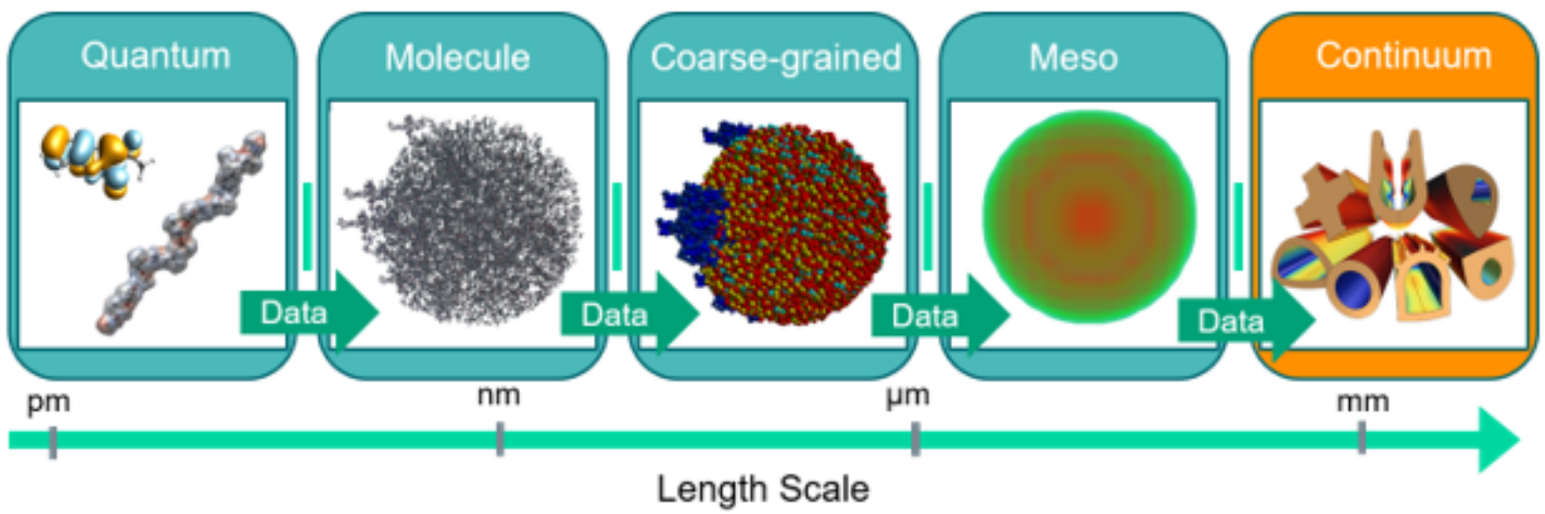
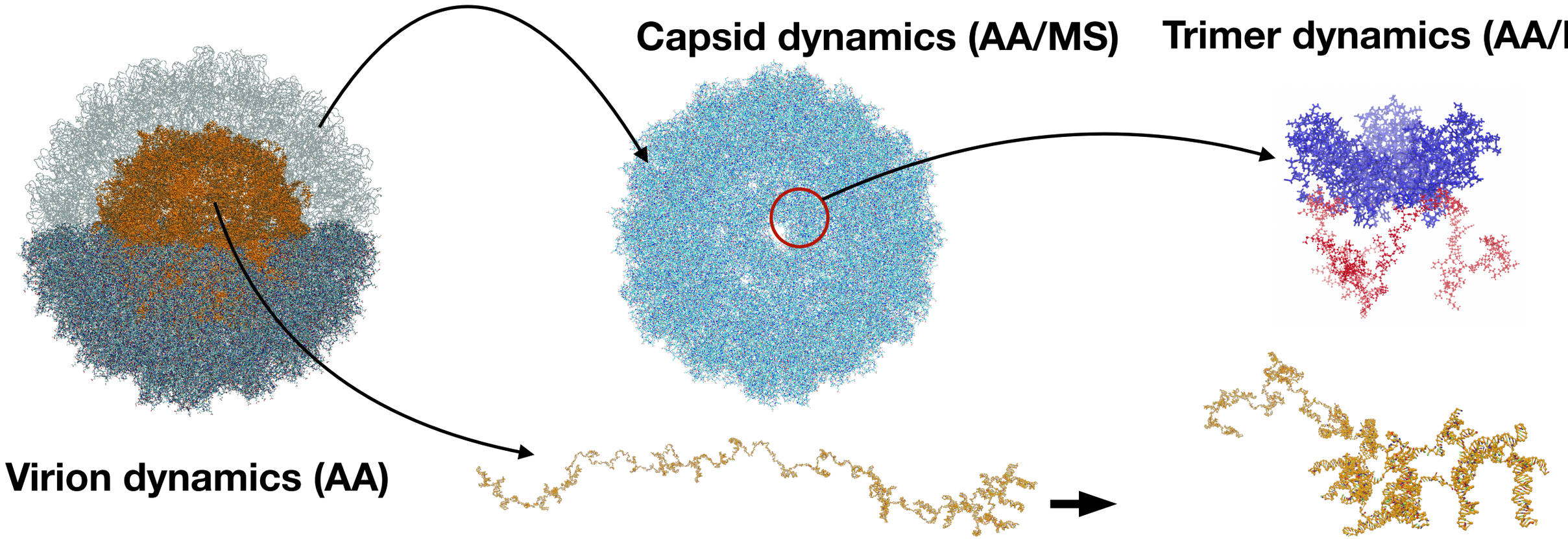
Computational 3D structure Characterization of CCMV RNA2 + Fragment =

