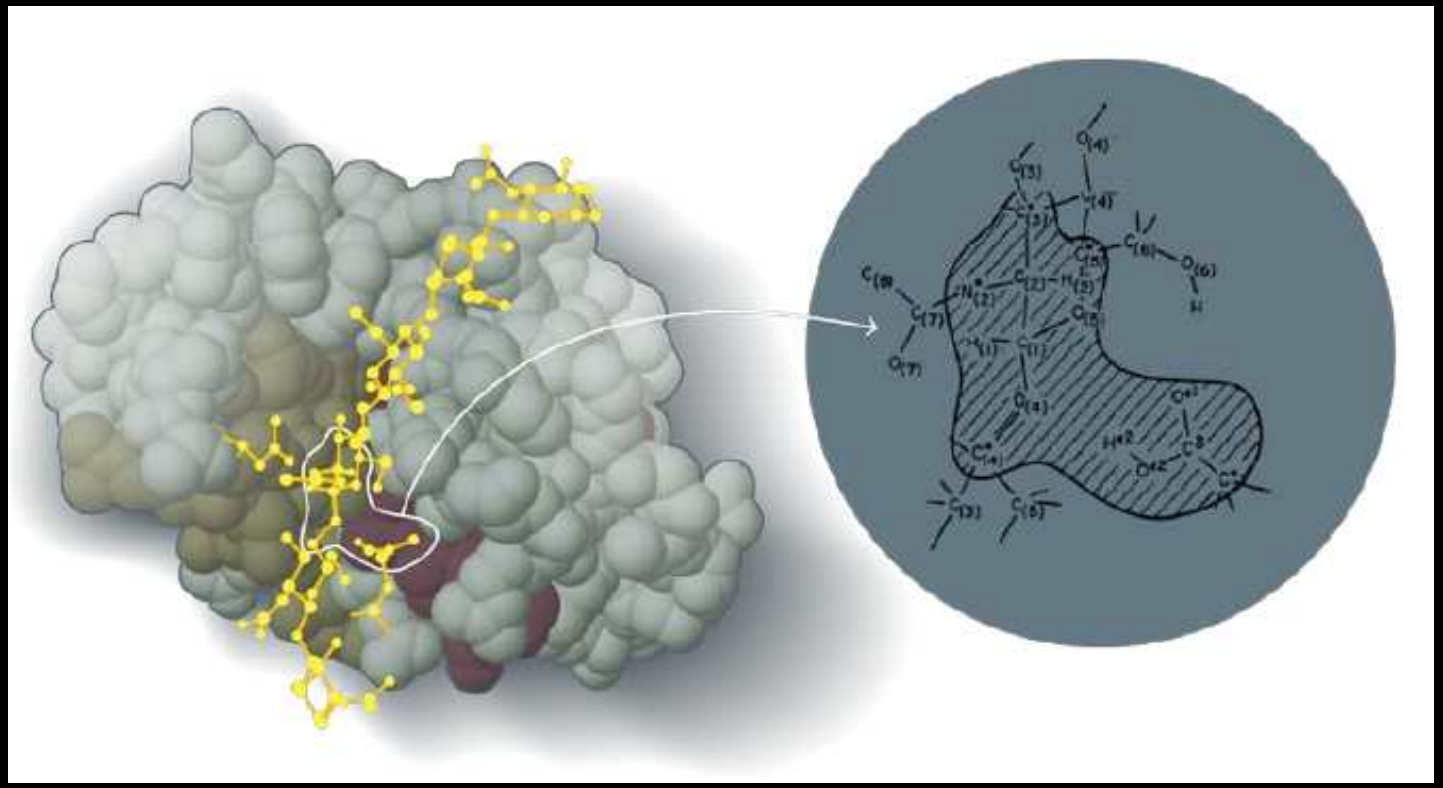
Retinal Molecule



Separating σ skeleton (MM) and π system (QM)

A. Warshel, M. Karplus, JACS 94 (1972), 5612

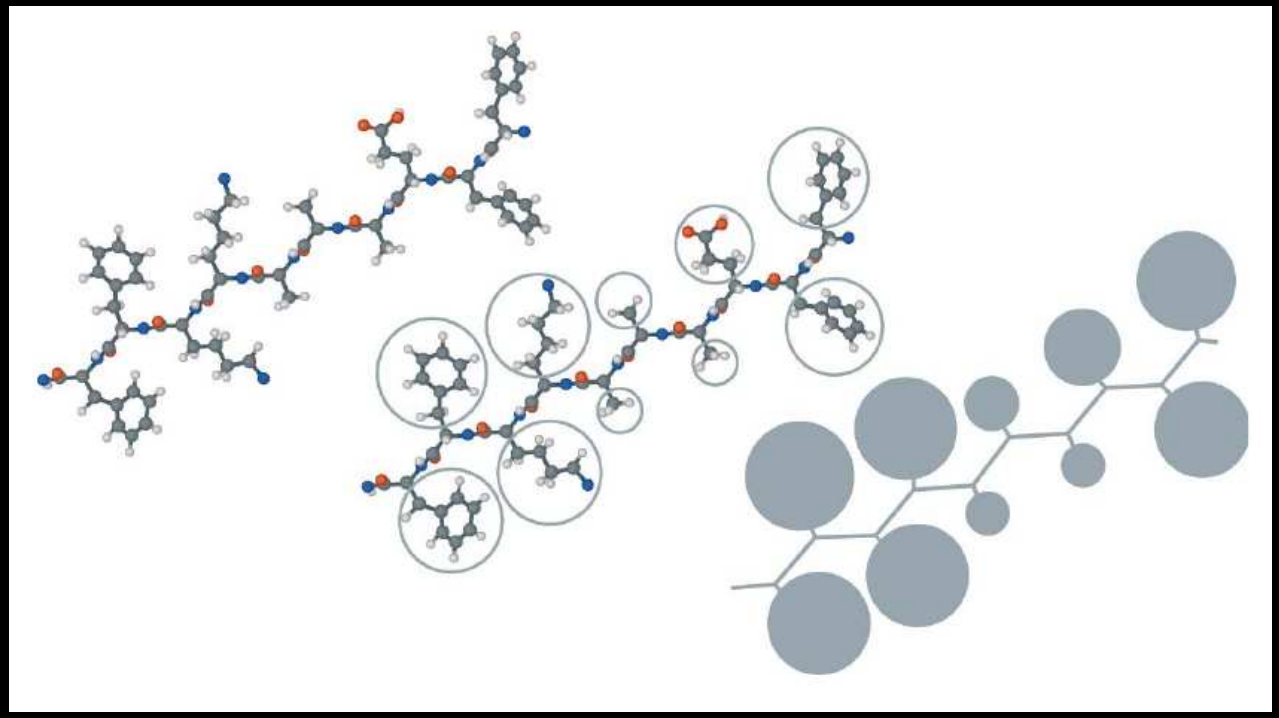
Lysozyme Reaction



QM/MM + continuum treatment

A. Warshel, M. Levitt, J.Mol.Biol. 103 (1976), 227

Protein Folding



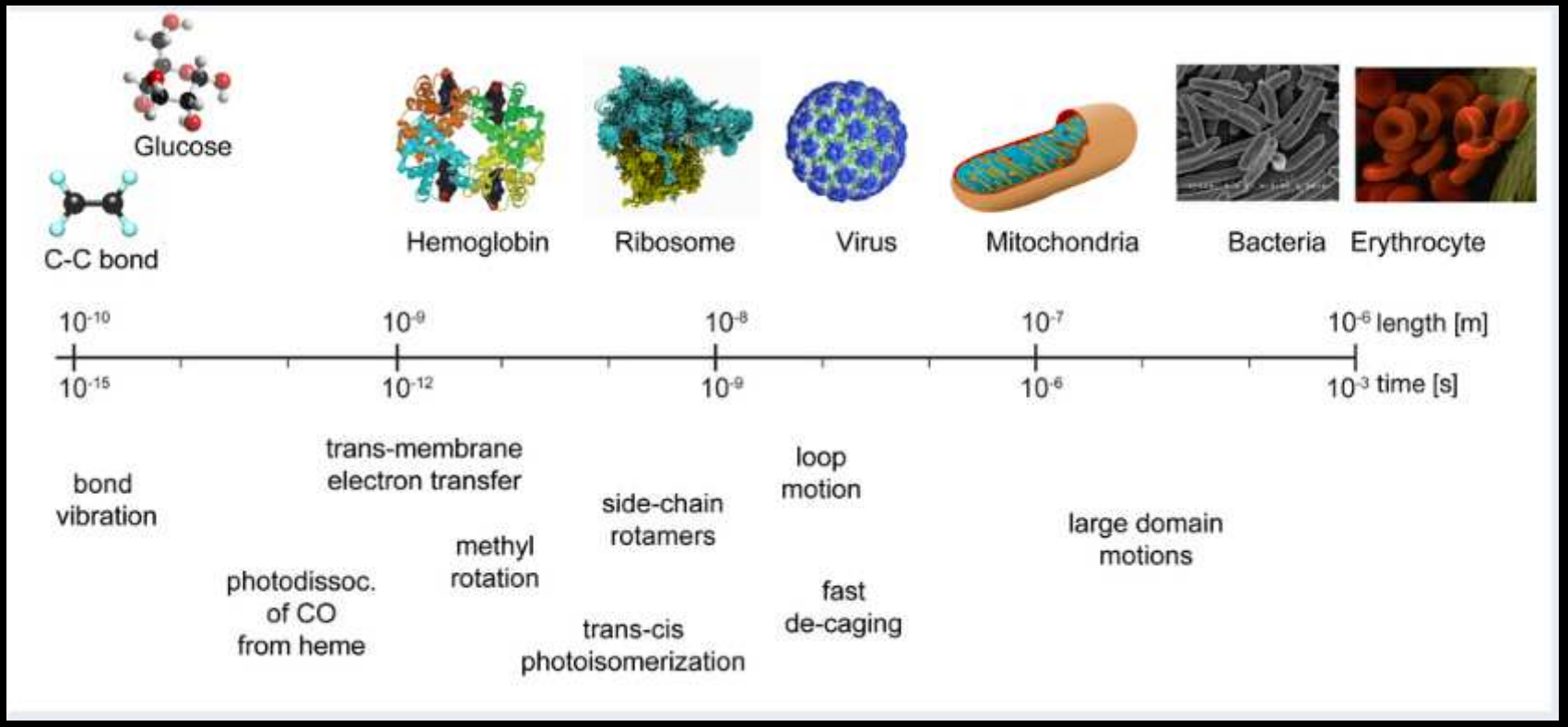
Simulated annealing of a "CG" Bovine Pancreas Trypsin Inhibitor

M. Levitt, A. Warshel, Nature 253 (1975), 694

Outline

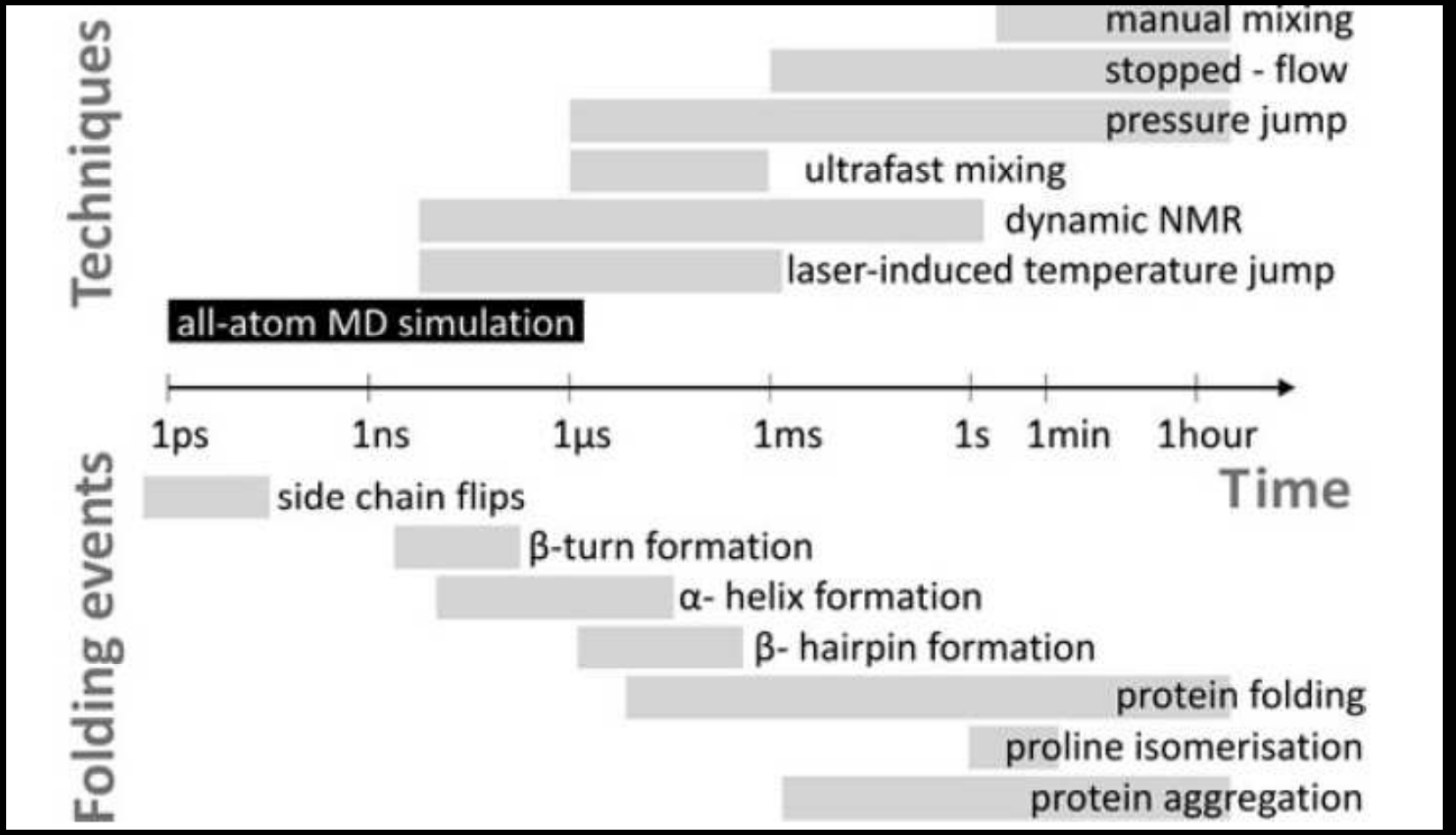
1. Time and length scales
2. Molecular Dynamics
3. Coarse-Graining
4. Protein membrane interaction
5. Jumping between scales

Time and Length Scales



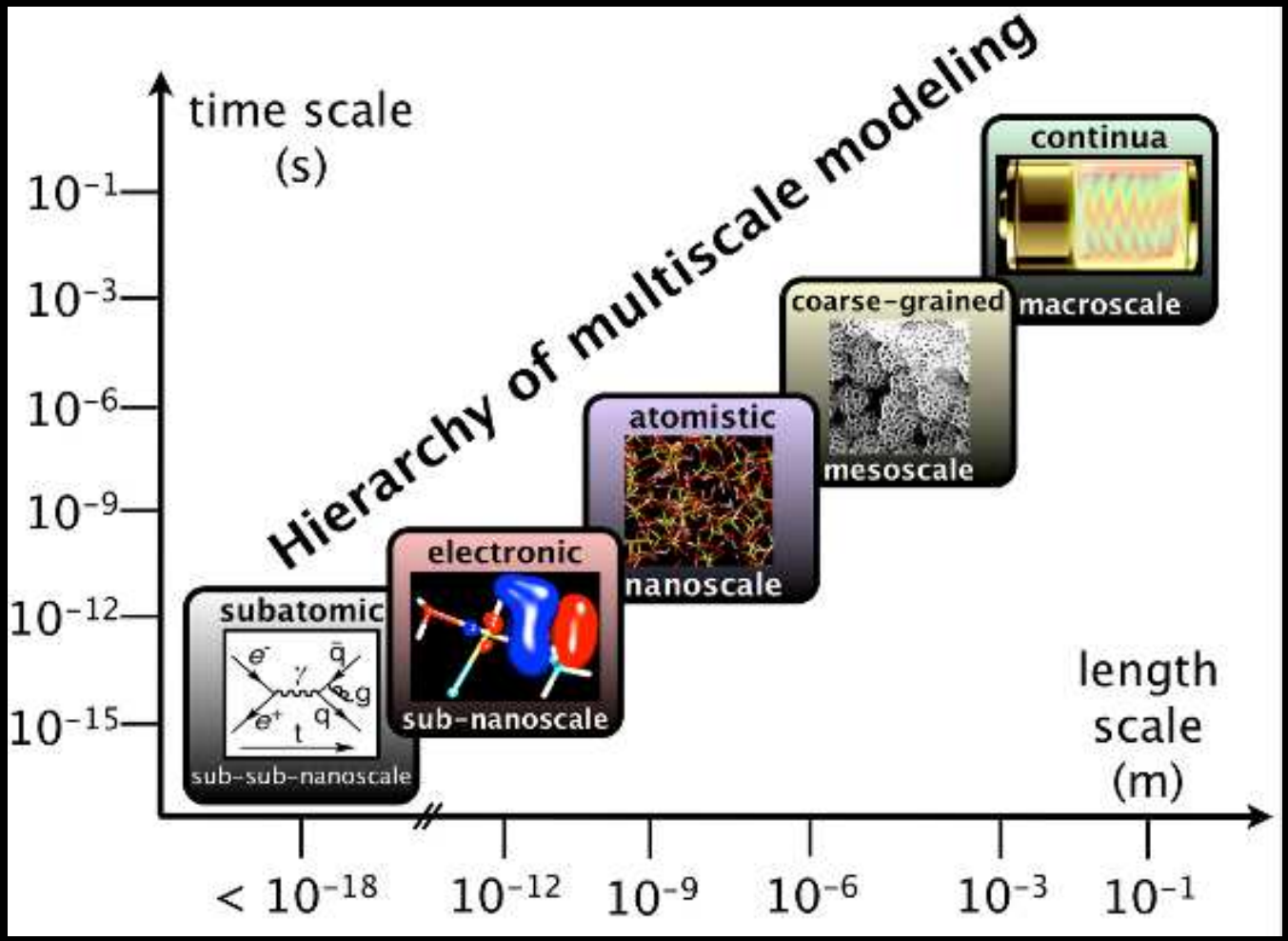
Biological scales

http://www.psi.ch/swissfel/time-and-length-scales-of-biochemical-reactions



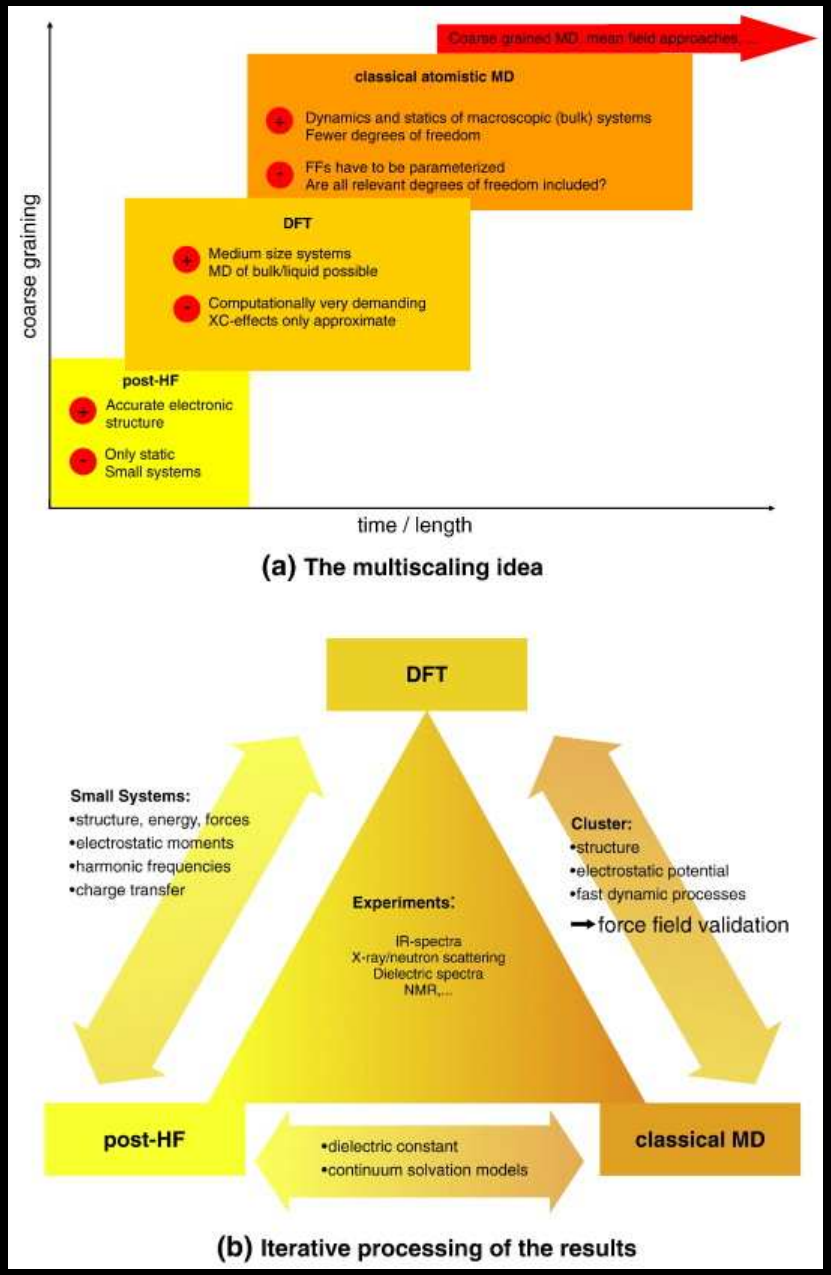
Protein dynamics

S. Kmiecik et al., em A. Kolinski (ed.) *Multiscale Approaches to Protein Modeling* (Springer, 2011), 281



Computational scales

http://www.hiu.kit.edu/104.php



QM - DFT - MD

F. Dommert et al., J. Mol. Liquids 152 (2010), 2

Molecular Dynamics

Numerical solution of equations of motion - classical case

$$\mathbf{r}_i(\Delta t) = \frac{\mathbf{F}_i}{2m_i}\Delta t^2 + \mathbf{v}_i(0)\Delta t + \mathbf{r}_i(0)$$

$$\mathbf{v}_i(\Delta t) = \frac{\mathbf{F}_i}{m_i}\Delta t + \mathbf{v}_i(0)$$

$$\mathbf{F}_i = -\frac{d}{d\mathbf{r}_i}V(\mathbf{r})$$

with $i = 1, 2, \ldots, N$

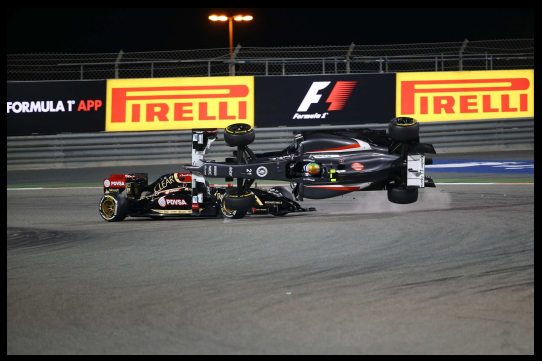
based on a force field (here: AMBER)

$$
\begin{aligned}
V(\mathbf{r}) &= \sum_{i=1}^{bonds} k_{b,i}\left(\mathbf{r}_i - \mathbf{r}_{o,i}\right)^2 + \sum_{i=1}^{angles} k_{\theta,i}\left(\theta_i - \theta_{o,i}\right)^2 \\
&+ \sum_{i=1}^{dihed.} \frac{V_{o,i}}{2}\left[1 + \cos\left(n_i\phi_i - \gamma_i\right)\right] \\
&+ \alpha \sum_{i,j>i}^{N} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \\
&+ \beta \sum_{i,j>i}^{N} \frac{q_i q_j}{4\pi\epsilon_o r_{ij}}
\end{aligned}
$$

$$
\varepsilon_{ij} = = \sqrt{\varepsilon_i \varepsilon_j} \quad \sigma_{ij} = \frac{\sigma_i + \sigma_j}{2}
$$

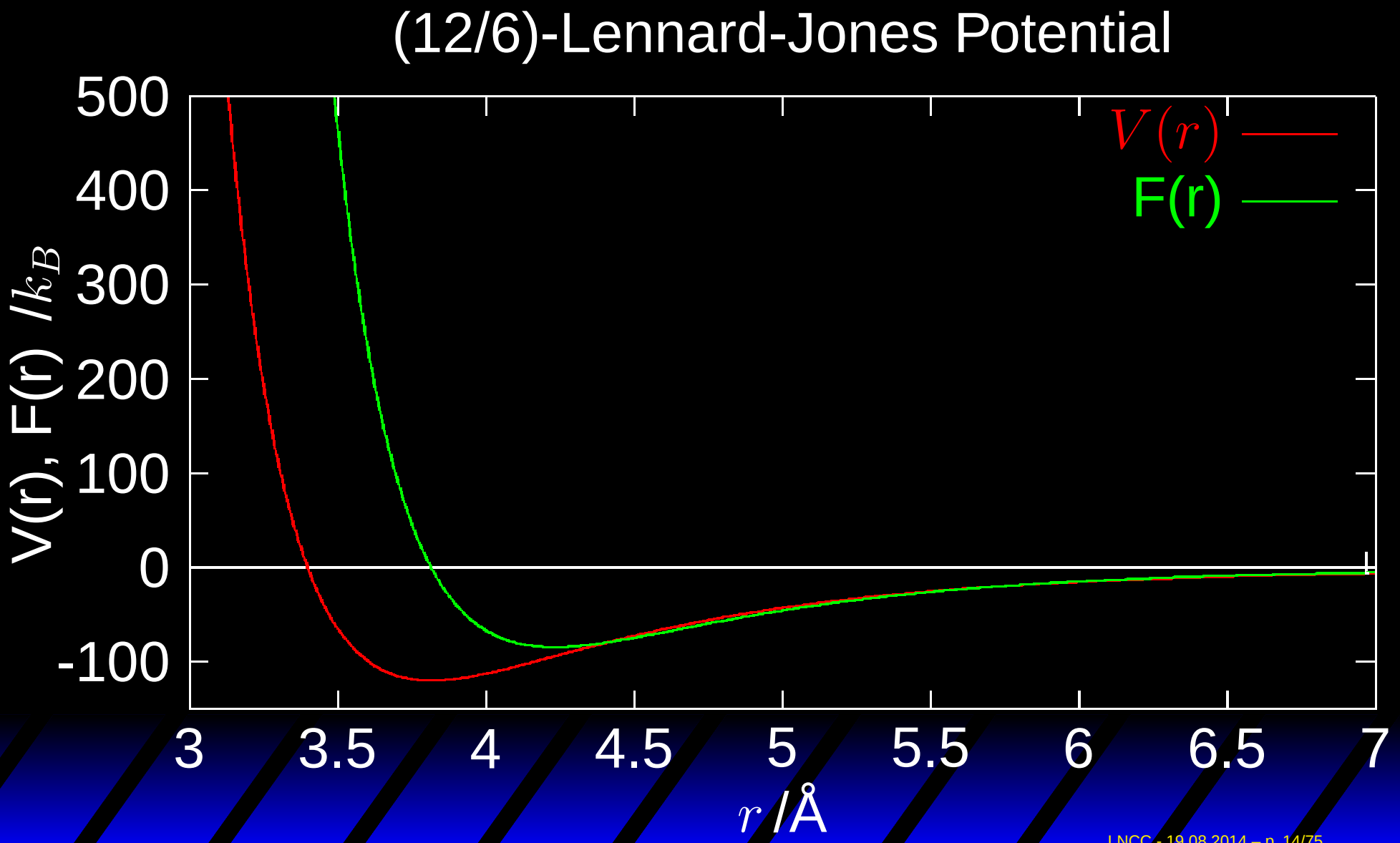
Time step Δt

- as large as possible
- resolve the dynamical process



- flexible atomistic simulations: $\Delta t \simeq 1 \cdot 10^{-15}$ s (fs)

- we need constant forces during Δt



Increasing the time step Δt

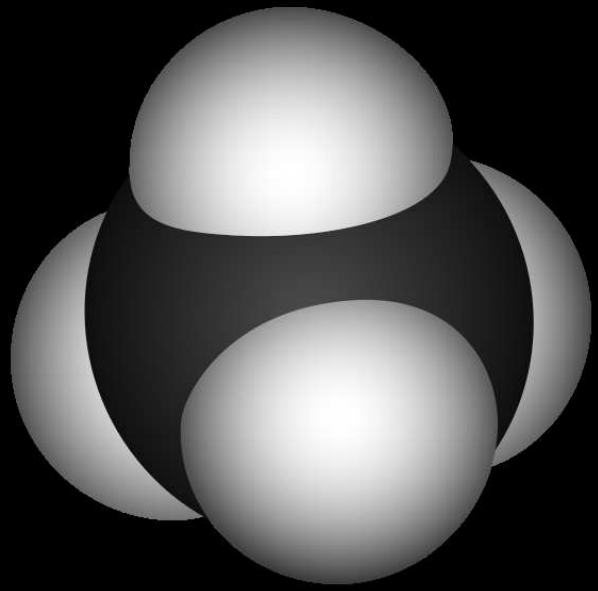
- remove the fastest processes (constraints)

  $\Longrightarrow \Delta t \simeq 2\,\text{fs}$
- smooth potentials
- Maxwell-Boltzmann

$$\langle v \rangle = \sqrt{\frac{8 k_B T}{\pi m}}$$

$\Longrightarrow$ increase the mass

Case study: methane

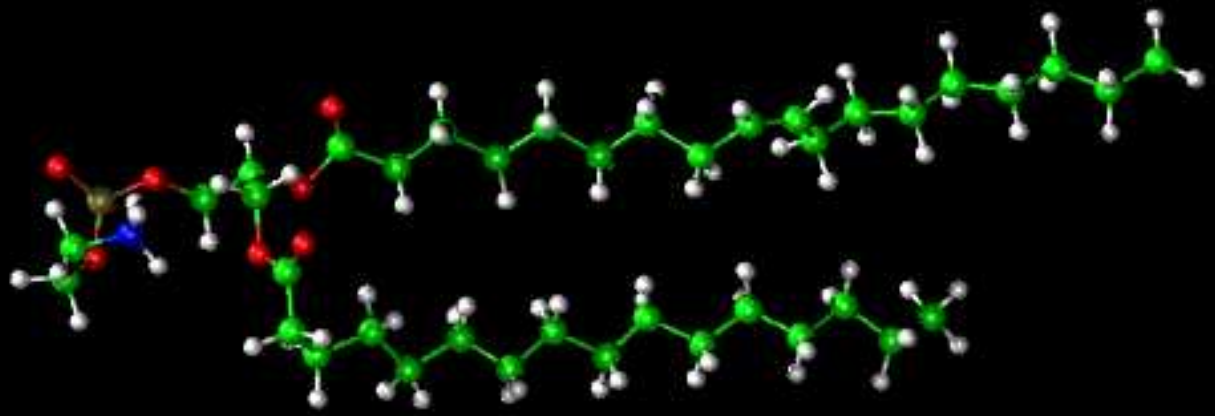


- OPLS-AA: C (opls138), H (opls140)
  $\Longrightarrow$ 25 pair interactions and electrostatics
- OPLS-UA: CH_4 (opls066)
  $\Longrightarrow$ 1 pair interaction without electrostatics

- Vibrational dephasing: all-atom, flexible
- Rotational dynamics, neutron scattering, . . .: all-atom, constraints
- Liquid density, diffusion constant, vaporization enthalpy, . . .: united-atom

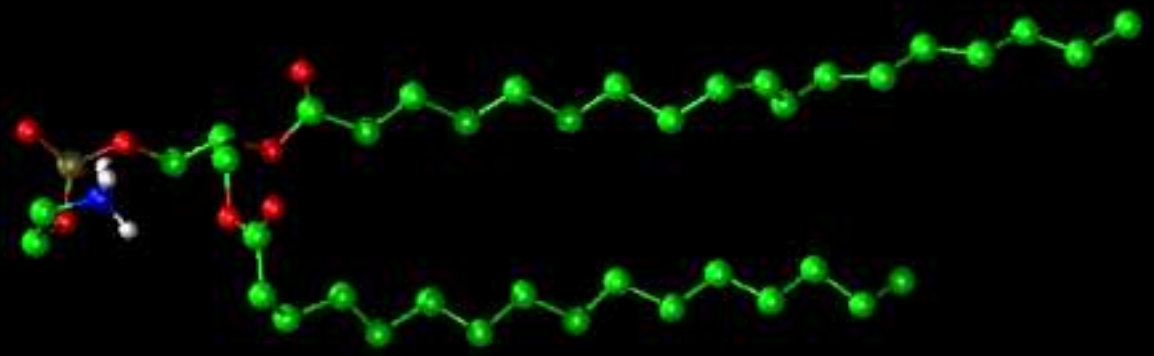
$\Longrightarrow$ we are already scaling on time and length scales

Case study: POPE all-atom
(1-Palmitoyl-2-oleoylphosphatidylethanolamine)



125 atoms
$\Longrightarrow$ 15625 pair interactions
constraints for membrane simulations: $\Delta t \simeq 2$ fs

Case study: POPE united-atom, but polar hydrogens



52 atoms

$\Longrightarrow$ 2704 pair interactions

constraints for membrane simulations: $\Delta t \simeq 2\,\text{fs}$

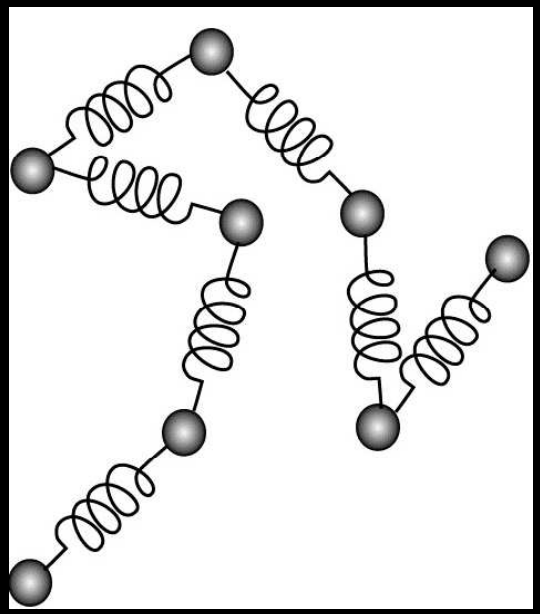
Coarse-Graining

- Goal: realistic simulation of large complex systems for a long time by
  - reduction of interaction sites (MD CPU time $\propto N^2$)
  - elimination of very fast dynamical processes (small masses, large force constants)
  - Maintaining the essential physics of the system

Polymer physics

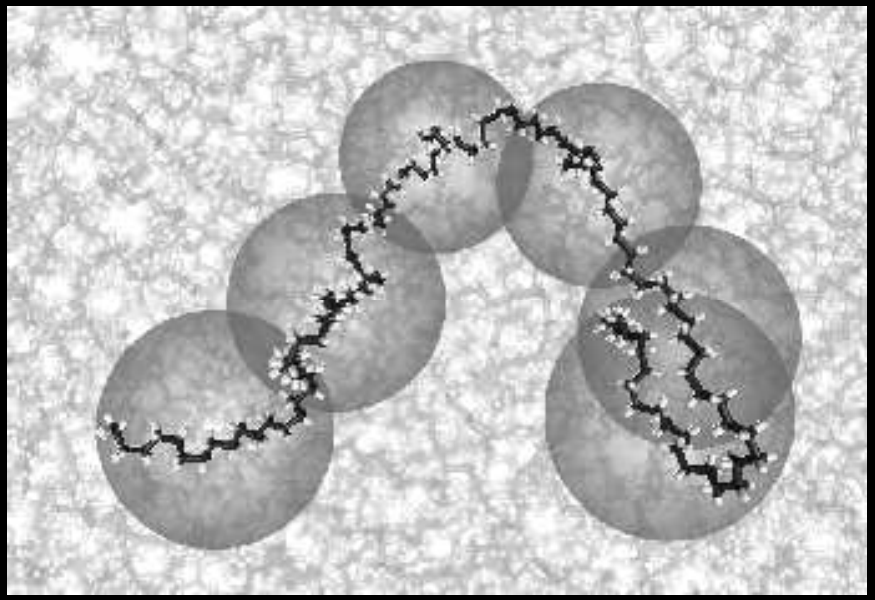
Rouse-Zimm

spring-bead



de Gennes

blob



Protein folding

- M. Levitt, A. Warshel, Nature 253 (1975), 694
- Go models

  2-D: H. Taketomi et al., Int. J. Pept. Res. **7** (1975), 445.

  3-D Example: V.I. Abkevich et al., Folding and Design **1** (1996), 221

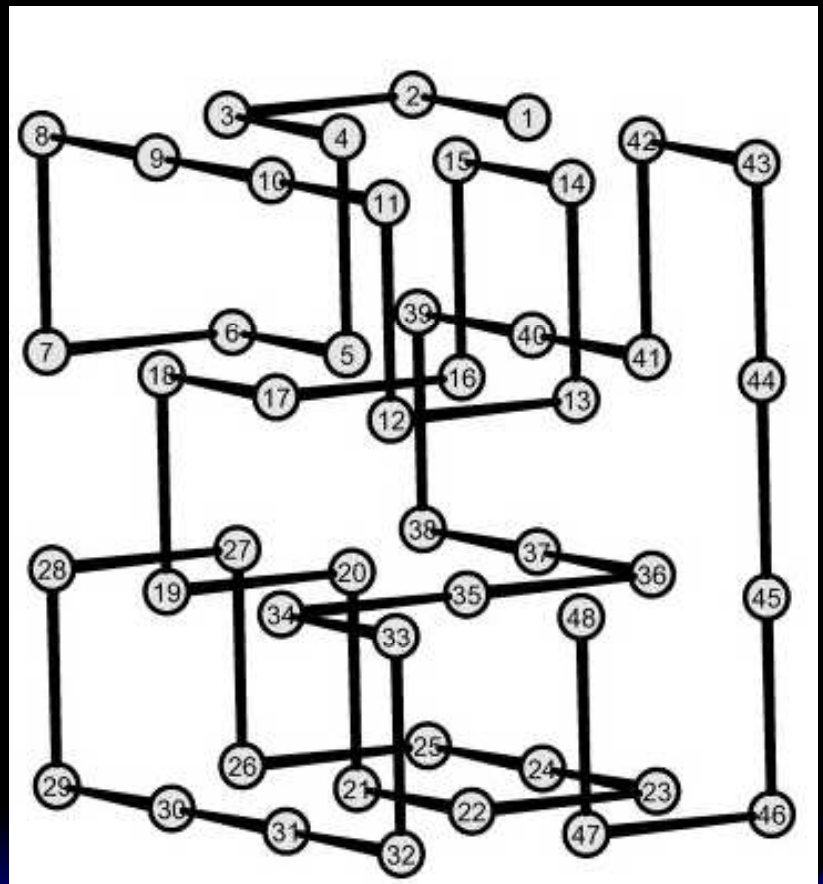
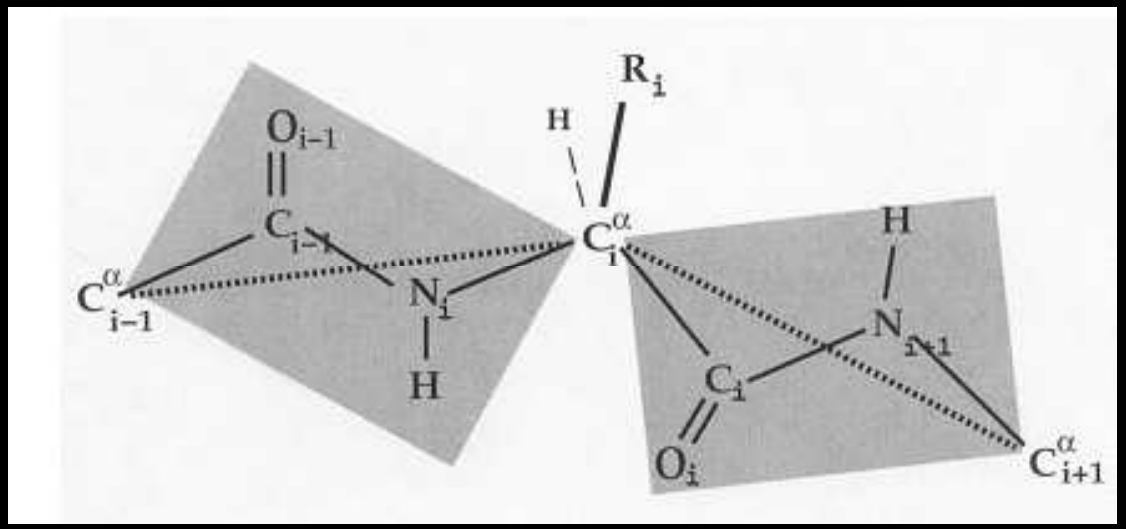


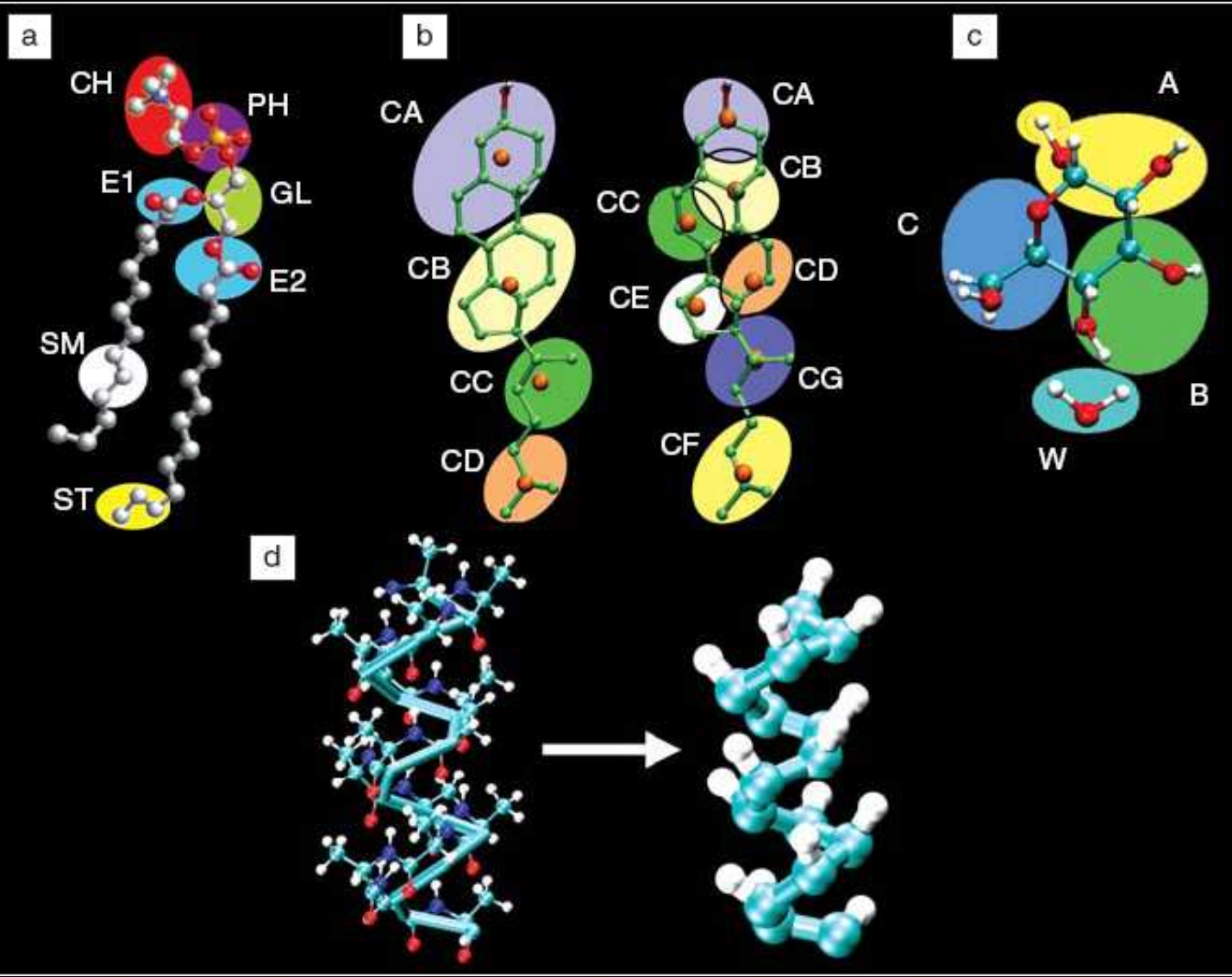
Figure 1. A randomly chosen maximally compact 48-mer on a cubic lattice.

Creating a CG model: Mapping

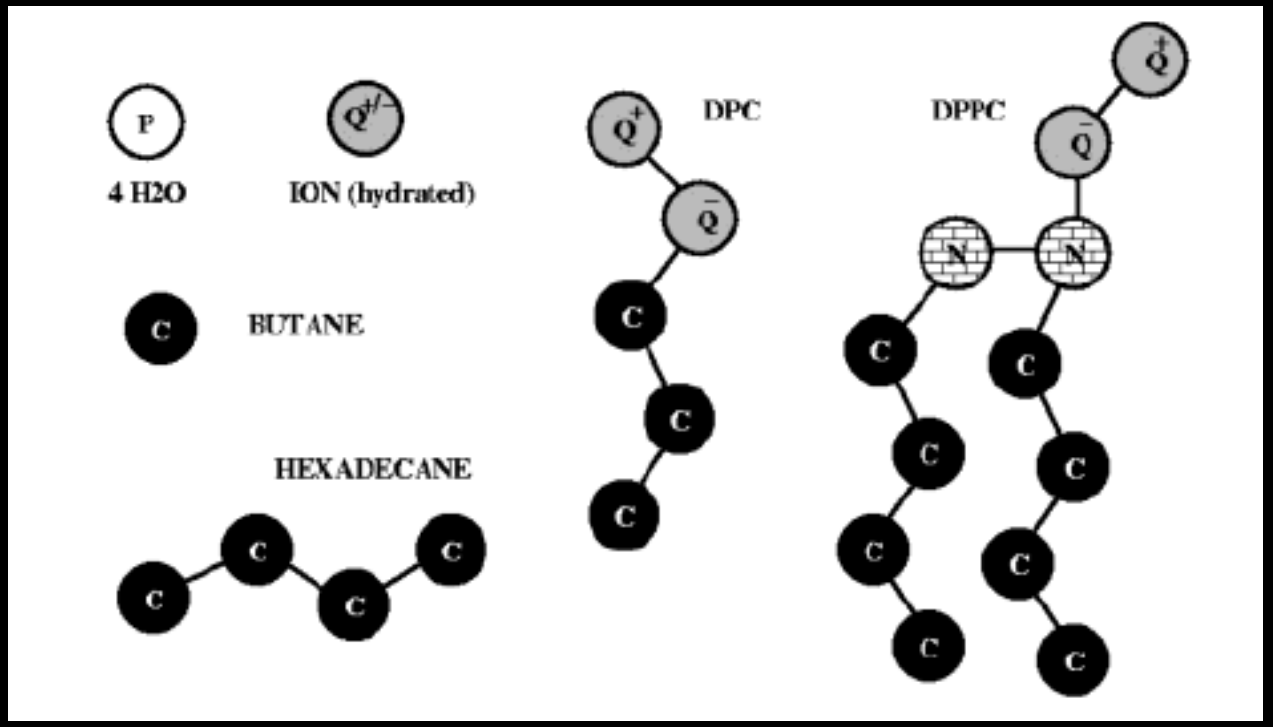
- Physical/chemical intuition: polar/non-polar, hydrogen bond donor/acceptor, charged/neutral, attractive/repulsive, . . .
- Bead at the center of mass of the atoms to be joined (proteins: center at C_α)



MS-CG: G.S. Ayton et al, MRS Bulletin 32 (2007), 929.