Giovanni Mattiotti & Manuel Micheloni

ECT Workshop - Structure and topology of RNA in living systems*
Trento, 2023



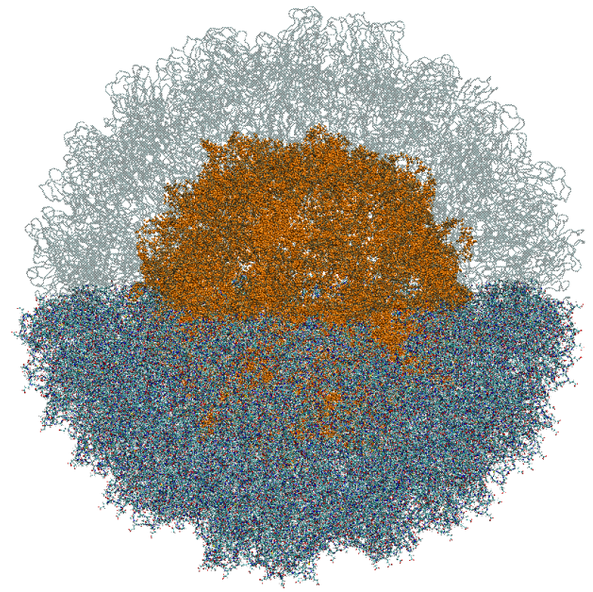
Overview of the project



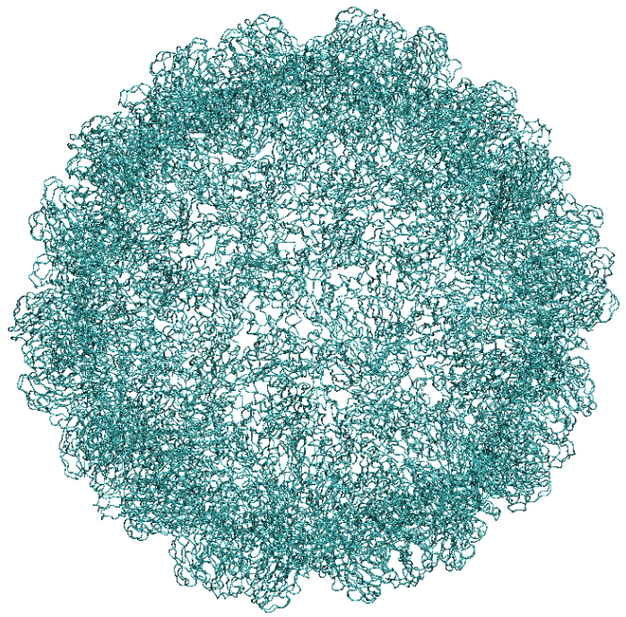
- In silico study of structural dynamics of a whole virus particle (virion)
- Limit-testing of state-of-the-art all-atom (AA) and coarse-grained (CG) force fields to simulate mixed systems (proteins and nucleic acids)

Chlorotic Cowpea Mottle Virus

- Small icosahedral plant virus
- Proteic capsid and (+)ssRNA genome
- Capsid has $5*10^5$ atoms, RNA $8*10^4$
  ⇒ suitable for MD simulations



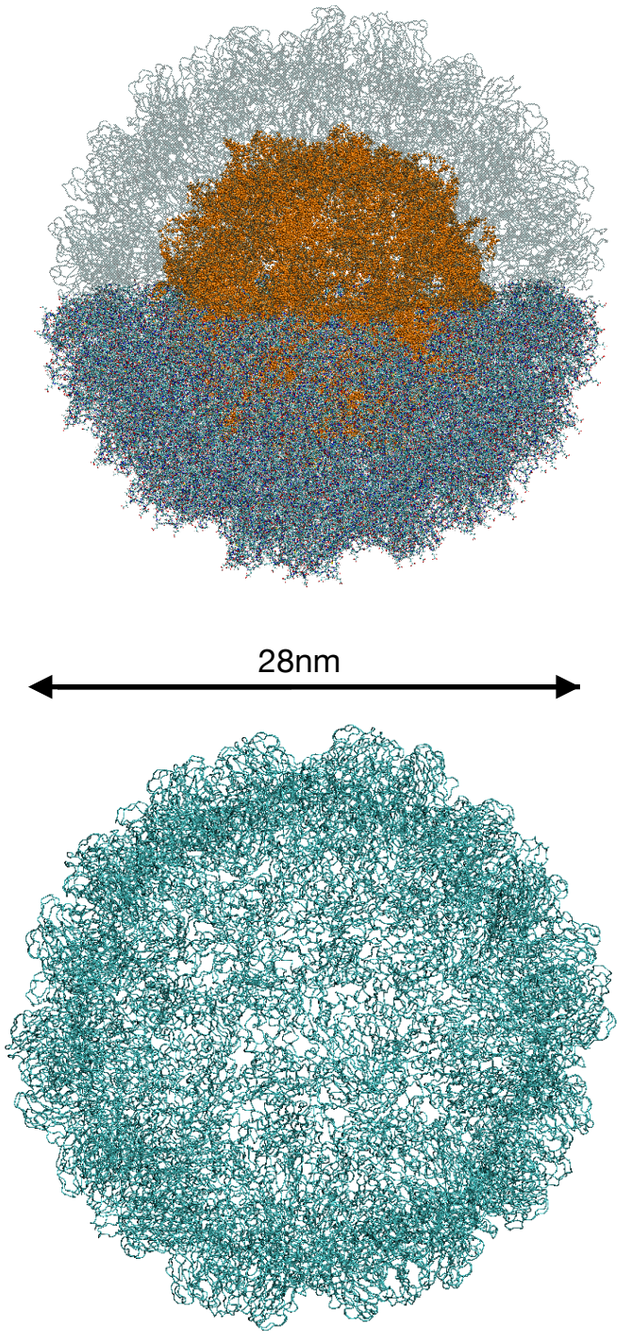
- First icosahedral virus reconstructed in vitro from its constituents [1] ⇒ studied a lot as benchmark virus
- Interesting nanotechnological applications



[1] *Garmann RF (2014), PhD Thesis*

Chlorotic Cowpea Mottle Virus

- PDB structure [2] of the capsid without tails (that are important to study RNA-capsid interactions)
- Multipartite virus: 4 different fragments of genome found inside it (RNA1, RNA2, RN3+RNA4)
- Sequences known [3], structures are an open problem ⇒ addressable via molecular dynamics simulations (?)

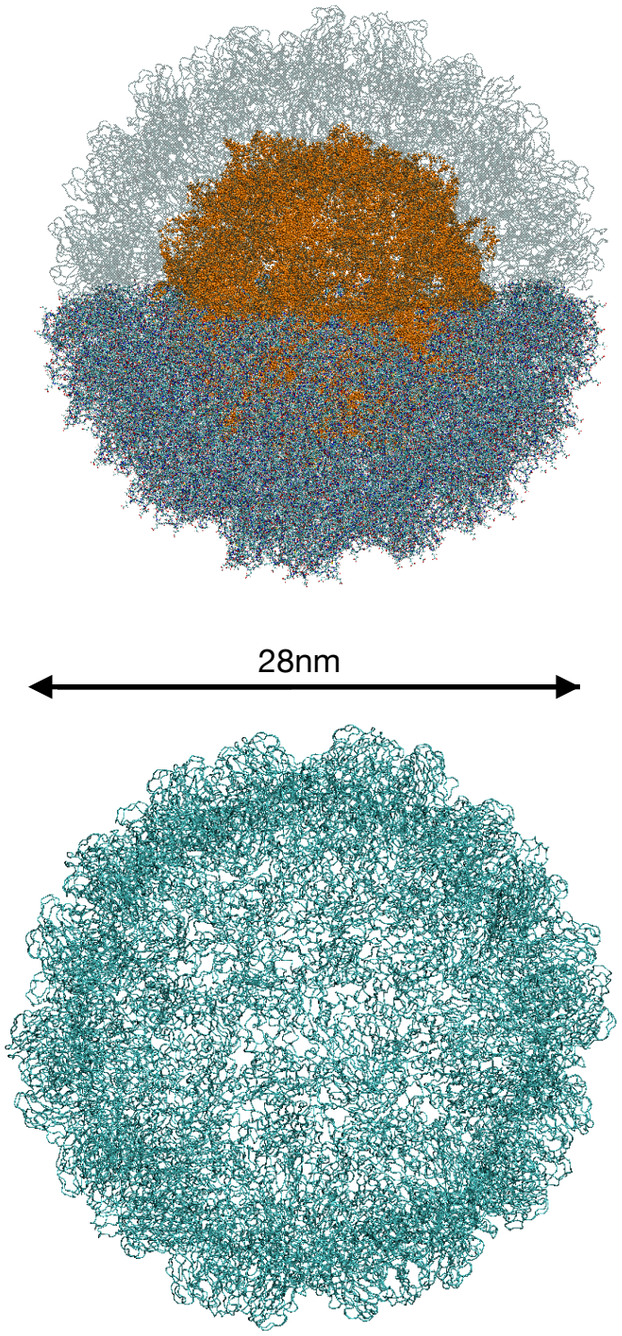


[2] *Speir et al., Structure (1995), 3(1), 63-78*

[3] *Allison et al., Virology (Sept. 1989), 172(1):321-30*

Chlorotic Cowpea Mottle Virus

- PDB structure [2] of the capsid without tails (that are important to study RNA-capsid interactions)
- Multipartite virus: 4 different fragments of genome found inside it (RNA1, RNA2, RN3+RNA4)
- Sequences known [3], structures are an open problem ⇒ addressable via molecular dynamics simulations (?)

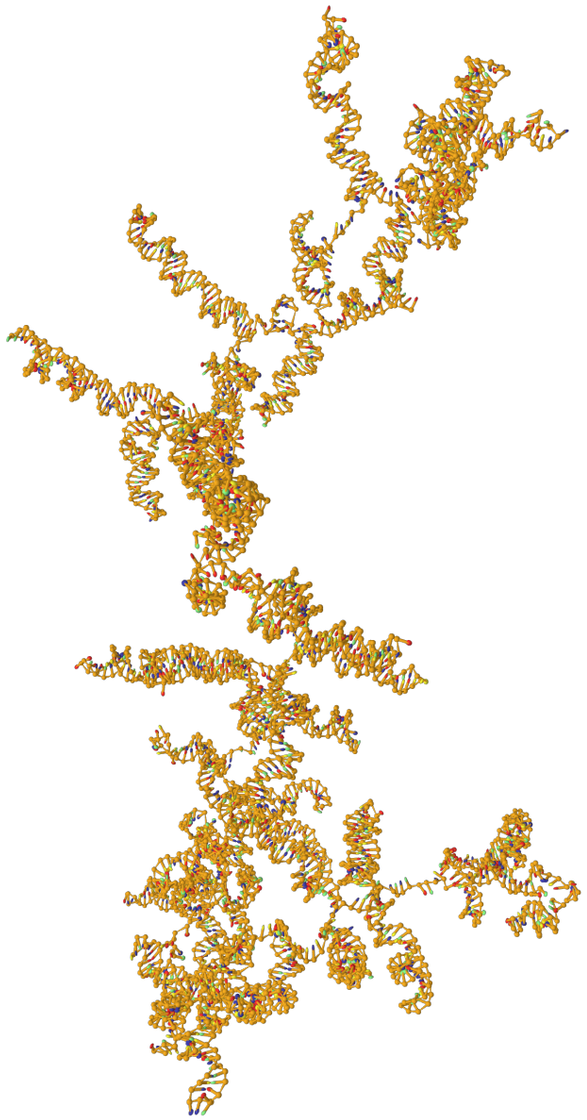


[2] *Speir et al., Structure (1995), 3(1), 63-78*

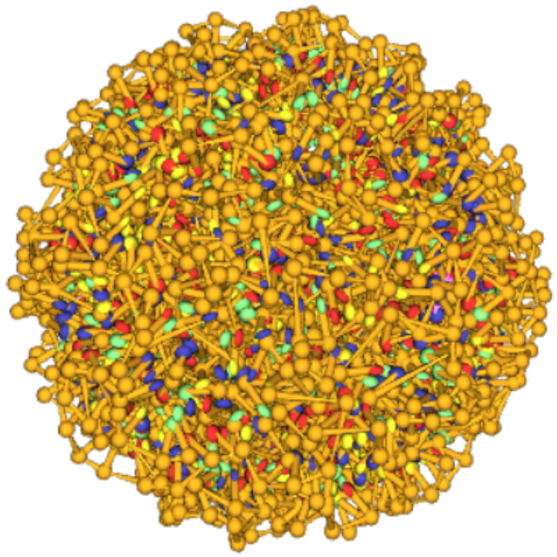
[3] *Allison et al., Virology (Sept. 1989), 172(1):321-30*