MARTINI, mapping scheme 4-1: S.J.Marrink et al., JPCB **108** (2004), 750.



- LJ + Coulomb (reaction field with dielectric constant of 20) + springs (bonds) + harmonic cosines (angles)

system	CG model	density,[b] g cm^{-3}	compressibility,[c] 10^{-5} bar^{-1}	diffusion,[d] 10^{-5} cm^2 s^{-1}
water	P	0.99 (0.99)	6 (4.5)	2.0 (2.3)
butane	C	0.68 (0.58)	28 (> 17)	1.9 (> 5)
hexane	C−C	0.58 (0.66)	14 (17)	0.7 (4)
octane	C−C	0.77 (0.70)	14 (13)	0.6 (2)
decane	C−C−C	0.67 (0.73)	12 (11)	0.35 (1)
dodecane	C−C−C	0.80 (0.75)	12 (10)	0.3 (−)
tetradecane	C−C−C−C	0.71 (0.76)	12 (9)	0.25 (−)
hexadecane	C−C−C−C	0.81 (0.77)	12 (9)	0.2 (−)
octadecane	C−C−C−C−C	0.74 (0.78)	11 (−)	0.2 (0.3)
eicosane	C−C−C−C−C	0.82 (0.79)	11 (−)	0.15 (−)

- Membrane properties

system	simulated area,[a] nm^2	experimental area,[a] nm^2
DPPC	0.47[b] (283 K)	0.46[b] (273 K), 0.48[c] (293 K)
DPPC	0.59 (300 K)	
DPPC	0.64 (323 K)	0.64 (323 K)
DPPC	0.66 (338 K)	0.64−0.67[d] (333 K), 0.67 (338 K)
DLPC	0.60 (300 K)	0.57 (293 K), 0.63 (303 K)
DLPC	0.66 (338 K)	0.64−0.68[d] (333 K), 0.71 (338 K)
DSPC	0.66 (338 K)	0.65 (333 K), 0.66 (338 K)
DOPC	0.67 (300 K)	0.72 (303 K)
DPPE	0.62 (338 K)	0.61 (342 K)
DOPE	0.61 (273 K)	0.65 (271 K)
DLPE	0.55 (300 K)	0.51 (308 K)
POPE	0.59 (300 K)	0.57 (303 K)

- Refinement (water/octanol partition) and inclusion of benzene and cholesterol
  $\Longrightarrow$ improper dihedrals

S.J.Marrink et al., JPCB **111** (2007), 7812.

Parameters for the CG force field

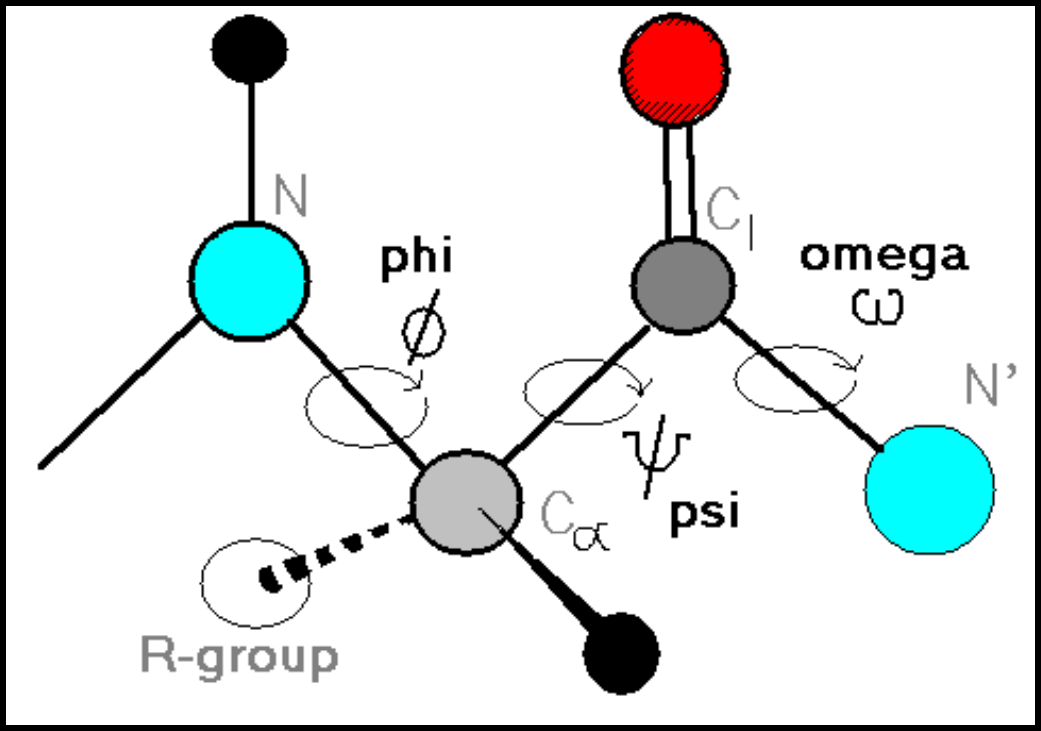
- Boltzmann factor:
  - Reverse Monte Carlo applied to radial distributions functions of atomistic simulations

    T. Murtola et al., JCP **121** (2004), 9156; **126** (2007), 075101; **131** (2009), 055101; PCCP **11** (2009), 1869
  - Force matching of atomistic systems

    S. Izvekov e G.A. Voth, JPCB **109** (2005), 2469; JCP **123** (2005), 134105; JCTC **2** (2006), 637
- Experimental properties (density, partition coeficients, membrane properties, solvation energies)

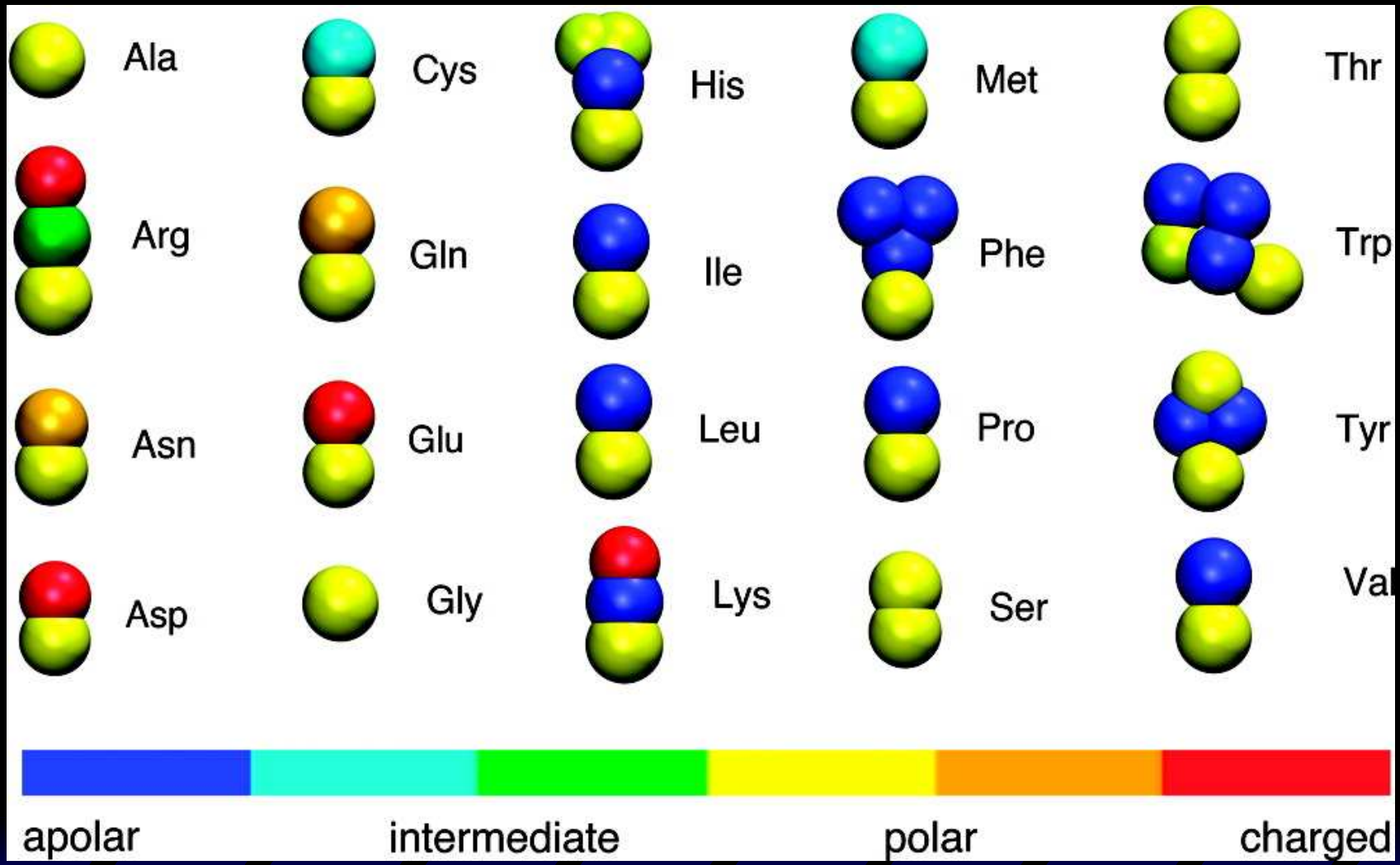
Proteins

- Secondary structure

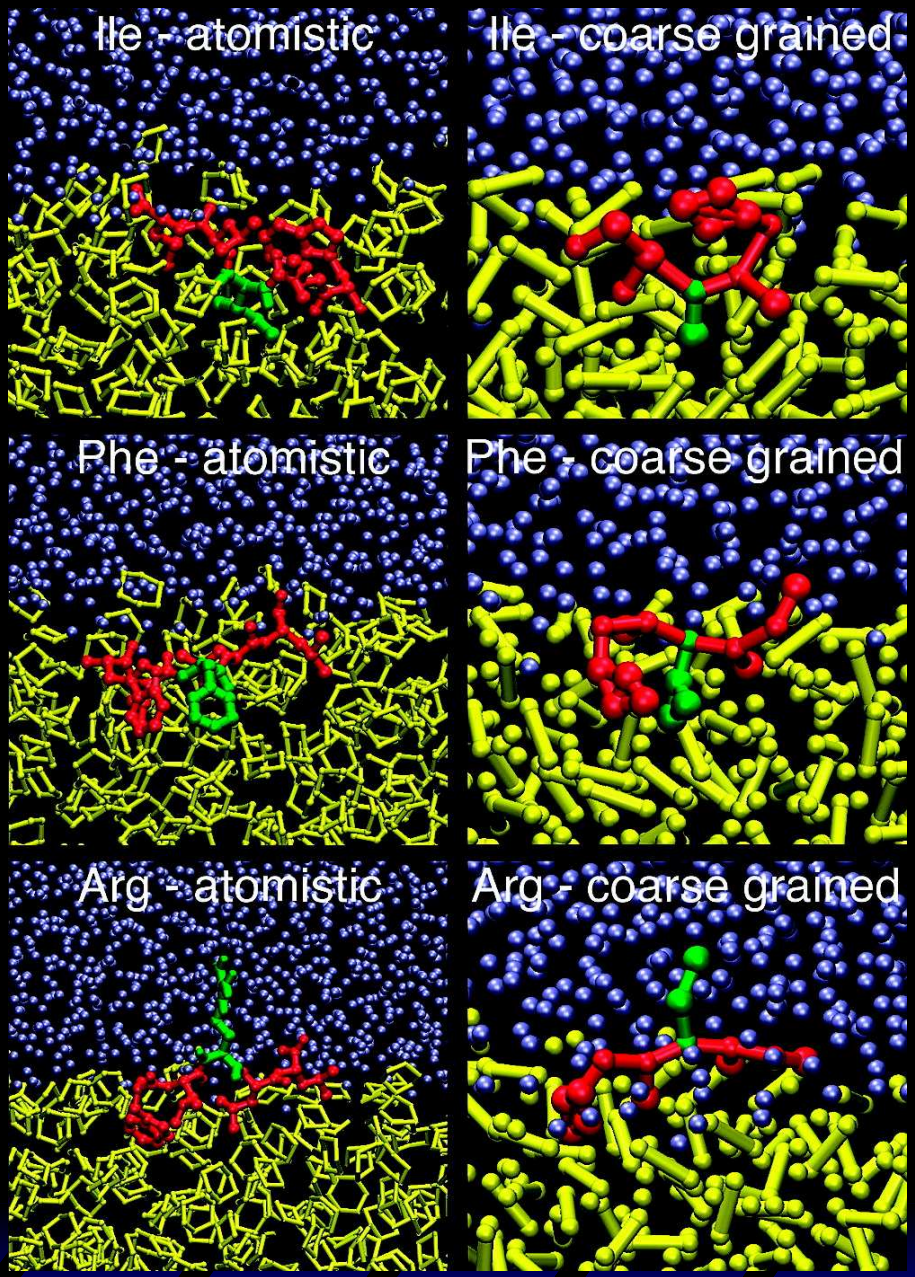


http://www.bioinf.org.uk/teaching/bbk/molstruc/practical2/peptide.html

- MARTINI proposal: beads at C_α (main chain) with different lateral groups (side chain)

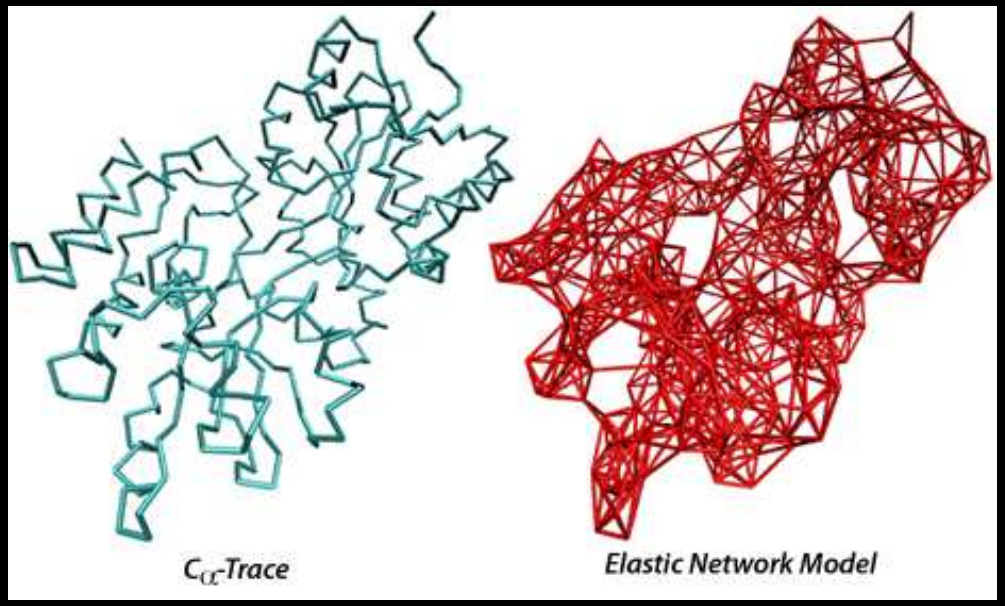


L. Monticelli et al., JCTC **4** (2008), 819.



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

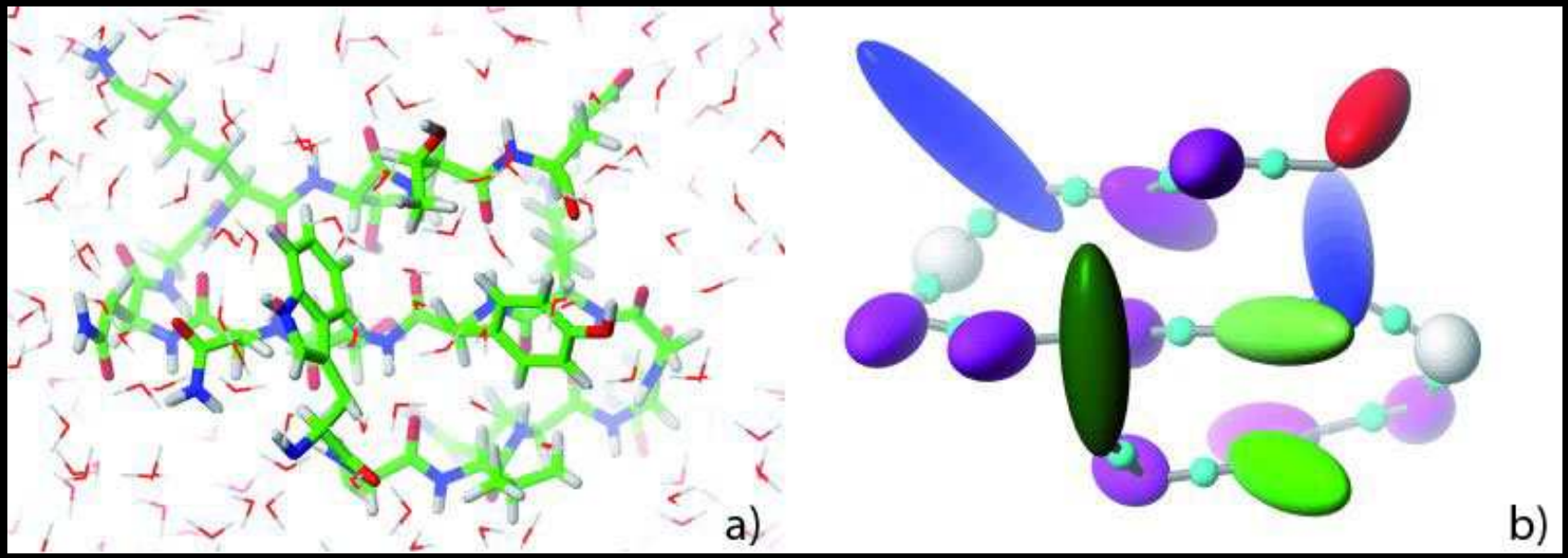
with elastic network



http://lorentz.immstr.pasteur.fr/nma

$\Longrightarrow$ conserved secondary structure during the simulation

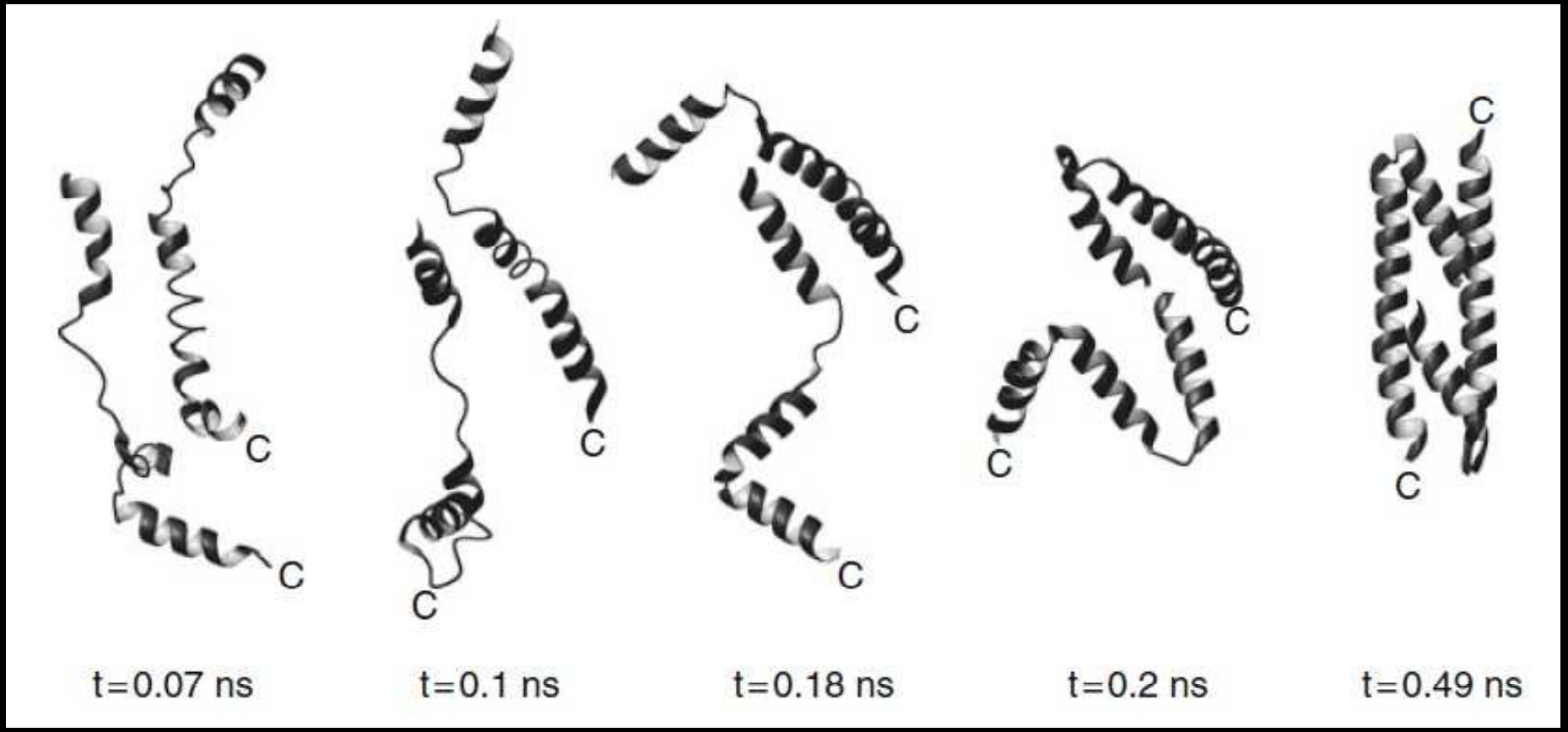
- UNRES proposal (H.A. Scheraga group): beads at C_α, side chains, and in the middle between the C_α's



www.unres.pl

not potential, but free energy based

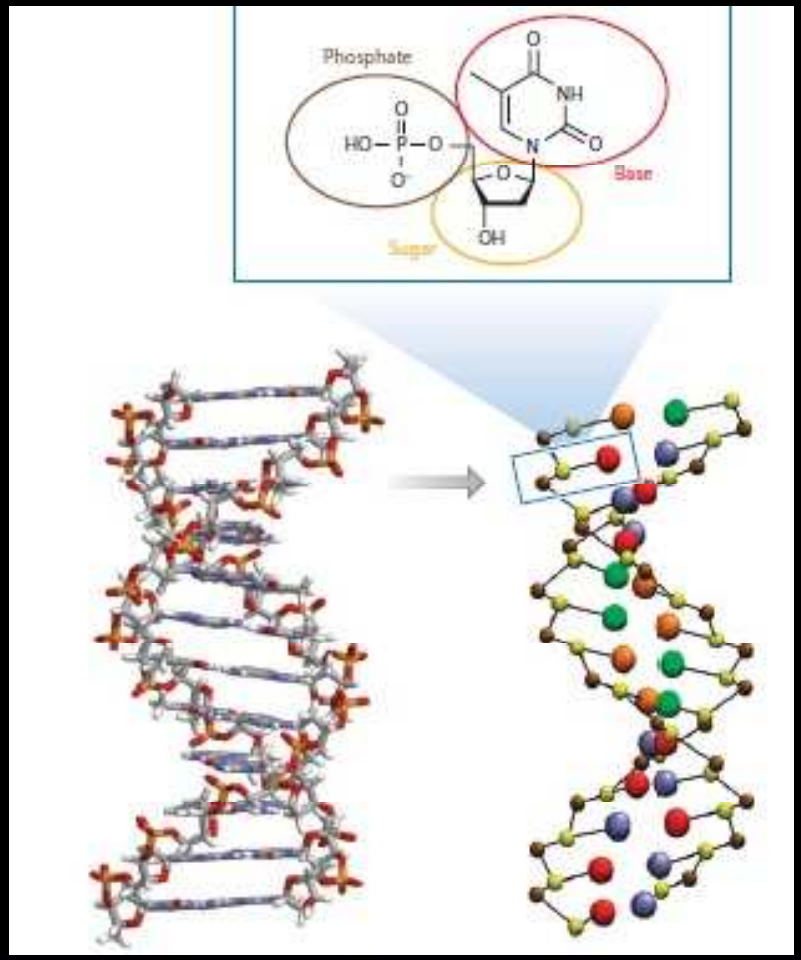
rather successful in CASP competition



1G6U (dimer): A.V. Rojas et al., JPCB **111** (2007), 293.