Three different environments

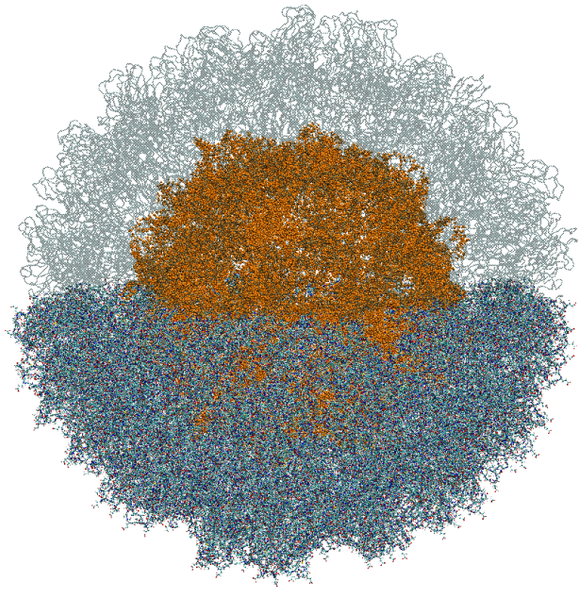
Free in (implicit) solution



Under spherical time-dependent constraint

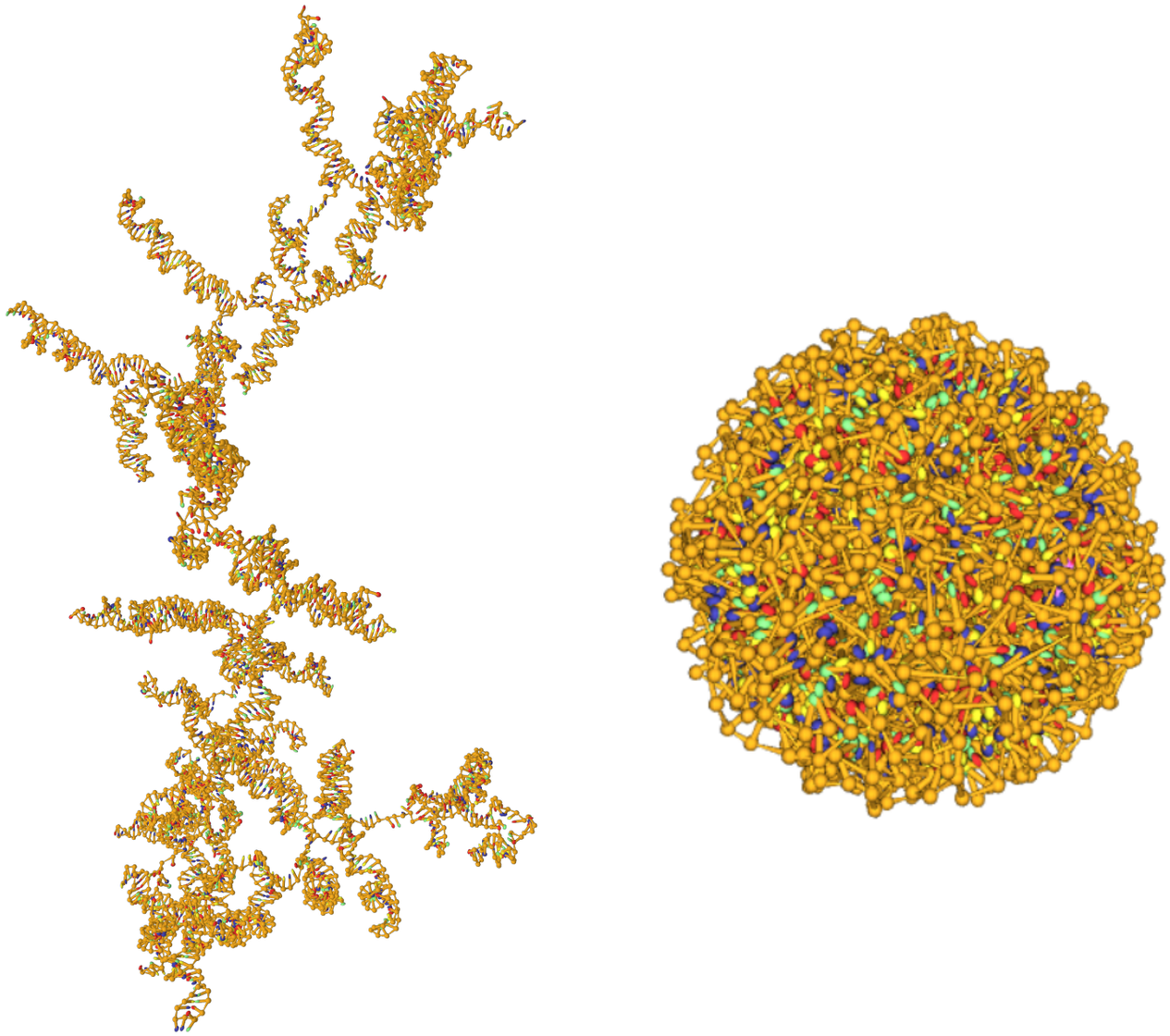


Packed into the high-resolution model of the capsid

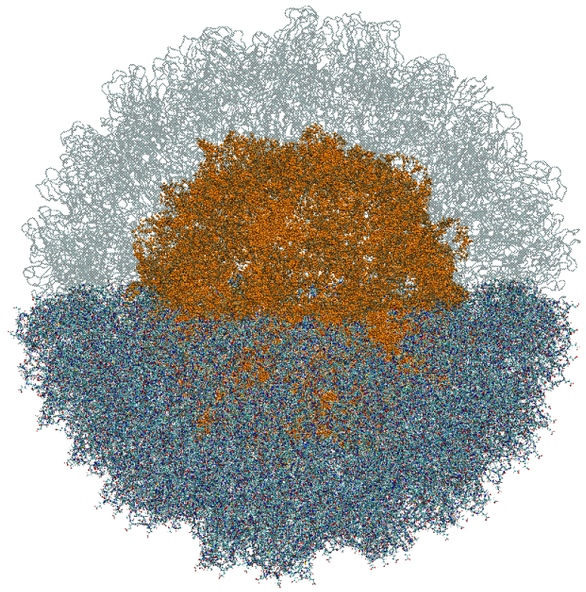


Three different environments

oxRNA [4] coarse-grained model

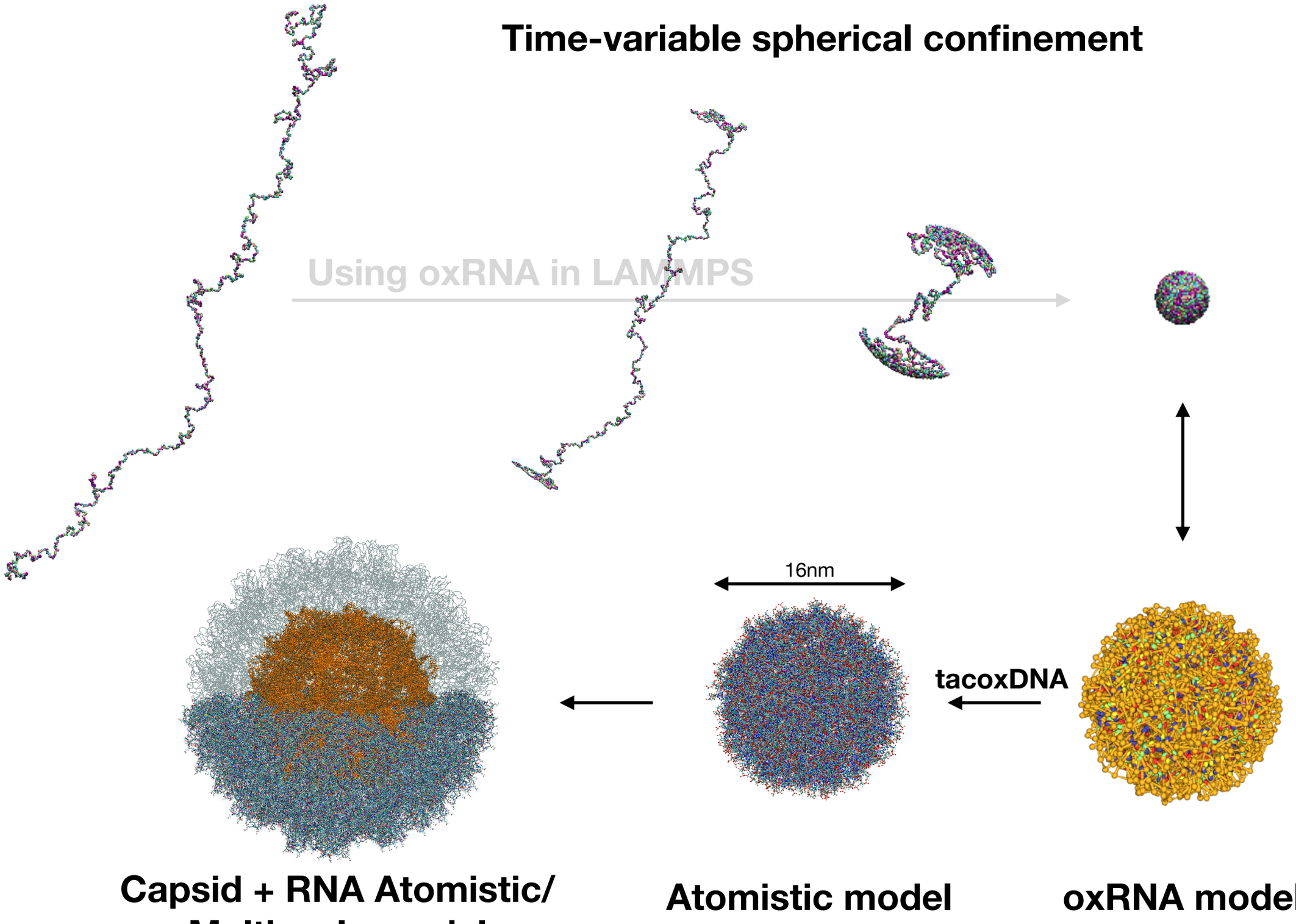


All-atom CHARMM36m force field in explicit solvent

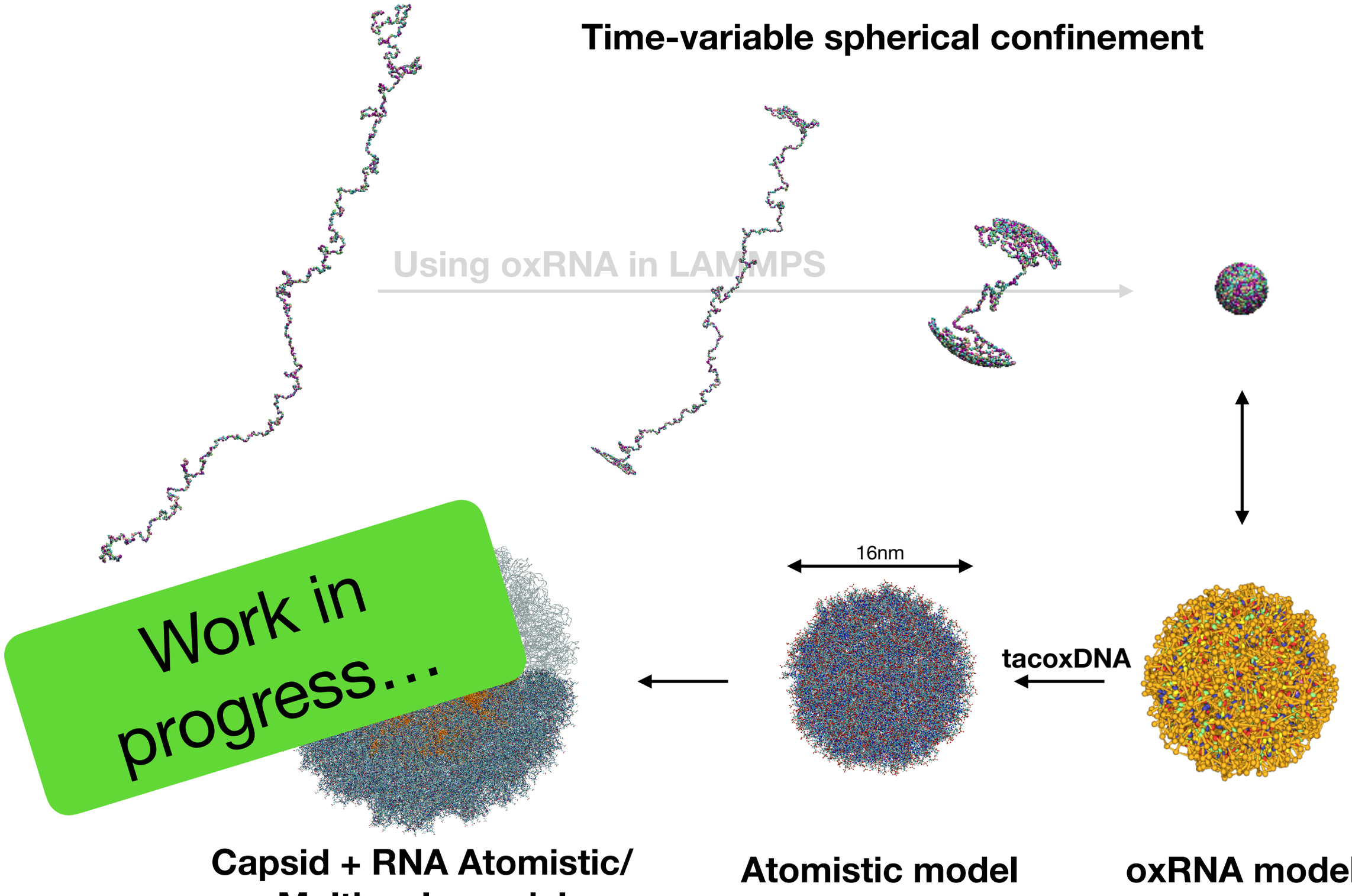


[4] *Matek et al., J. Chem. Phys. 143, 243122 (2015)*

RNA2 packing into the capsid



RNA2 packing into the capsid



oxRNA free filament setup

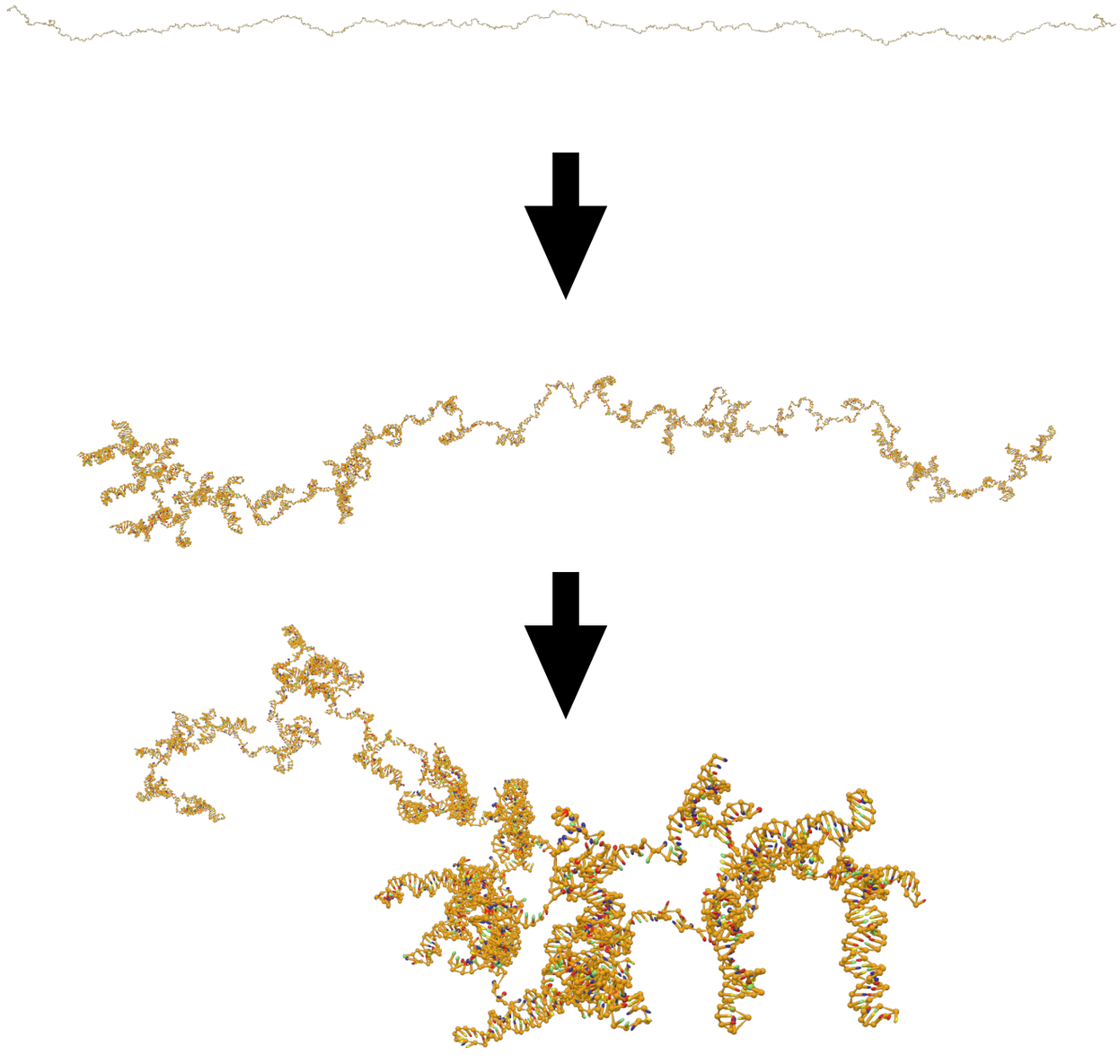