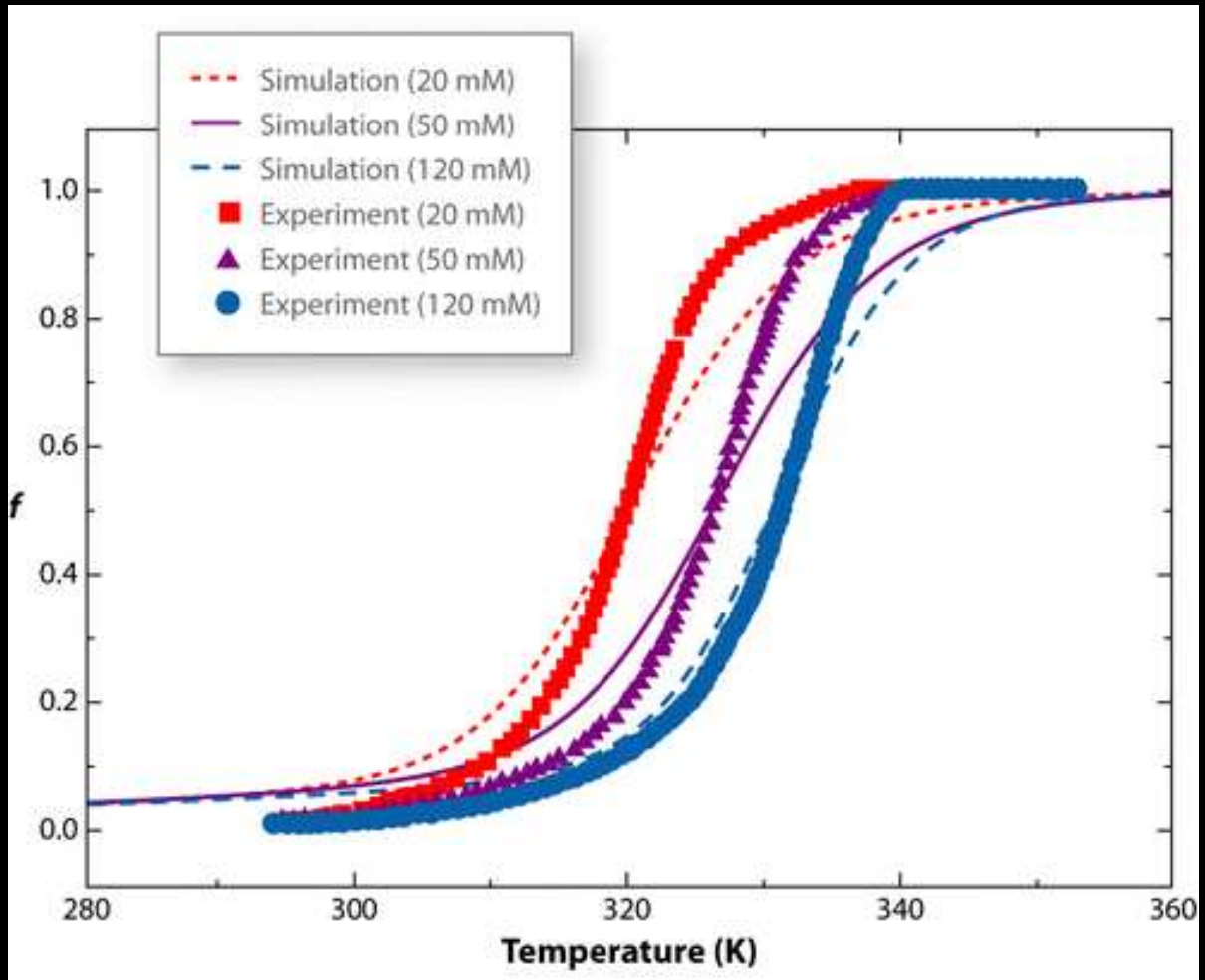
DNA, RNA

- 3SPN (three sites per nucleotide)



J.J. de Pablo, Annu. Rev. Phys. Chem. **62** (2011), 555.

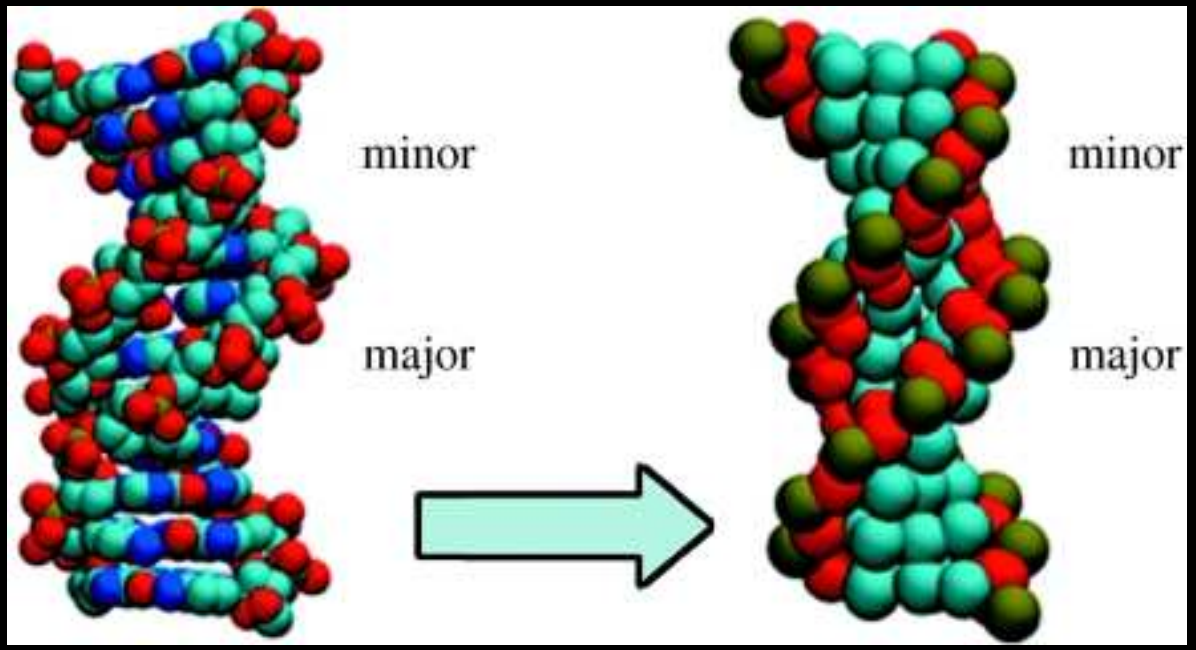
Melting curves: shape of curves and concentration effects



de Pablo JJ. 2011.
Annu. Rev. Phys. Chem. 62:555–74

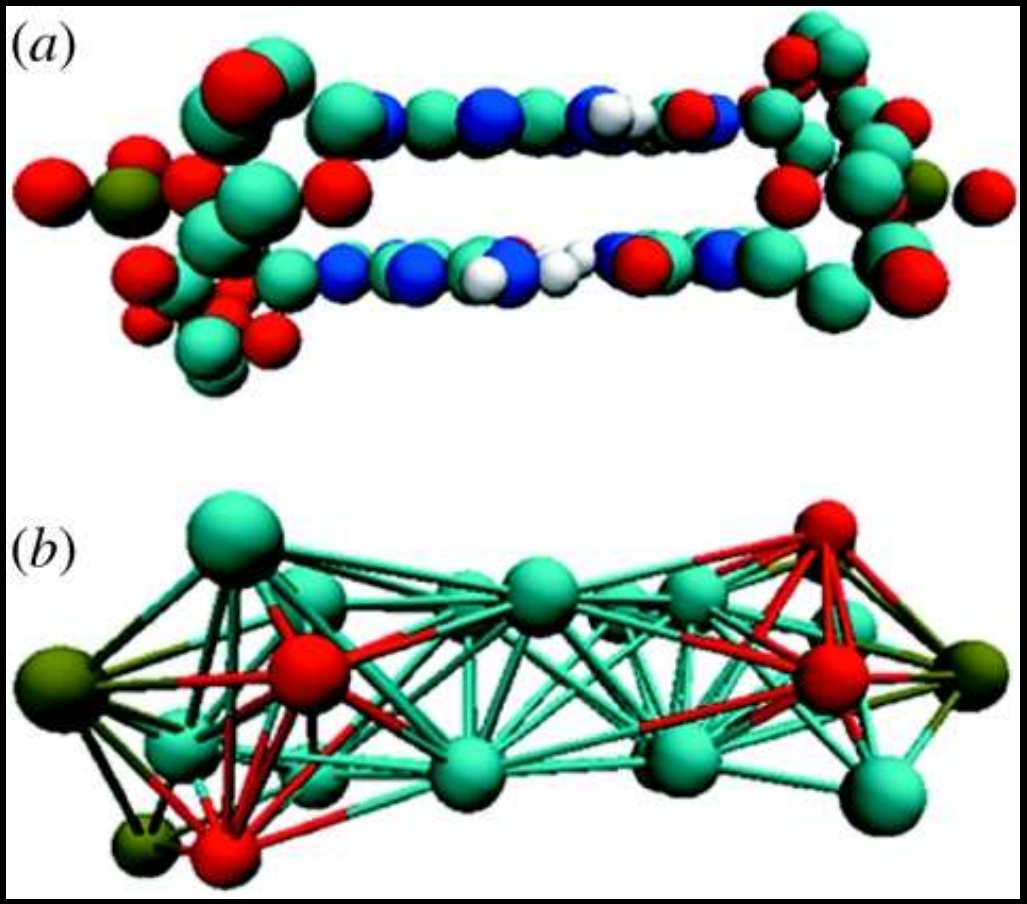
- MARTINI

  12 base pairs (1BNA.pdb)

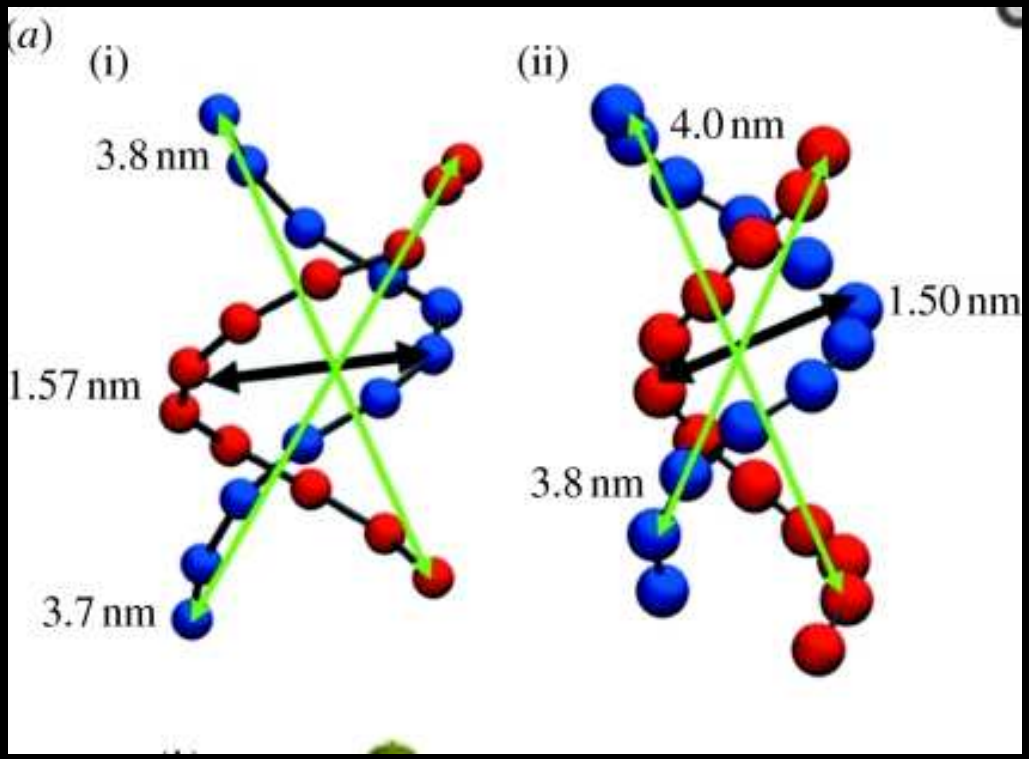


S. Khalid et al., J. R. Soc. Interface **5** (2008), 241.

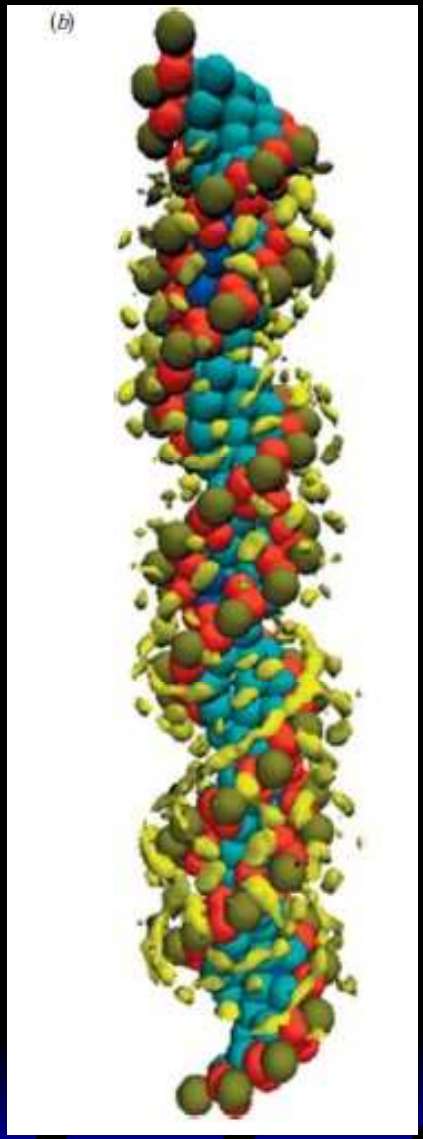
elastic network



distances (pdb vs. 30 ns)

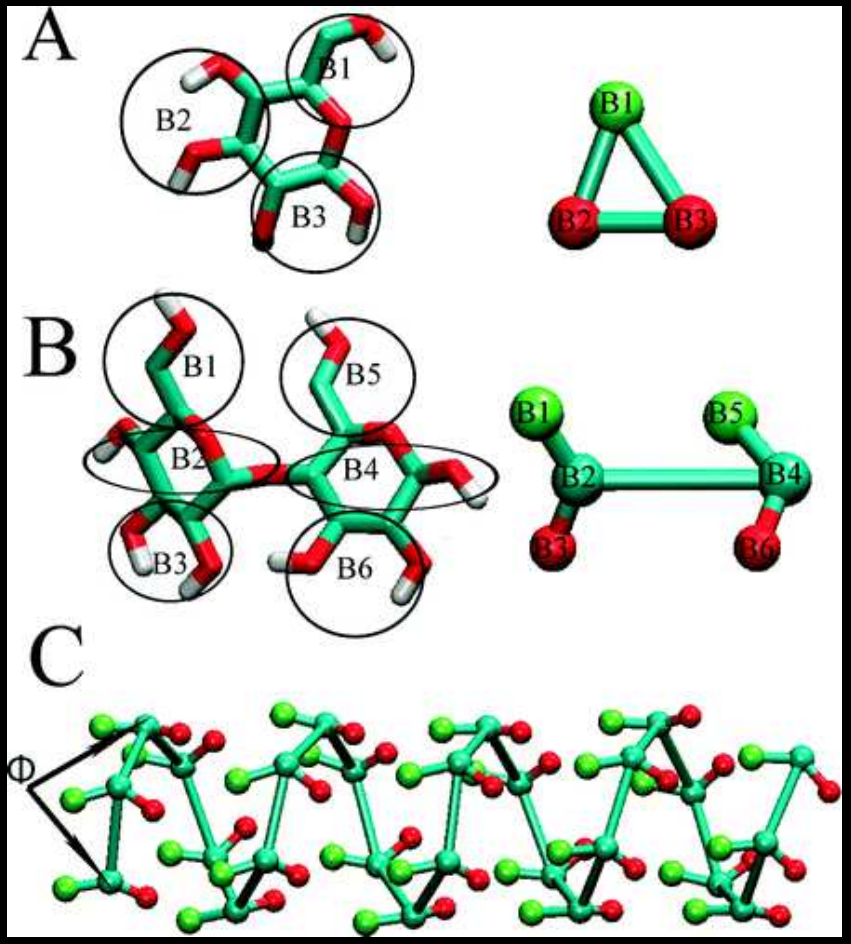


Density of Na^+ around the 12 base pairs



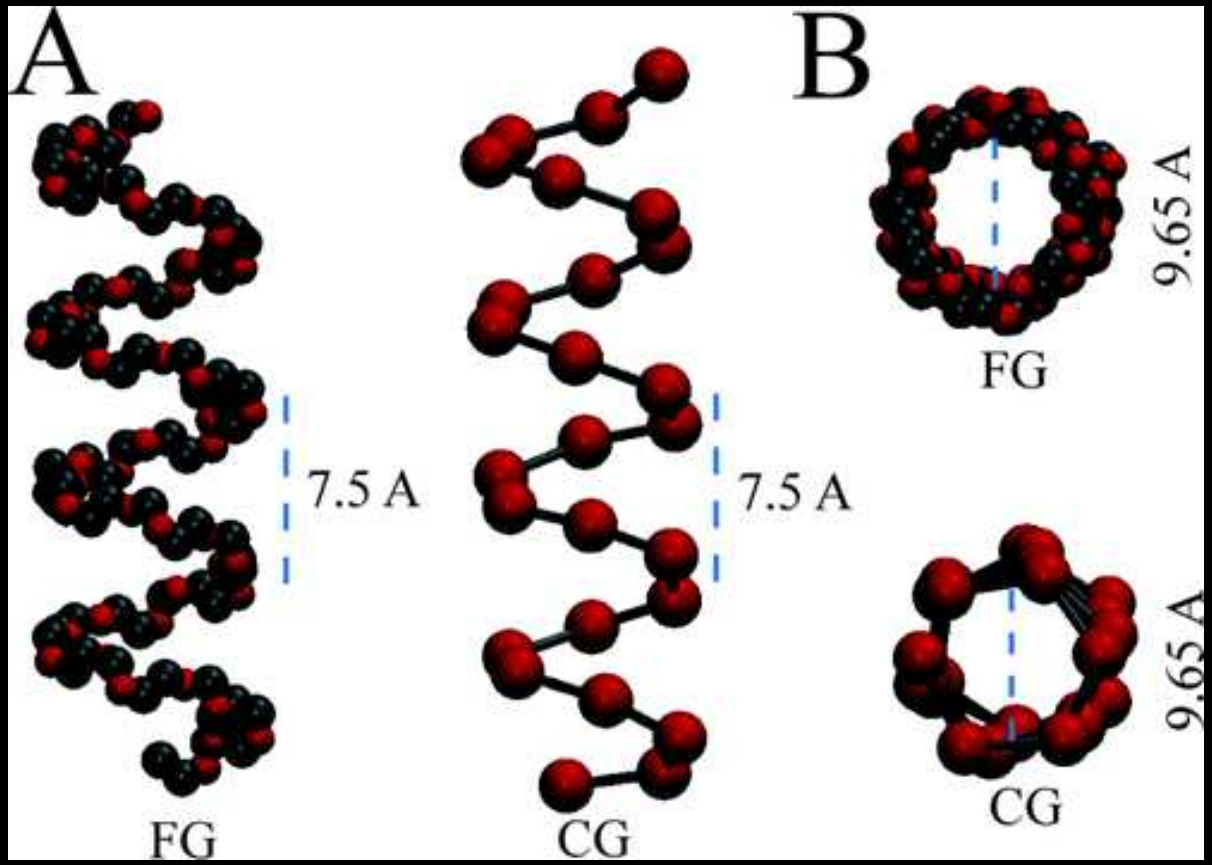
Sugars

- Sweet MARTINI



C.A. López et al., JCTC **5** (2009), 3195.

Amylose in nonane (atomistic vs. CG)

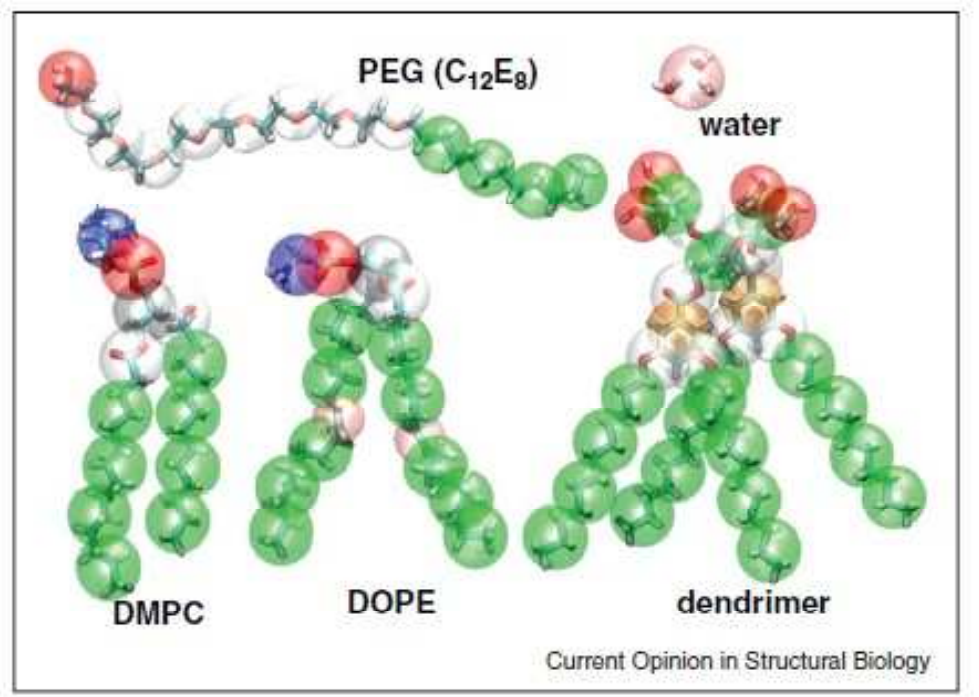


Amphiphilic Molecules

- SDK: Shinoda, de Vane, Klein

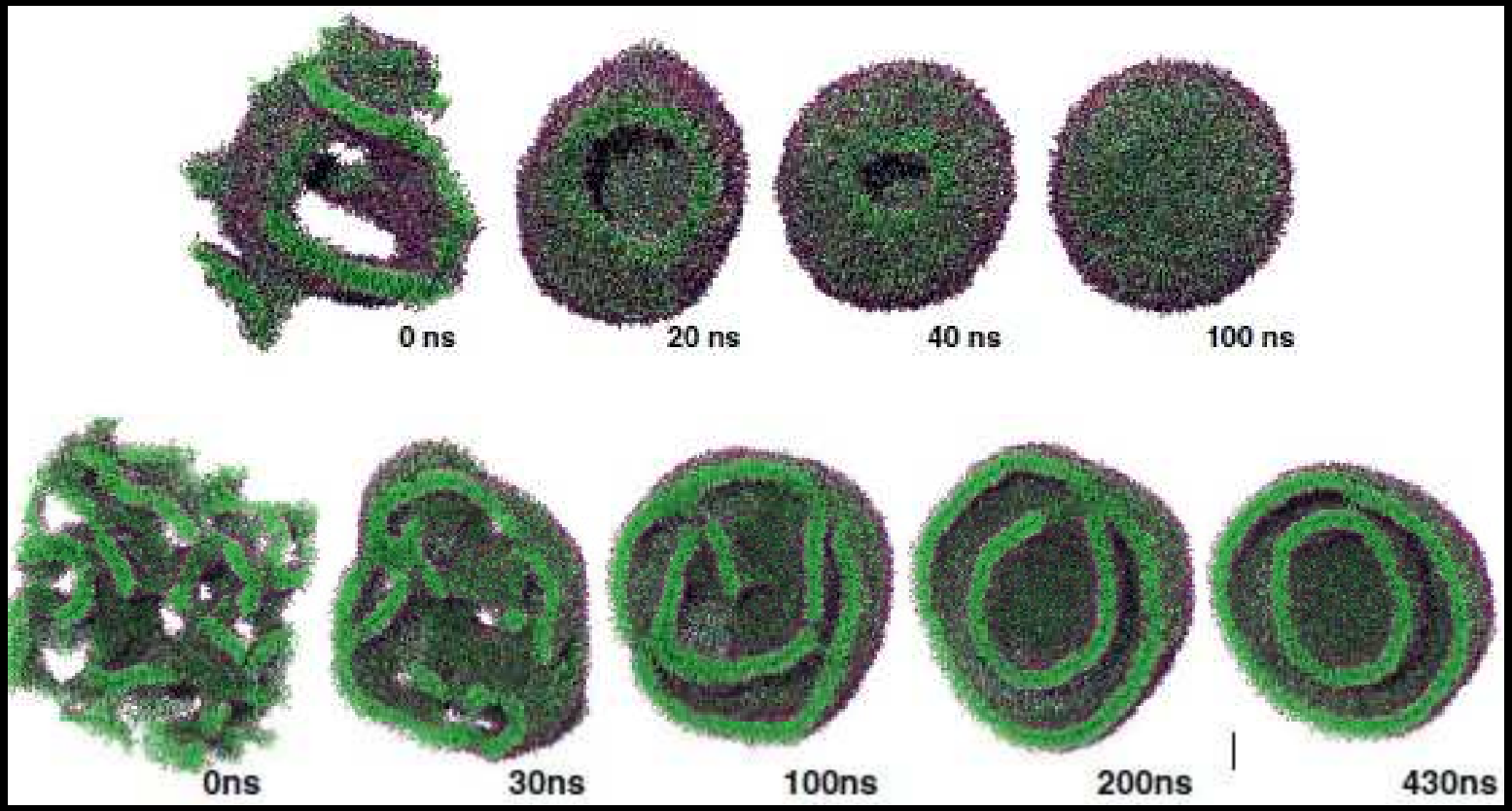
  Parameterization against experimental data

  Mapping:



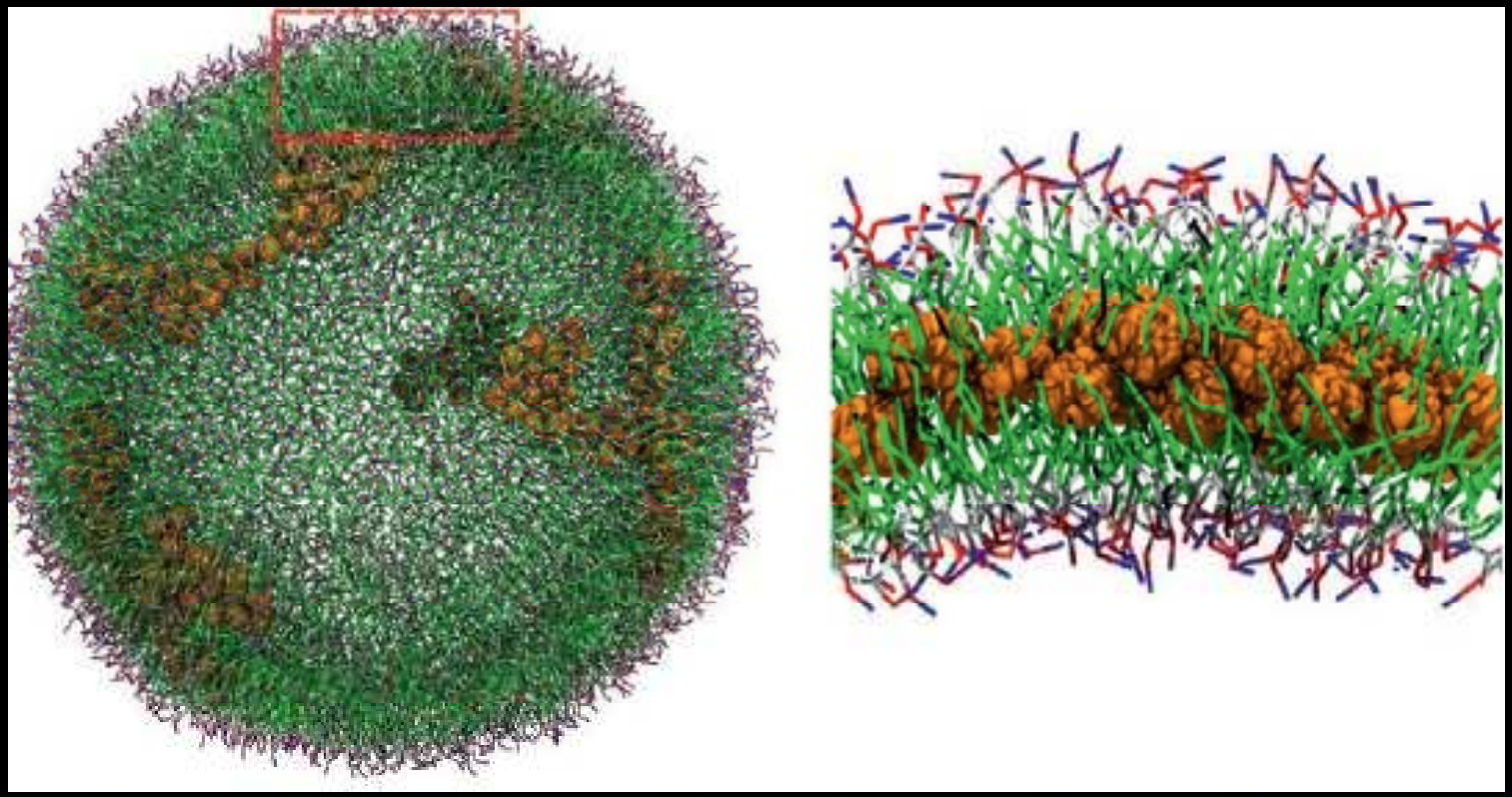
W. Shinoda et al., Curr. Opin. Struct. Biol. **22** (2012), 175.

Simulation of 5000 (vesicle) and 20000 DMPC (multilamelar vescicle) in water

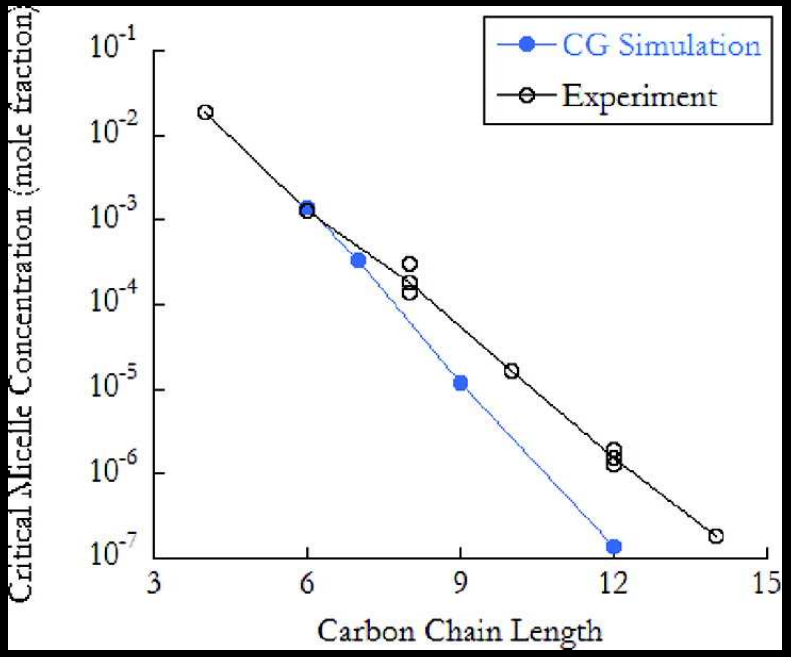


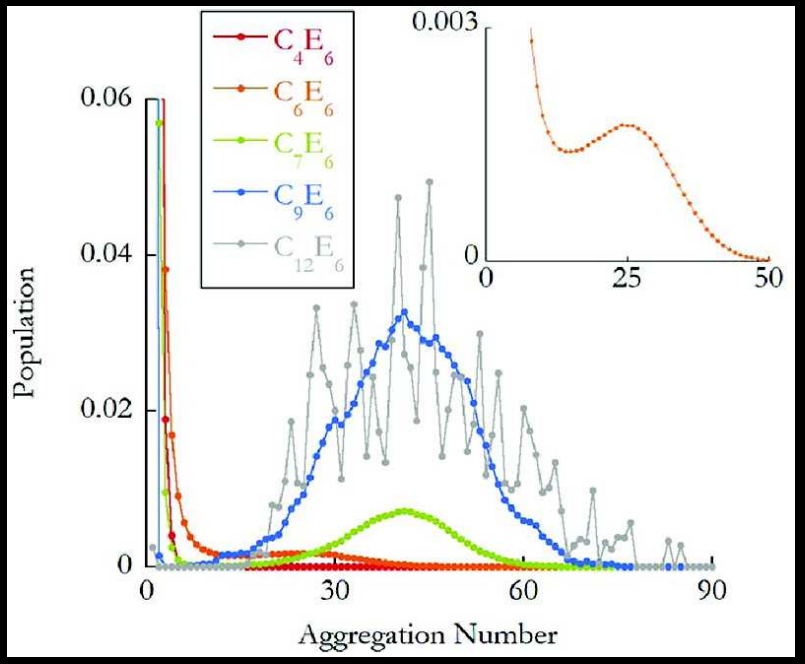
W. Shinoda et al., Curr. Opin. Struct. Biol. **22** (2012), 175.

5000 DMPC with 330 C_{60} in water after 1.7 μs



Micellization of C_nE_6





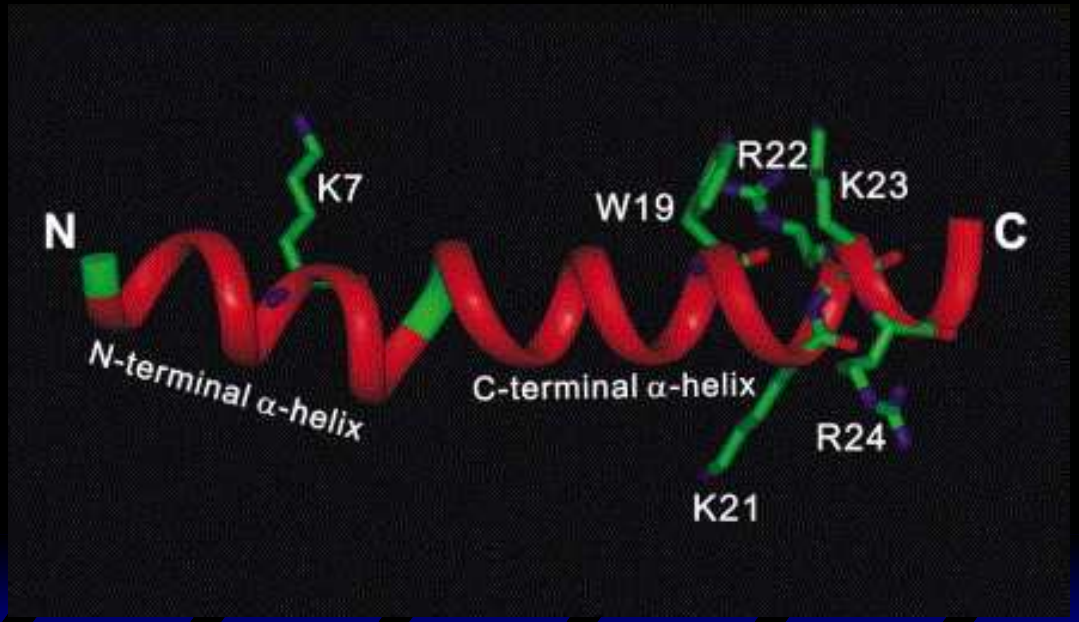
B.G. Levine et al., JCTC **7** (2011), 4135

Protein Membrane Interaction

- AMPs:

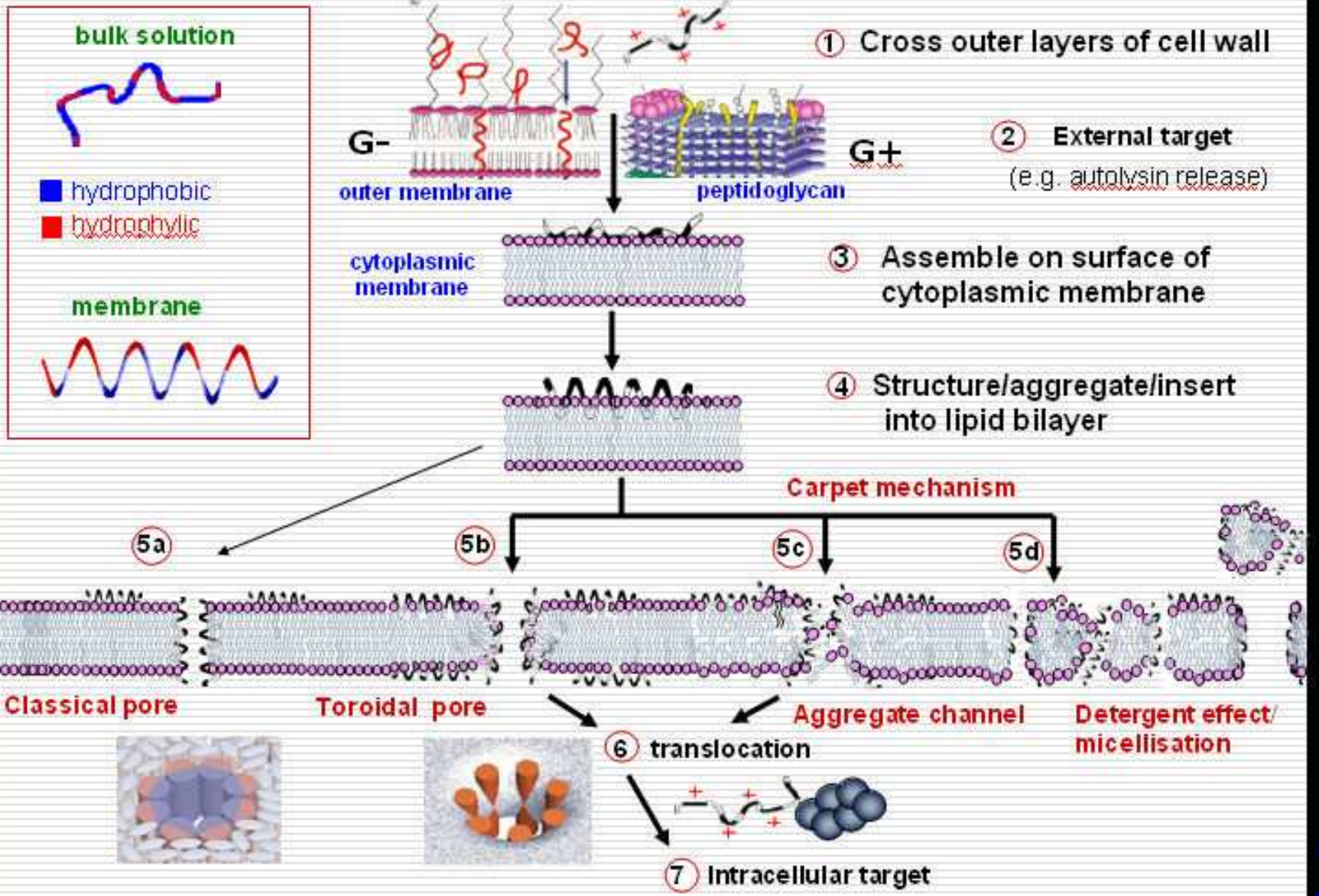
  Magainin-H2 (23aa, +3)

  Melittin (26aa, +6)

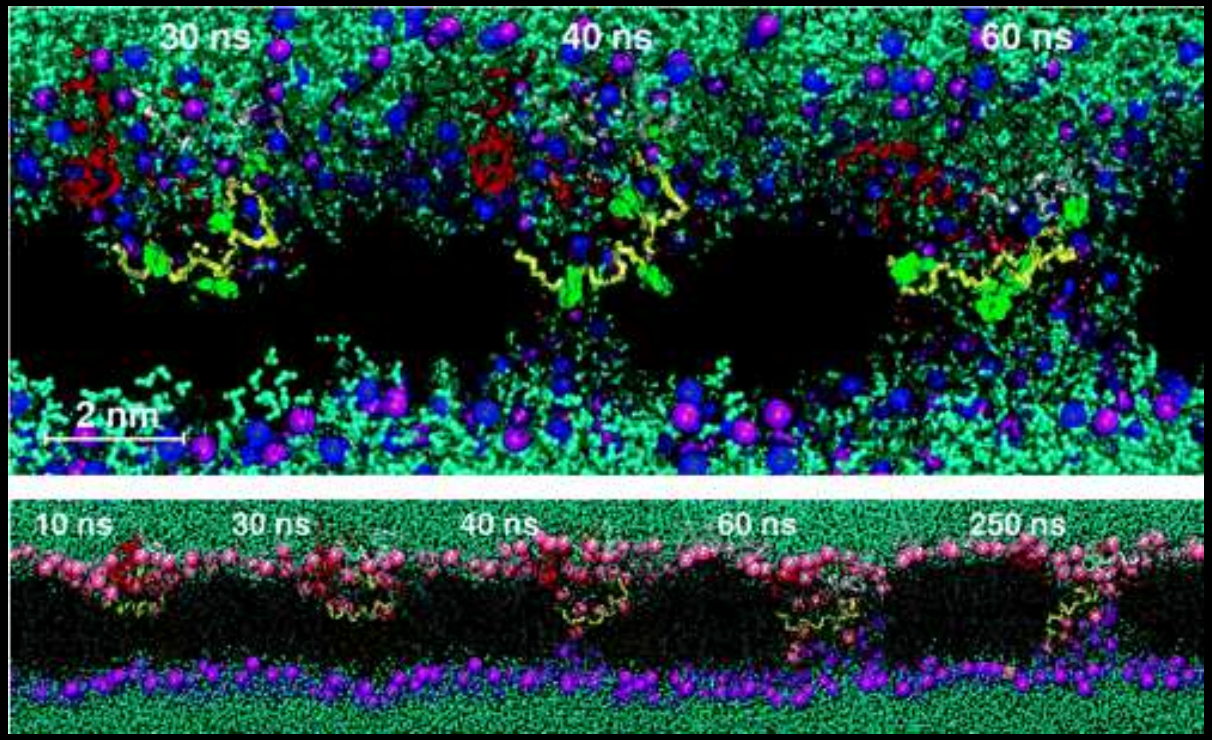


H.L. Liu, C.M. Hsu, CPL **375** (2003), 119.

Mode of action of α-helical AMPs



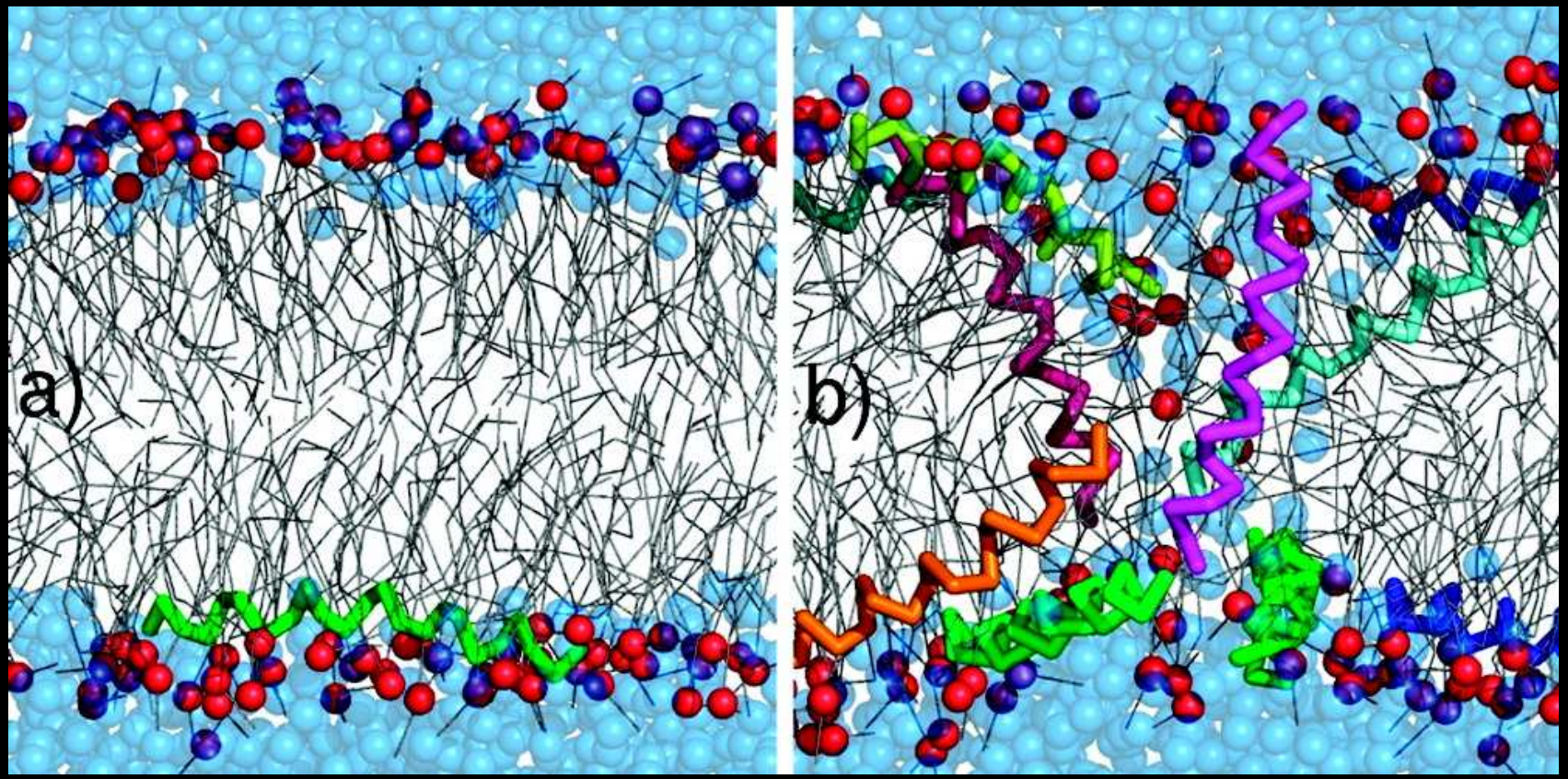
Magainin-H2: atomistic simulation of 4 polypeptides at a DPPC bilayer