- oxDNA2 Force-Field ⇒ implicit solvent

  with Debye screening due to salt concentration

- Two different concentration chooses: 0.15M and 0.5M

- Experimentally [5] CP-RNA interaction is strong at 0.1M and absent at 1M

- Chain relaxation at T = 333K and then production run at T = 310K



[5] *J Mol Biol. 2014 March 06; 426(5): 1050-1060*

Internal energy and equilibrium

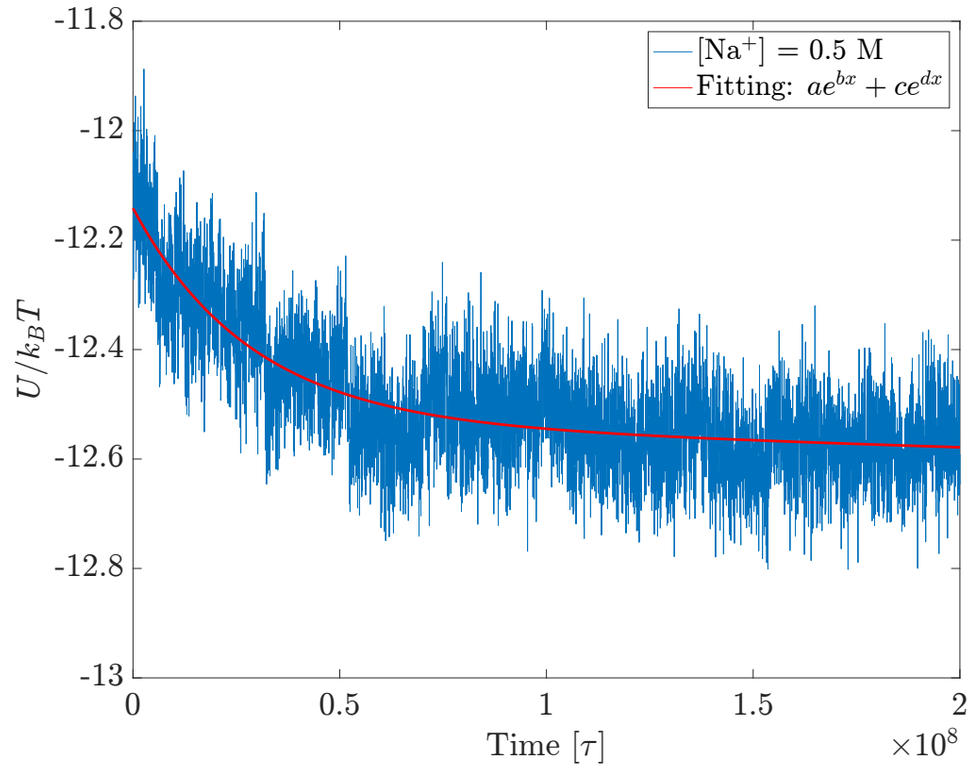
- Monitoring inter-nucleotide energy to assess equilibrium

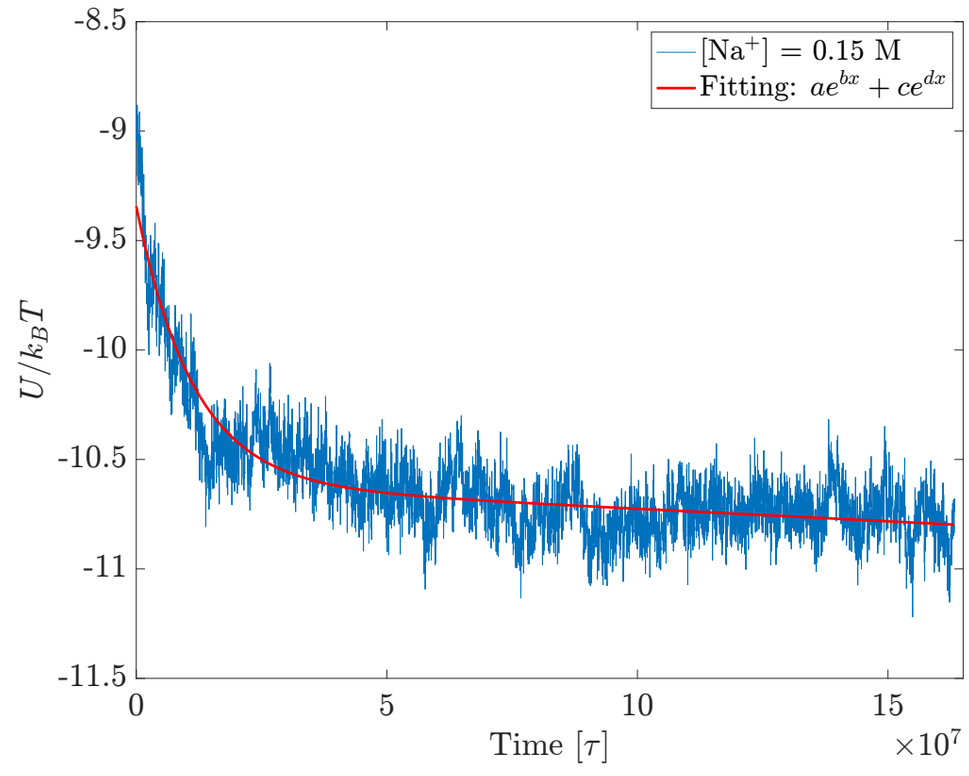
- Fit to a double-exponential to discard pre-equilibration trajectories for analysis

- At higher salt concentration the value is lower ⇒ more stability due to weaker screening effects