H. Leontiadou et al., JACS **128** (2006), 12156.

disordered toroidal pore

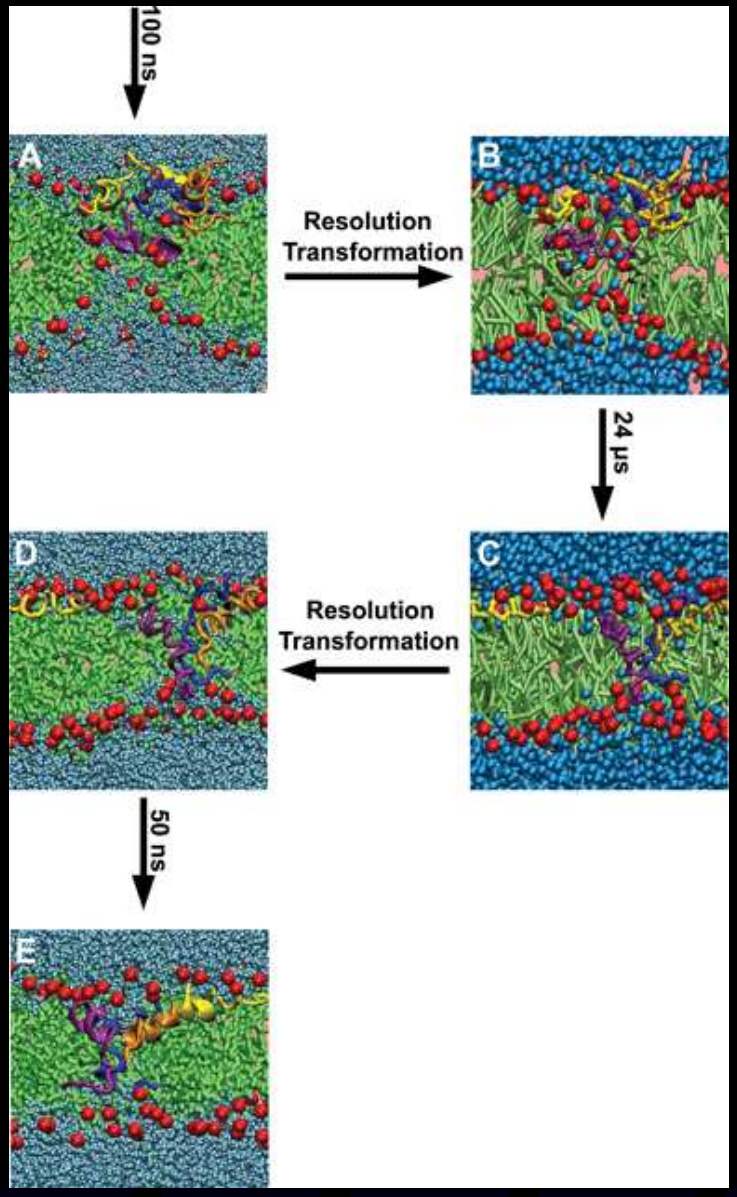
CG self-aggregation of polypeptides with DPPC in water



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

disordered toroidal pore with two magainins in the center and others at the edge of the pore

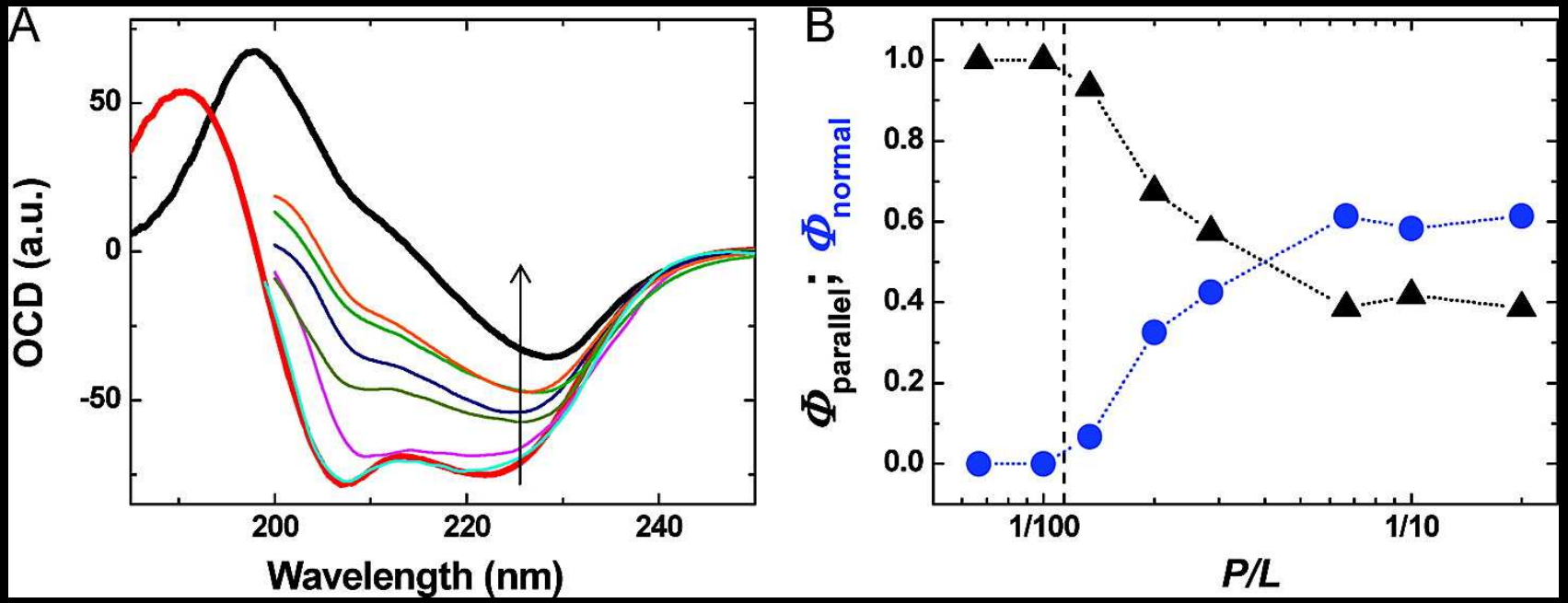
"multiscaling" of 4 polypeptides in a DPPC bilayer



A.J. Rzepiela et al., Faraday Discuss. **144** (2010), 431.

- different orientations and transitions of the polypeptides
- flip-flop of DPPC in the pore region
- water content within the center of the pore depends on the helicity (max. for 65 % helicity; pore closes for complete α-helical polypeptides)

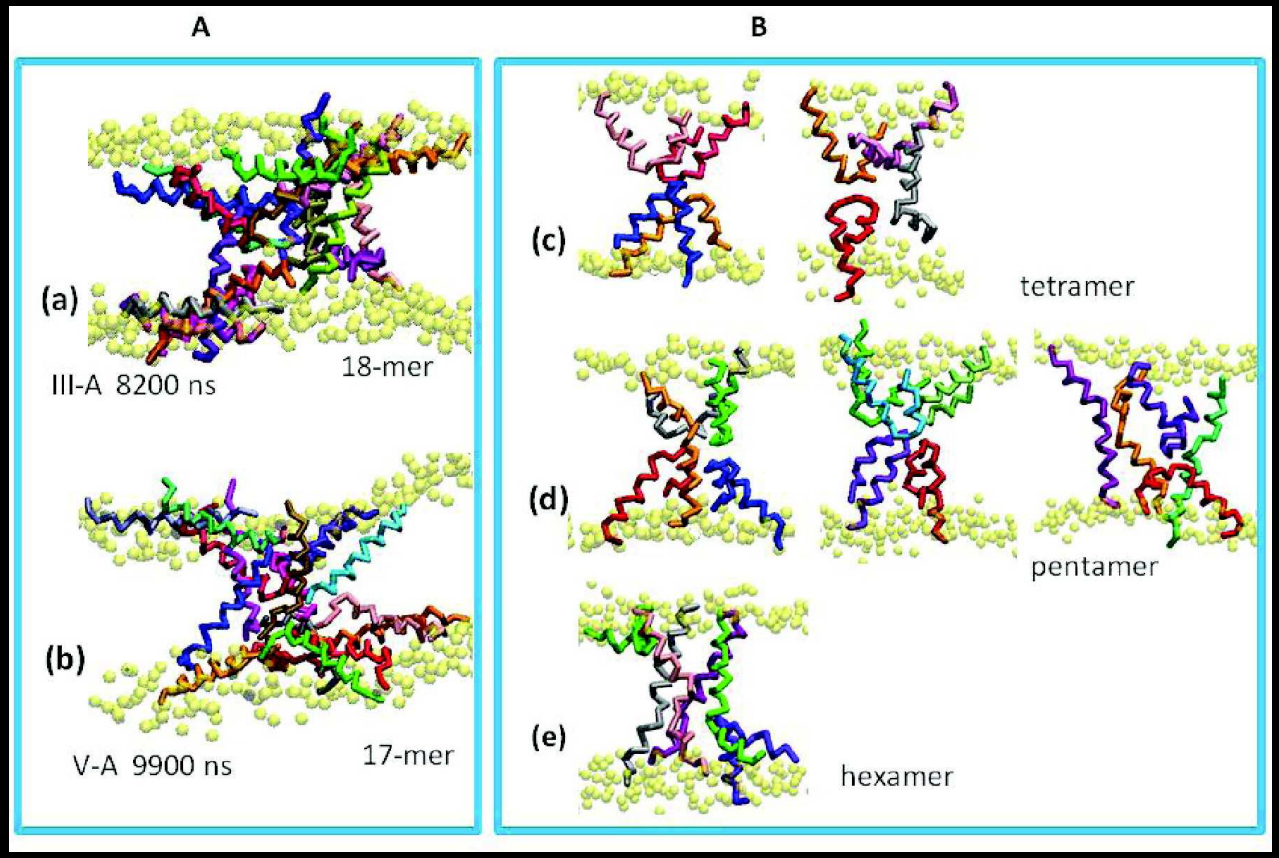
Melittin: recent experimental observations (POPC/DPPC)



Y. Yu et al., JACS **132** (2010), 195.

$\Longrightarrow$ transmembrane orientation for concentrations above 1 melittin per 40 lipid molecules

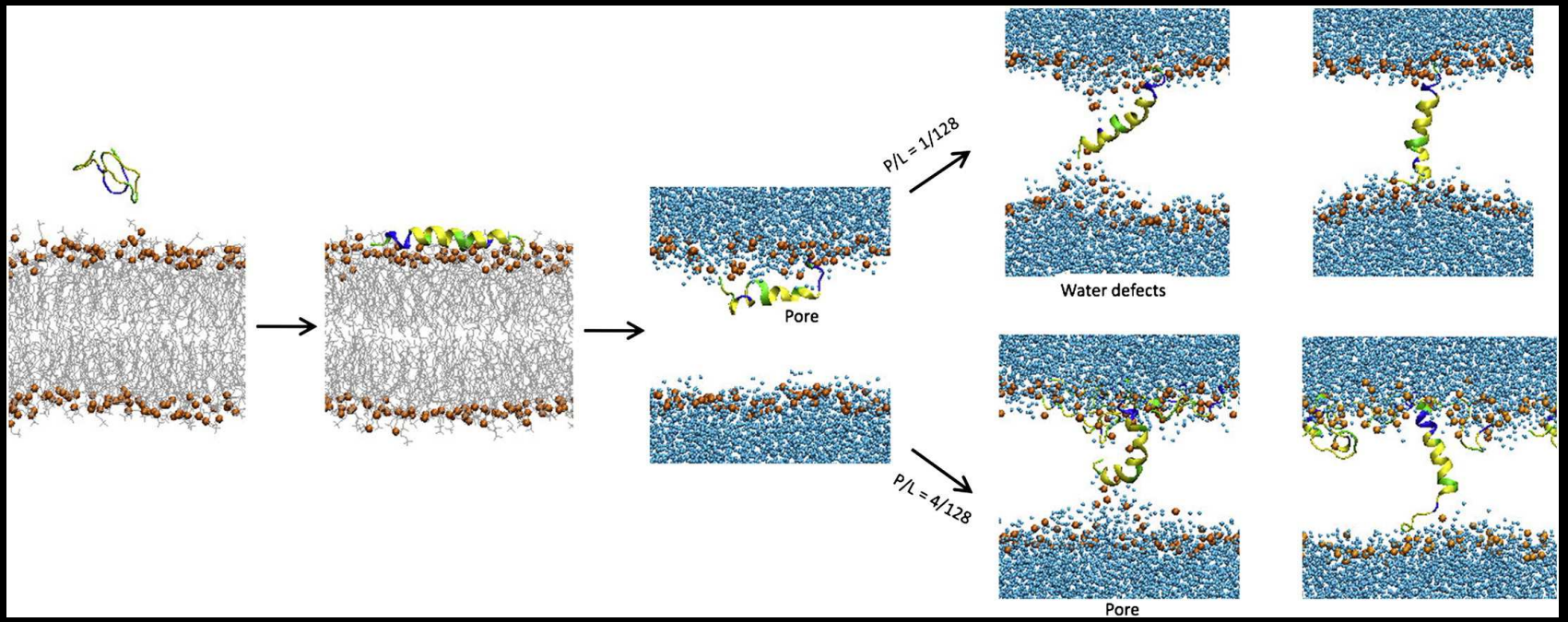
CG of 24 magainins (A) and melittins (B) at a DPPC bilayer (512 molecules)



K. Santo and M.L. Berkowitz, JPCB **116** (2012), 3021.

Melittin: Pore formation only for initial configurations with melittins at both interfaces

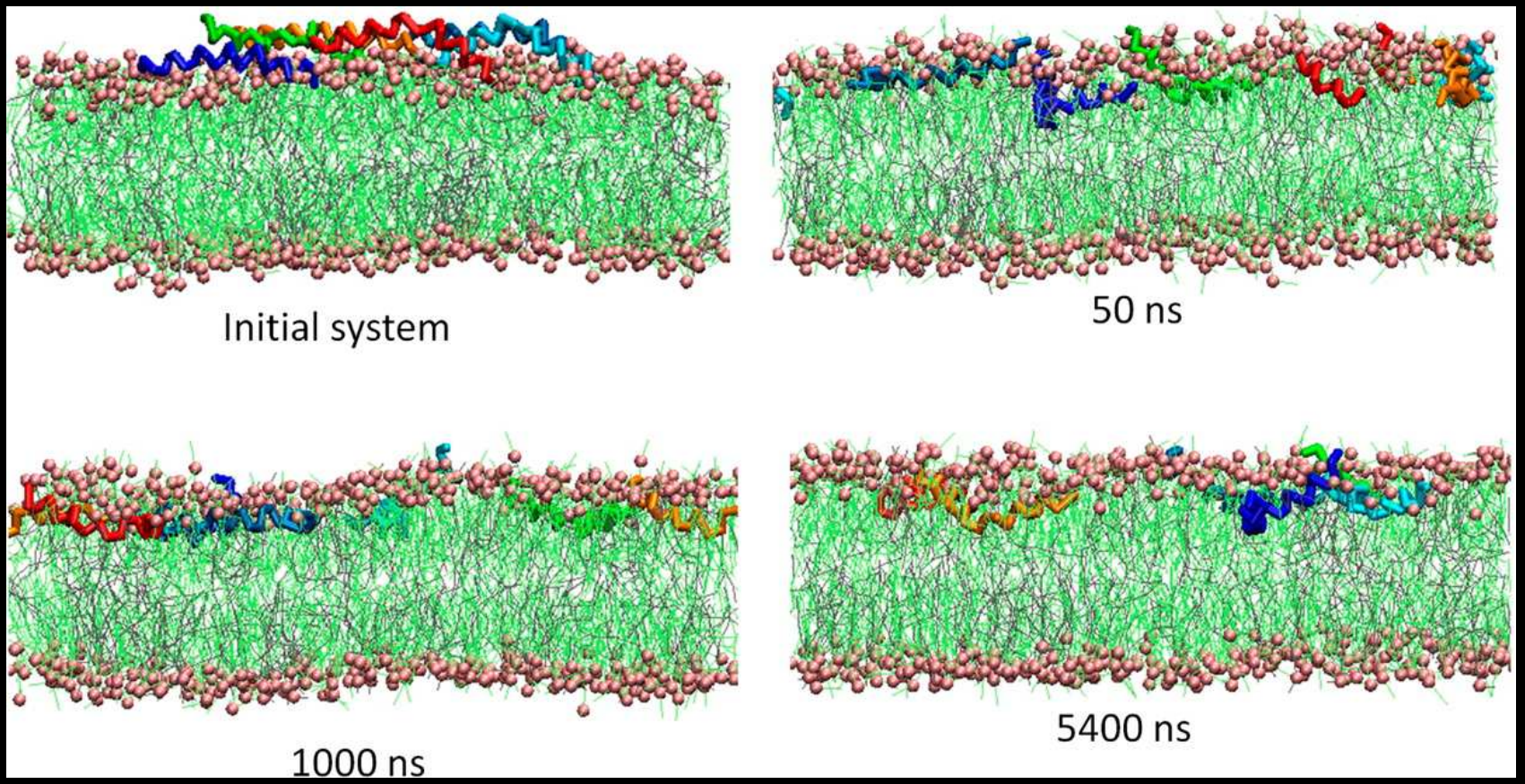
Pulled atomistic molecular dynamics of melittin at a POPC bilayer



S.J. Irudayam and M.L. Berkowitz, BBA **1818** (2012), 2975.

$\Longrightarrow$ higher melittin concentrations for reorientation process

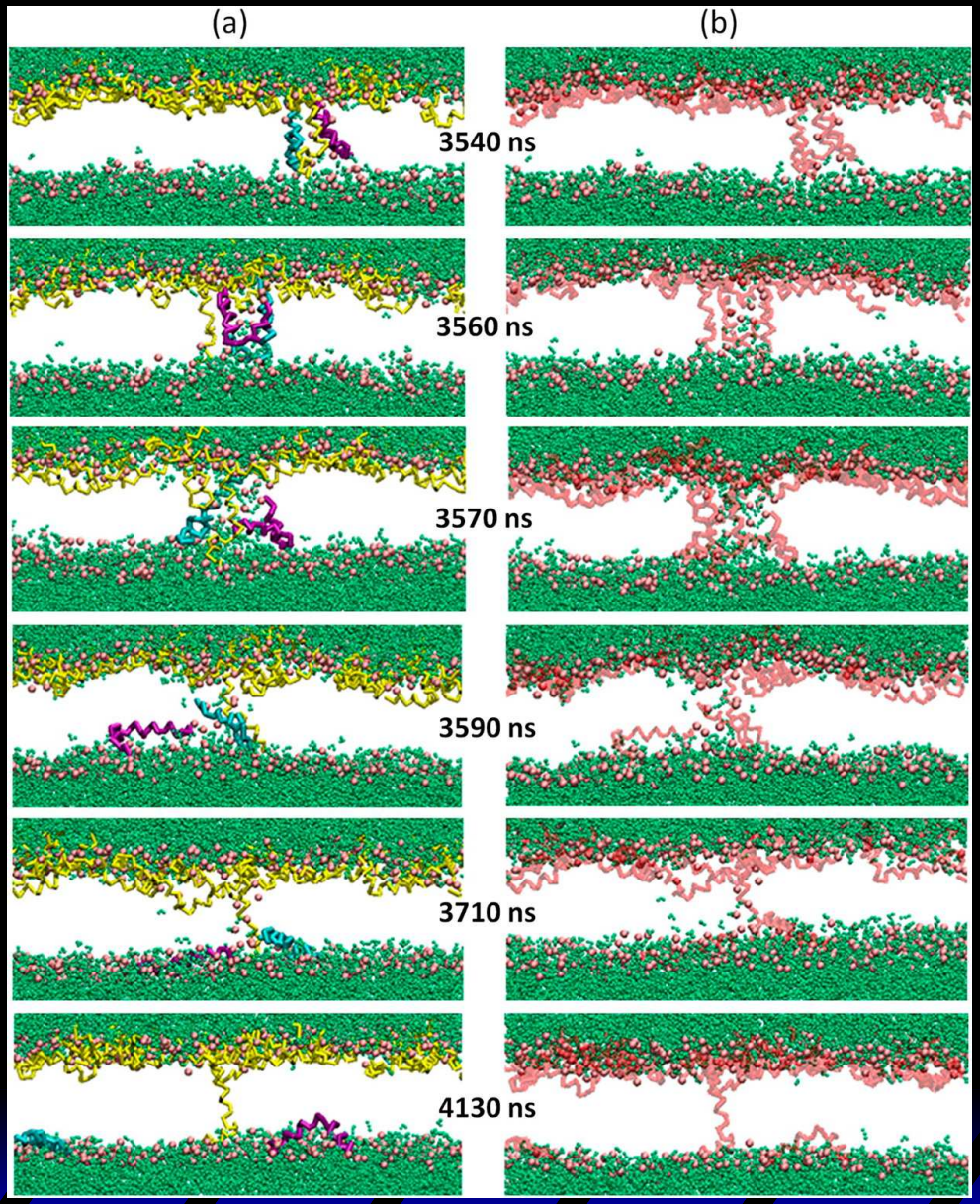
CG of melittin at DPPC/POPG (7:3) bilayer (melittin/lipid=1/100)



K.P. Santo et al., JPCB **117** (2013), 5031.

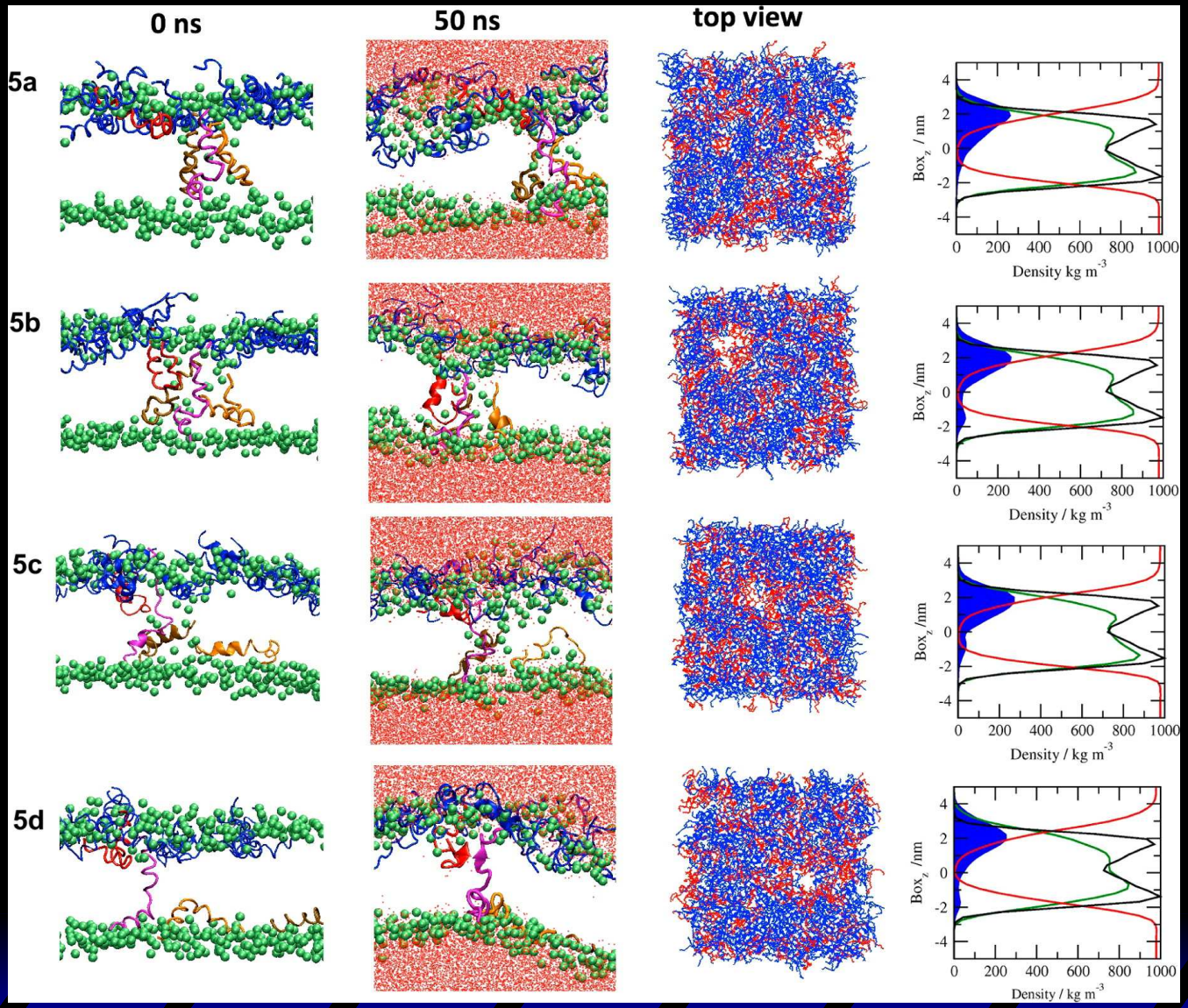
$\Longrightarrow$ no melittin reorientation

melittin/lipid=1/21: transient pore



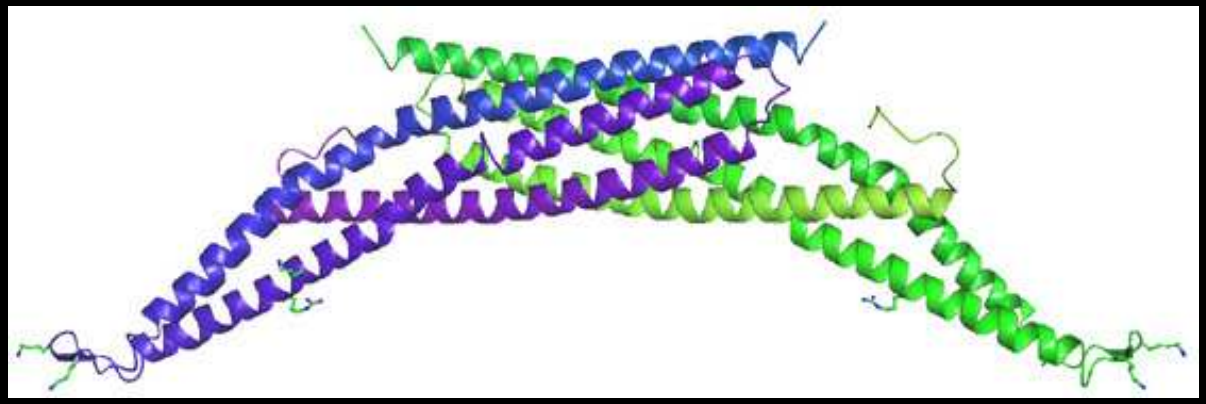
K.P.Santo et al., JPCB 117 (2013) 5031

continuation by atomistic simulations



K.P. Santo et al., JPCB **117** (2013), 5031.

- BAR (bin/amphiphysin/rvs)



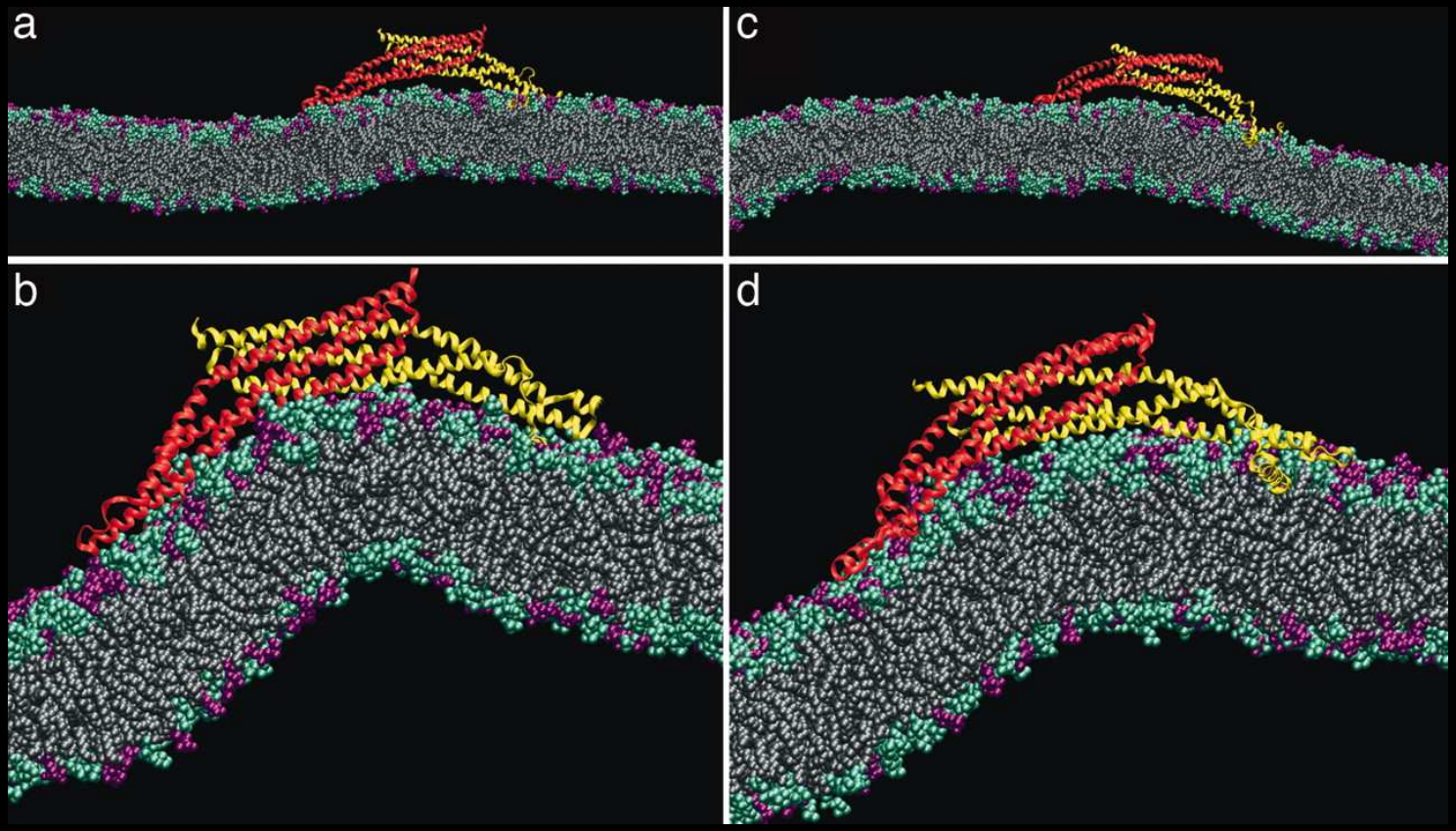
http://www.endocytosis.org/BARdomains/BARs.html

banana shaped (concave) α-helical dimer

induces/recognizes membrane curvature

N-BAR domain: amphiphilic N-terminals

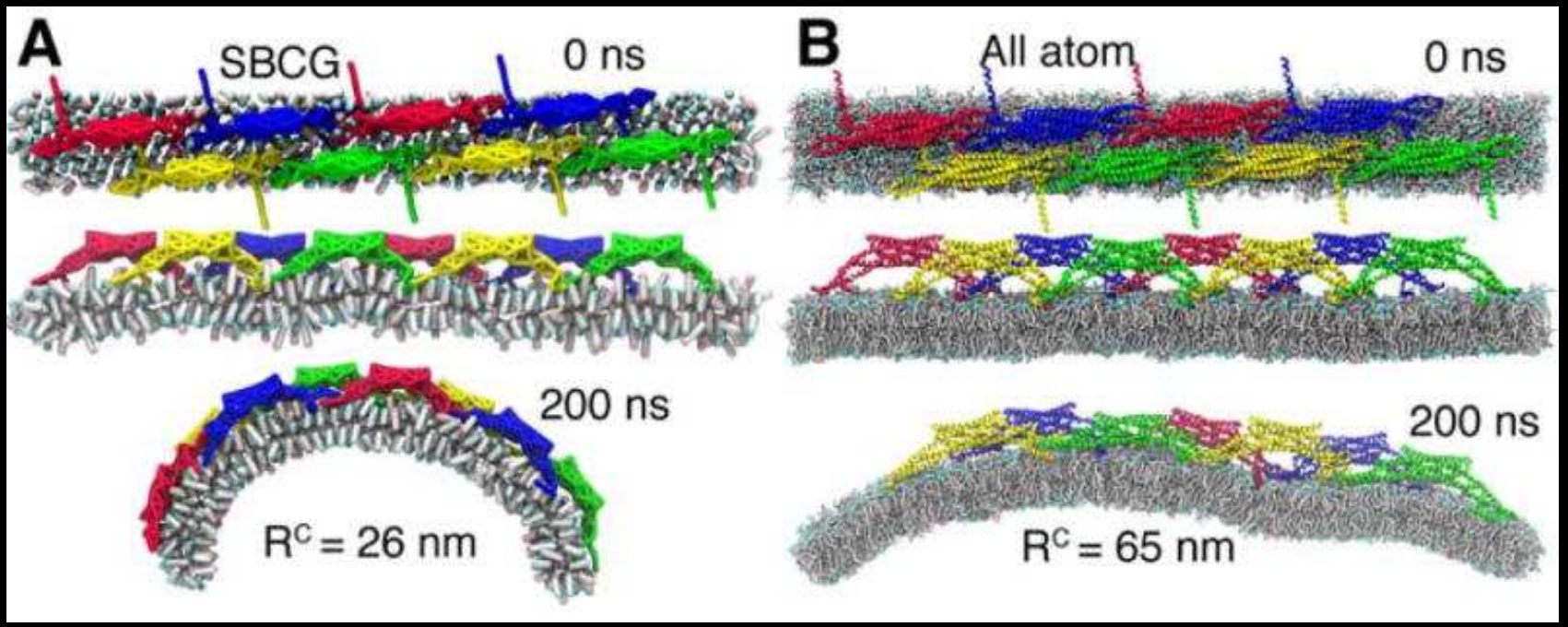
Atomistic simulations at DOPC/DOPS (7/3)



P.D. Blood, G.A. Voth, PNAS **103** (2006), 15068.

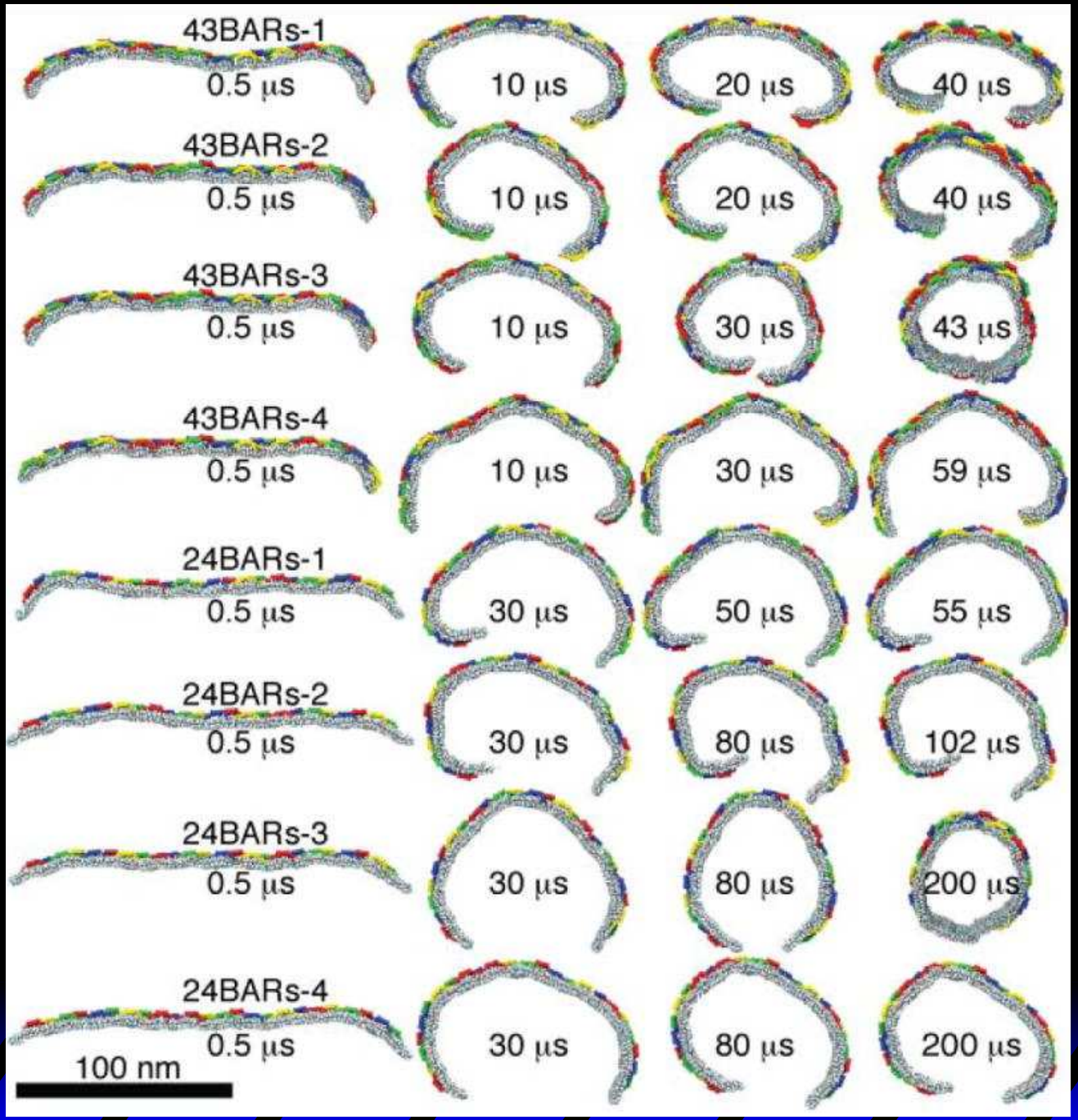
snapshots at 10 and 27 ns

Shape-based CG



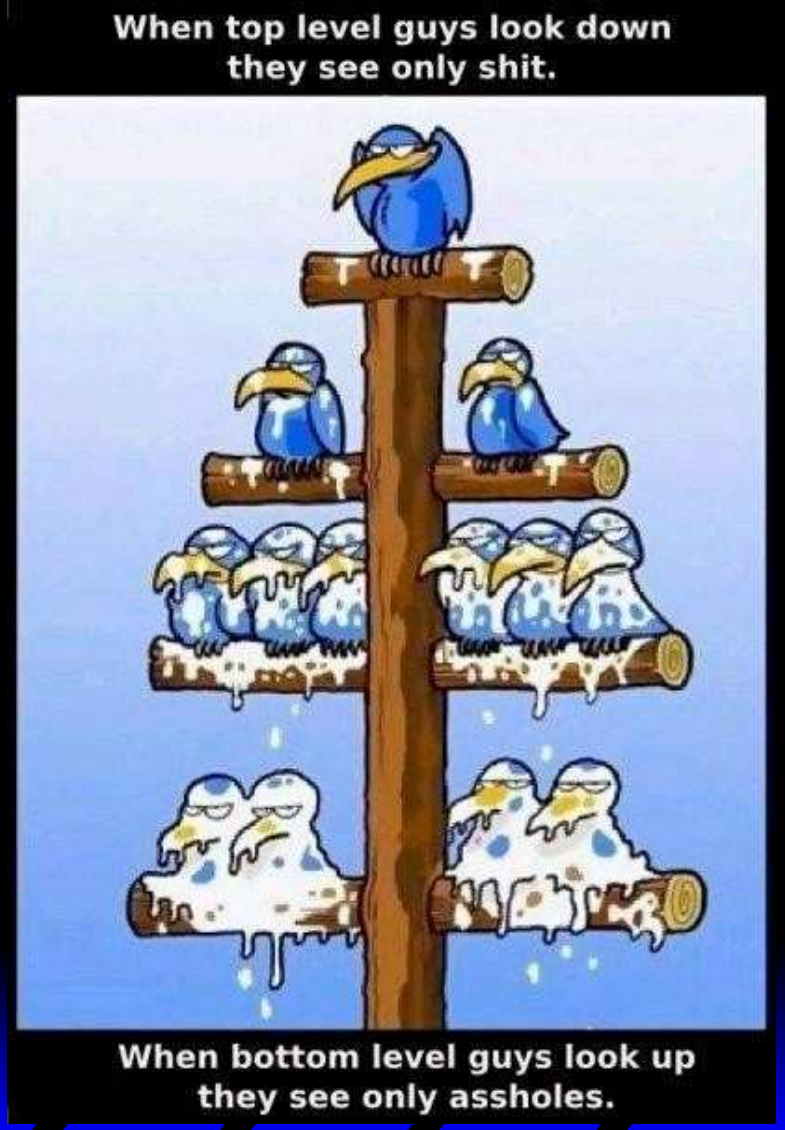
Y. Yin et al., Structure **17** (2009), 882.

$\Longrightarrow$ tubulation


43BARs-1
0.5 μs
10 μs
20 μs
40 μs
43BARs-2
0.5 μs
10 μs
20 μs
40 μs
43BARs-3
0.5 μs
10 μs
30 μs
43 μs
43BARs-4
0.5 μs
10 μs
30 μs
59 μs
24BARs-1
0.5 μs
30 μs
50 μs
55 μs
24BARs-2
0.5 μs
30 μs
80 μs
102 μs
24BARs-3
0.5 μs
30 μs
80 μs
200 μs
24BARs-4
0.5 μs
30 μs
80 μs
200 μs
100 nm

Jumping between Scales

- bottom-up and top-down



- MARTINI: bottom-up

  Rules: bead at the center of mass (C_α for proteins)

  flowchart at http://www.cgmartini.nl/cgmartini/index.php/flowchartfile indicating utilities for the transformation

  pdb file

  Proteins: sequence and secondary structure

  $\Longrightarrow$ that was easy

- MARTINI: top-down (the old version)

  A.J. Rzepiela et al., J. Comput. Chem. **31** (2010), 1333.

  Defining randomly AA positions around the CG bead ($r_{CG} = 0.3\,\text{nm}$)

  Restraining the AA to the corresponding CG bead

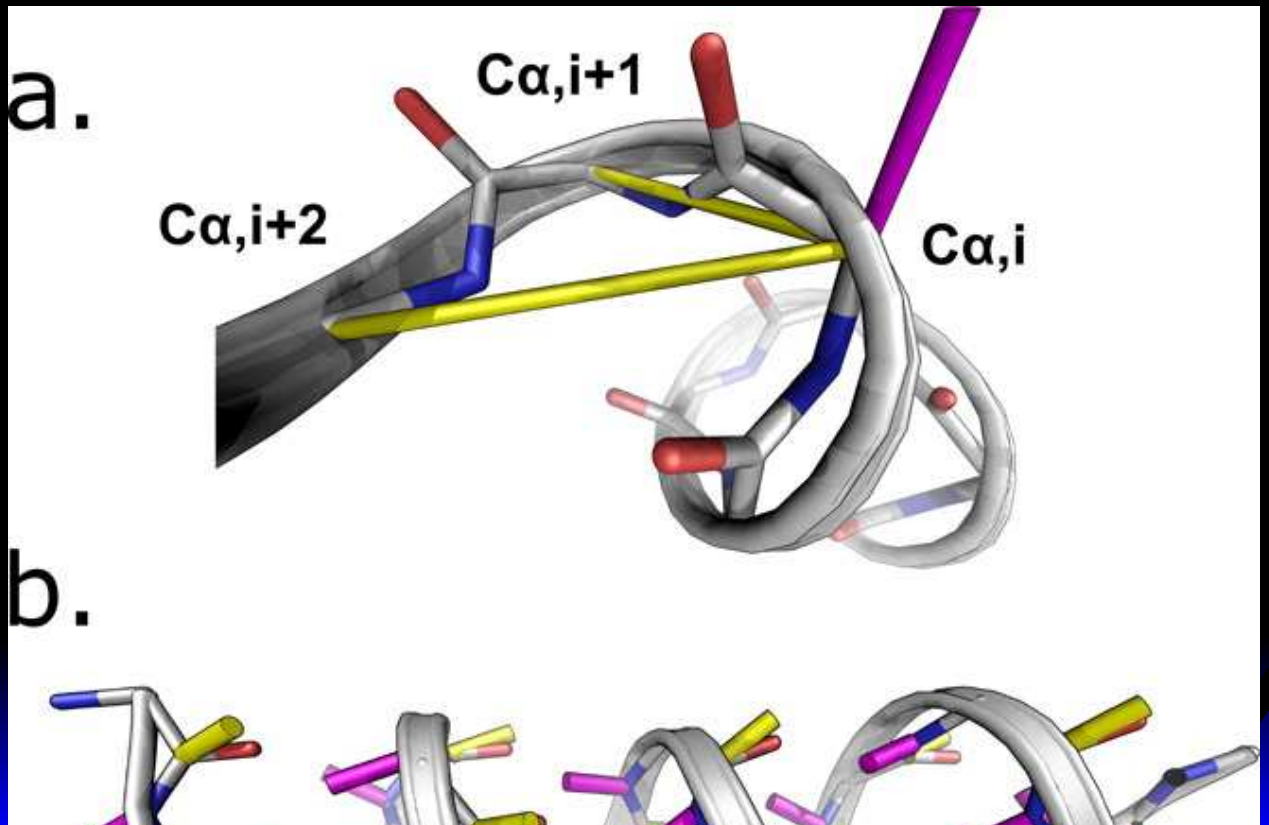
$$U_{restr} = \sum_{i=1}^{n} \frac{k}{2} \left(\mathbf{r}_i^{CG} - \mathbf{r}_i^{AA,com} \right)^2$$

  Performing simulated anealing starting at high temperature (1300 K) with a small cooling rate

- MARTINI: top-down (the new version)

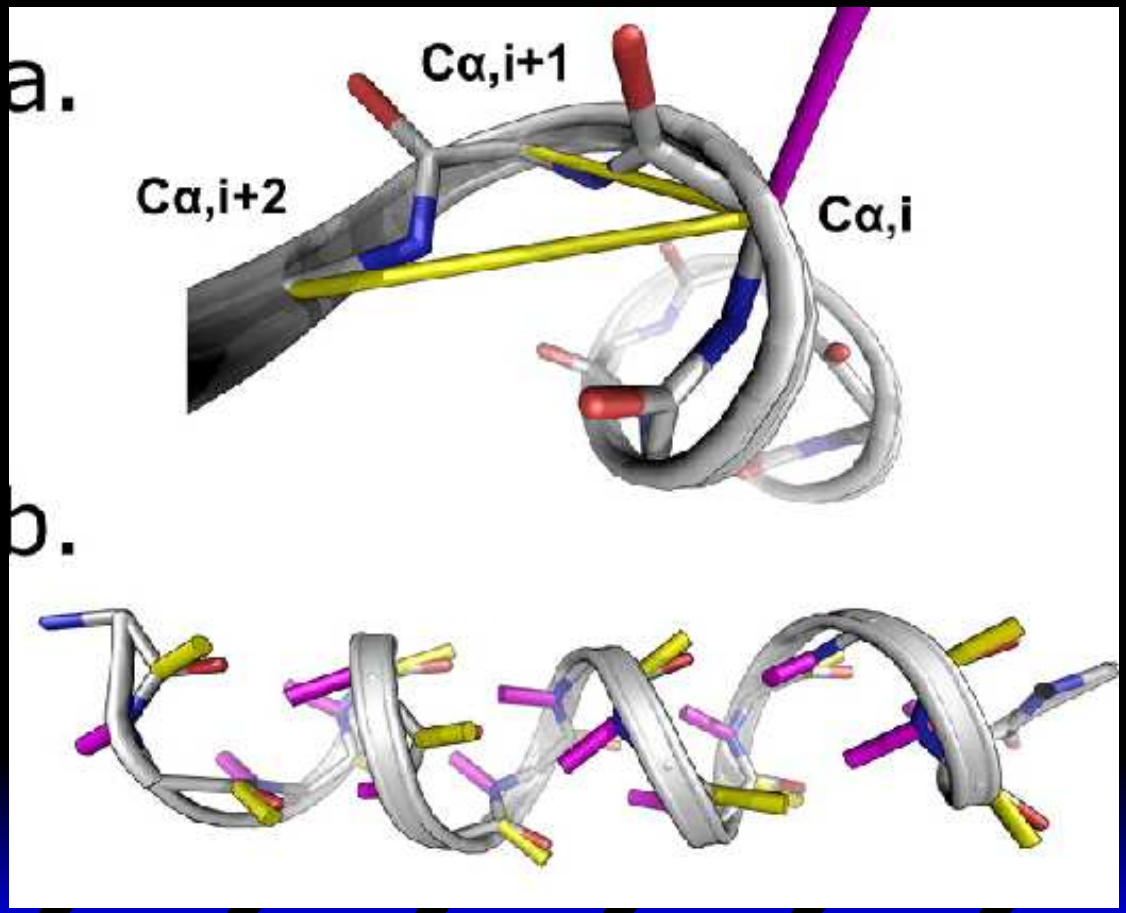
  T.A. Wassenaar et al., JCTC **10** (2014), 676.

  Observation: vector cross product of $C_{\alpha,i+1}$-$C_{\alpha,i}$ and $C_{\alpha,i+2}$-$C_{\alpha,i}$ approximates the direction of C=O at $C_{\alpha,i}$

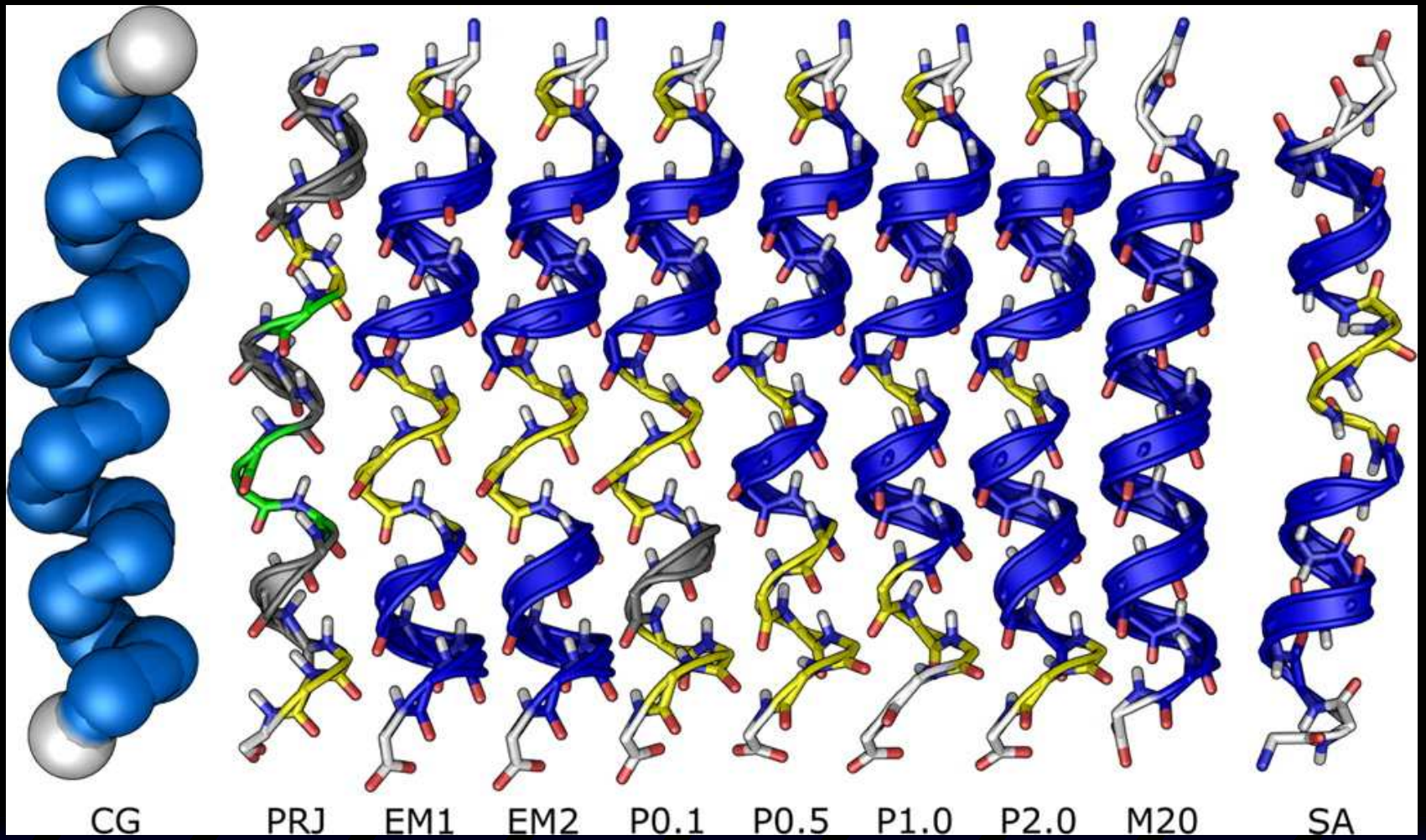


C=O at 1/3 of the distance from $C_{\alpha,i}$ to $C_{\alpha,i+1}$

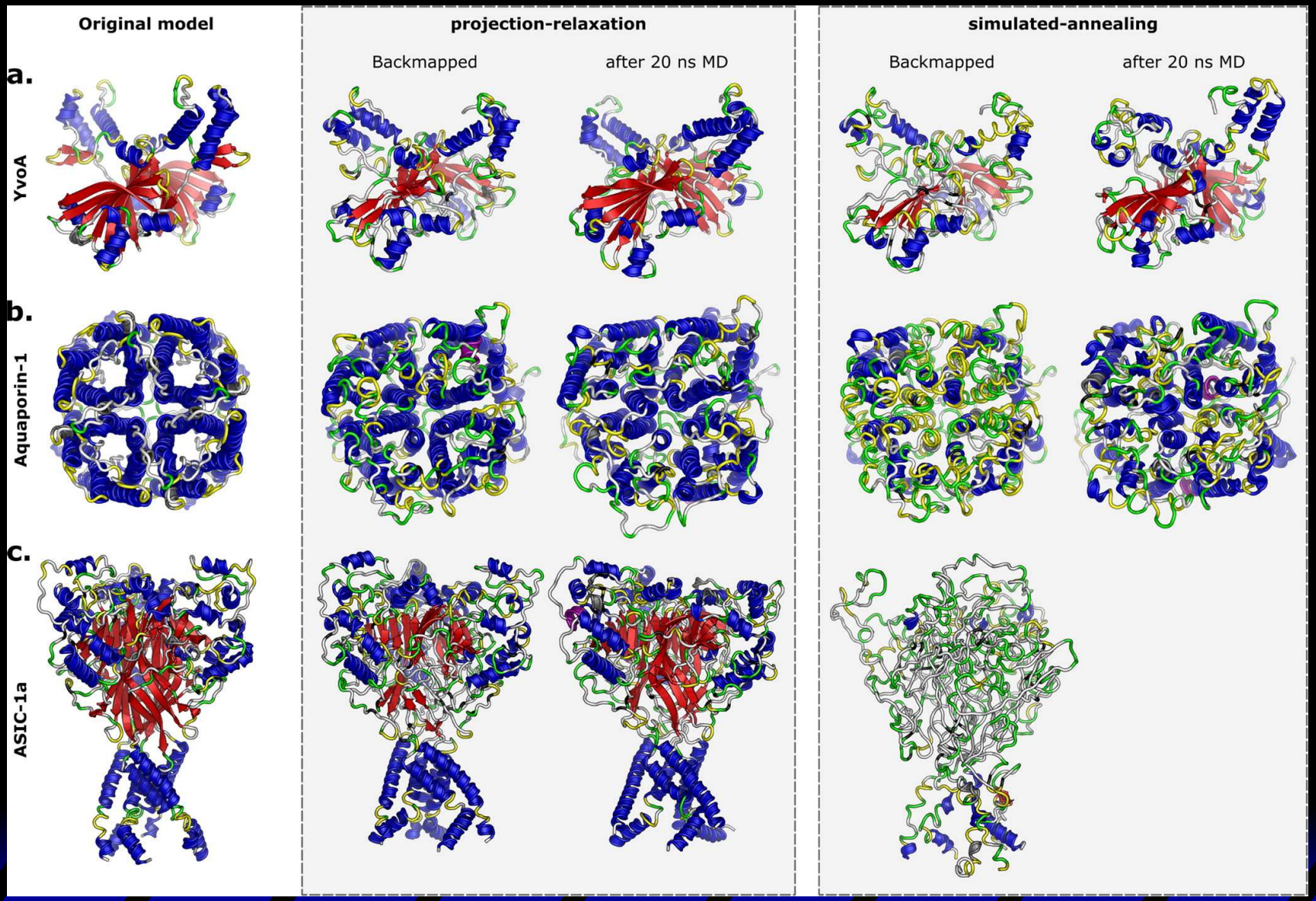
N-H at 2/3 of this distance in opposite direction



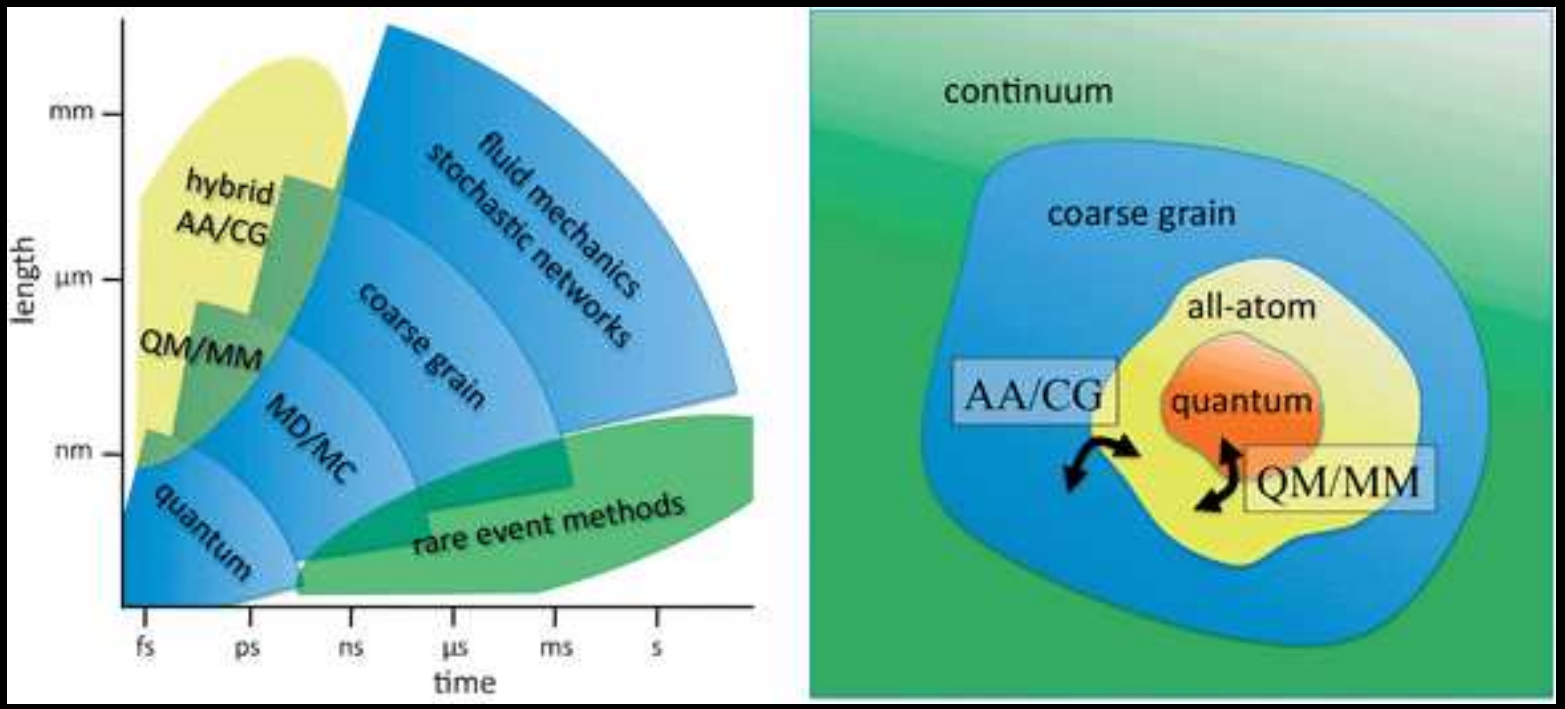
Example: WALP20 (helical)



Other examples



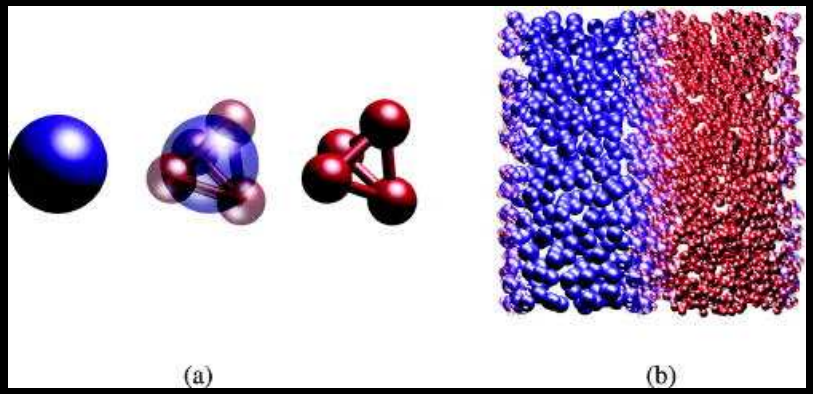
- Real Multiscaling

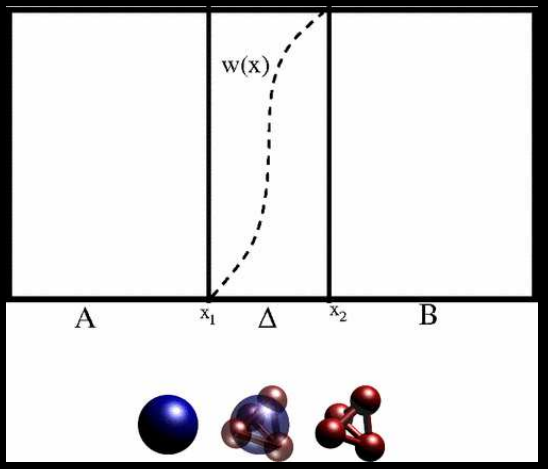


S.O. Nielsen et al., PCCP **12** (2010), 12401.

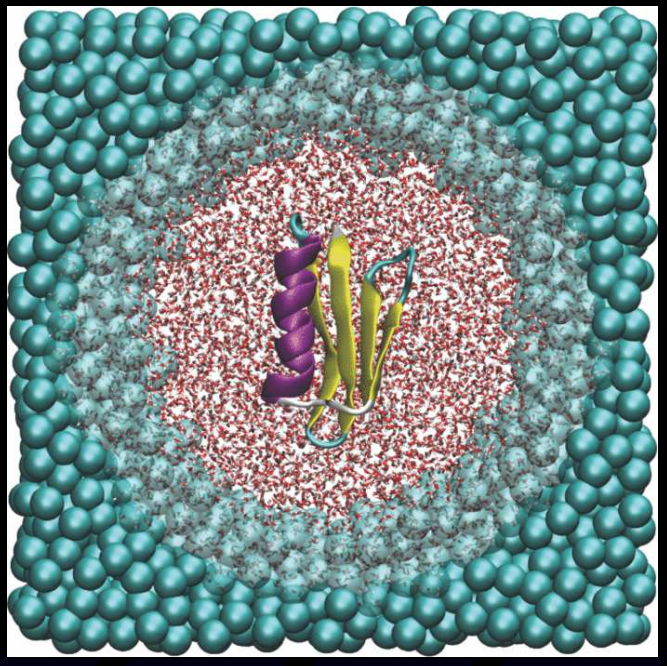
need for multiple time step algorithms

- Adaptive Resolution





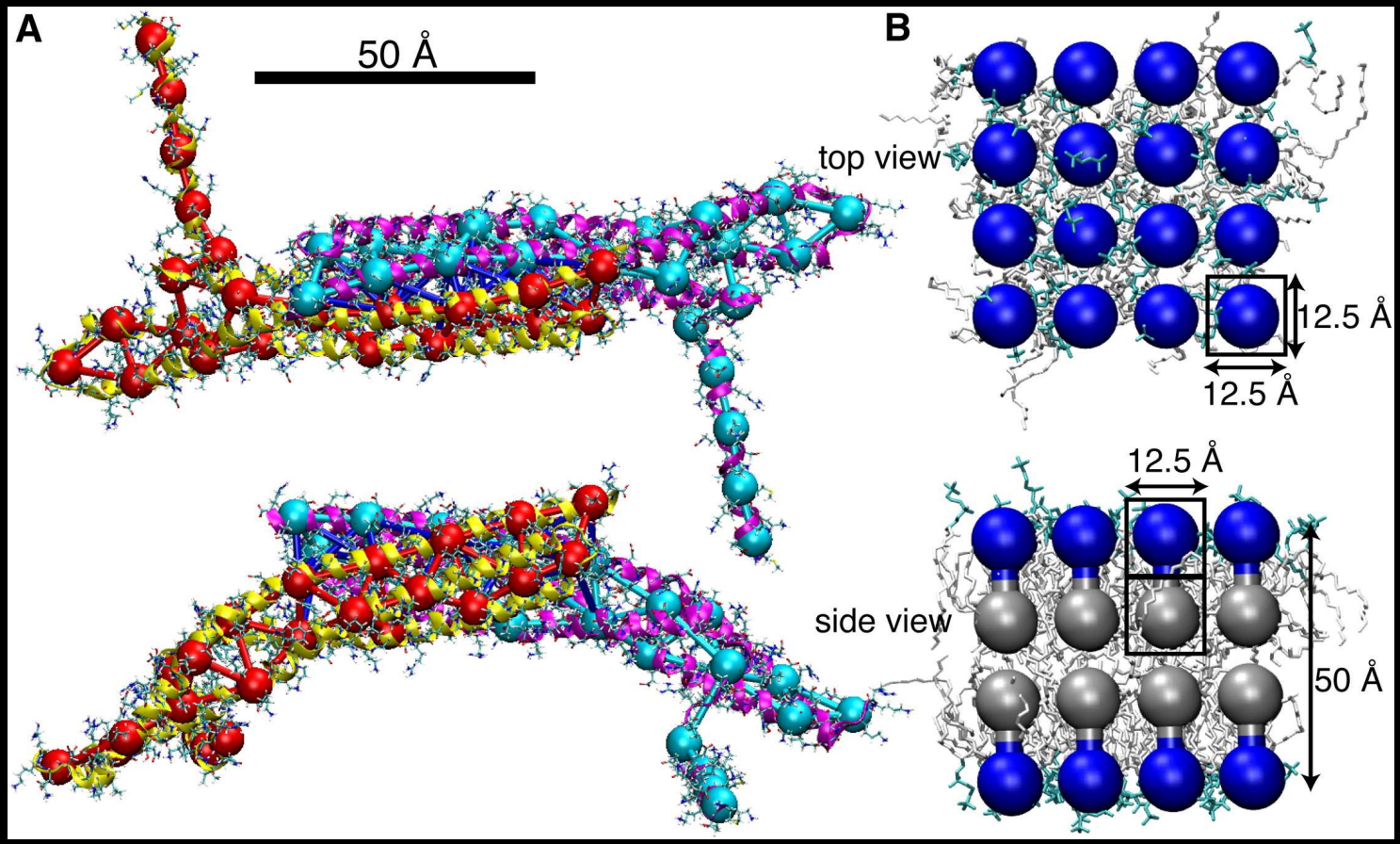
M. Praprotnik et al., JCP **123** (2005), 224106.



G-protein: J. Zavadlav et al., JCP **140** (2014), 054114.

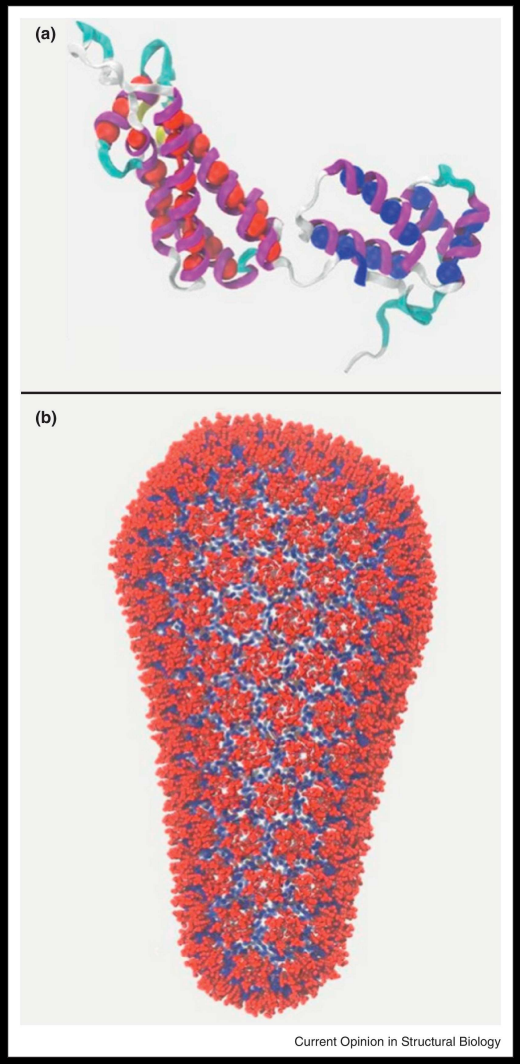
- Ultra-CG (shape-based CG)

  BAR domain



Y. Yin et al., Structure **17** (2009), 882.

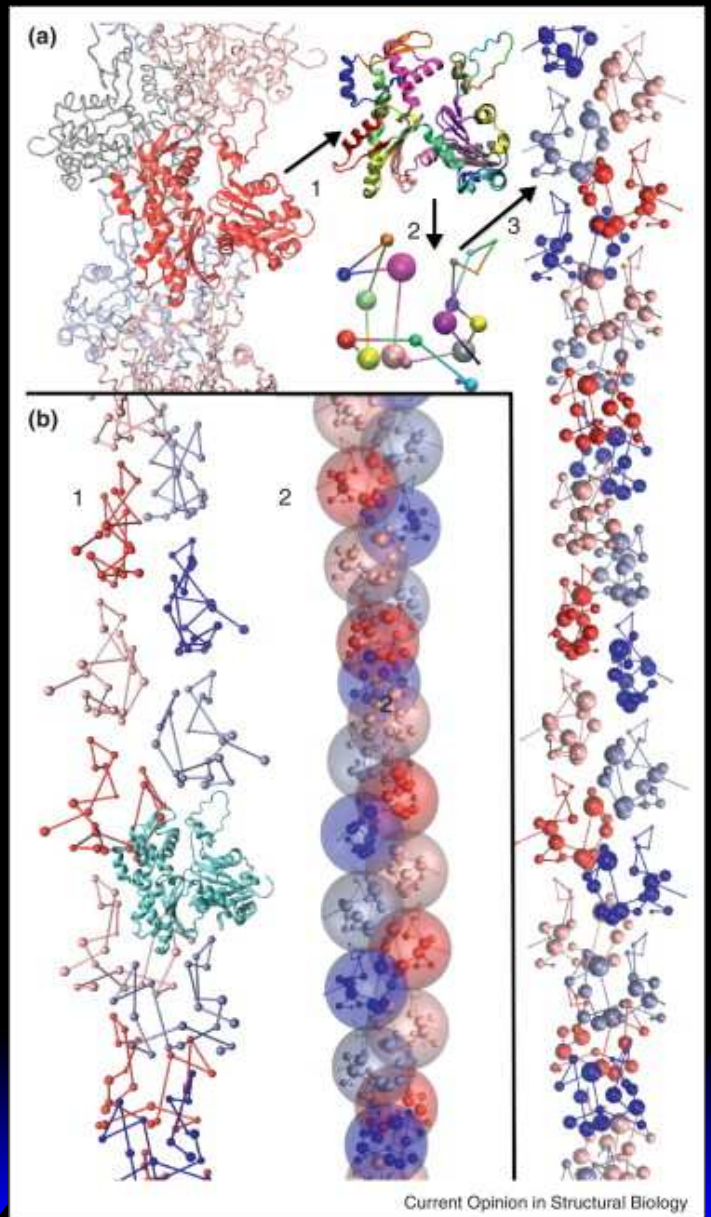
HIV capsid protein: shape based CG



M.G. Saunders, G.A. Voth, Curr. Opin. Struct. Biol., **22** (2012), 144.

- Bridging: Actin filament (muscle)

M.G. Saunders, G.A. Voth, Curr. Opin. Struct. Biol., **22** (2012), 144.



"Everything should be made as simple as possible, but not simpler."

Albert Einstein