$$U_{\text{fit}}(t) = A \cdot e^{bt} + C \cdot e^{dt}$$

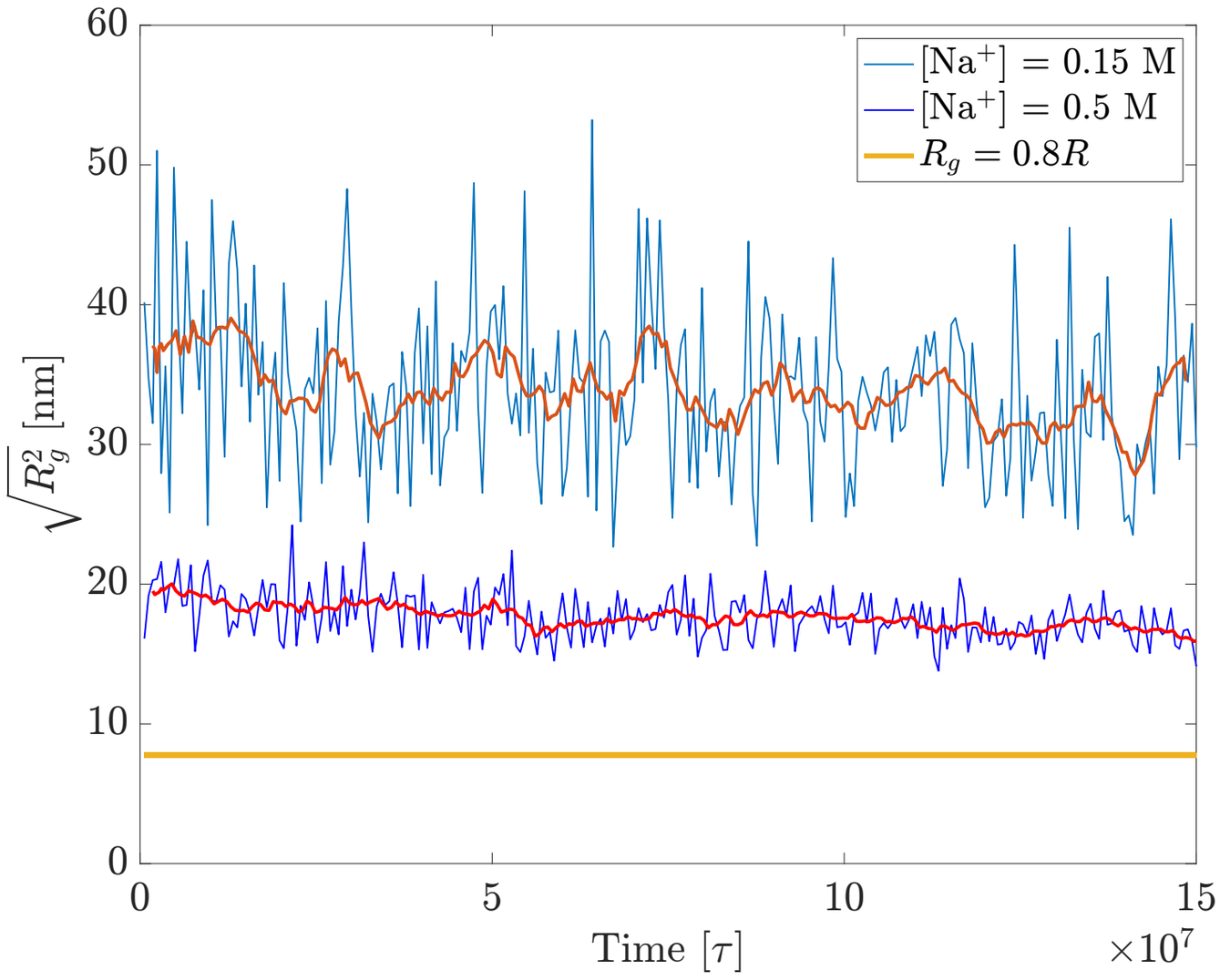
$$\Rightarrow t_{\text{eq}} \equiv 3 \cdot \left(-\frac{1}{d}\right)$$





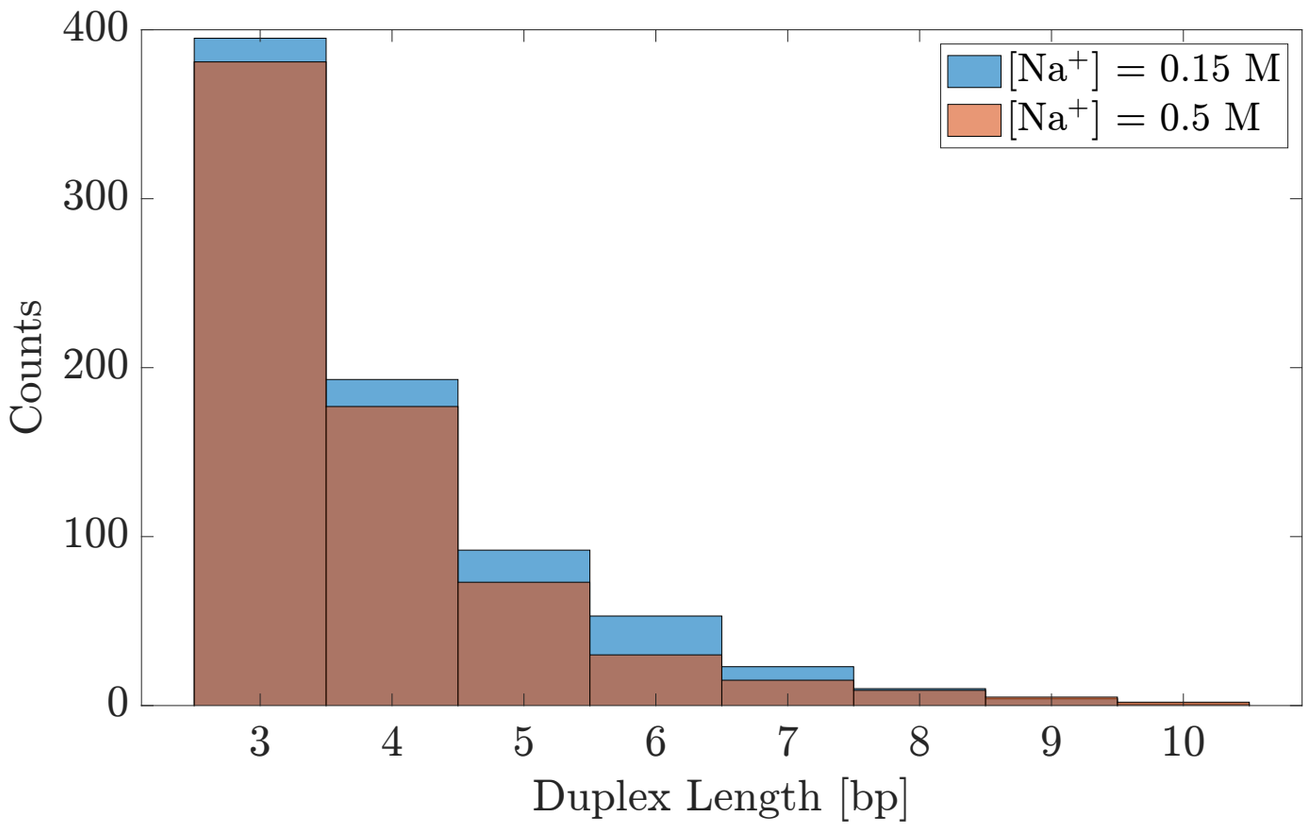
Radius of Gyration

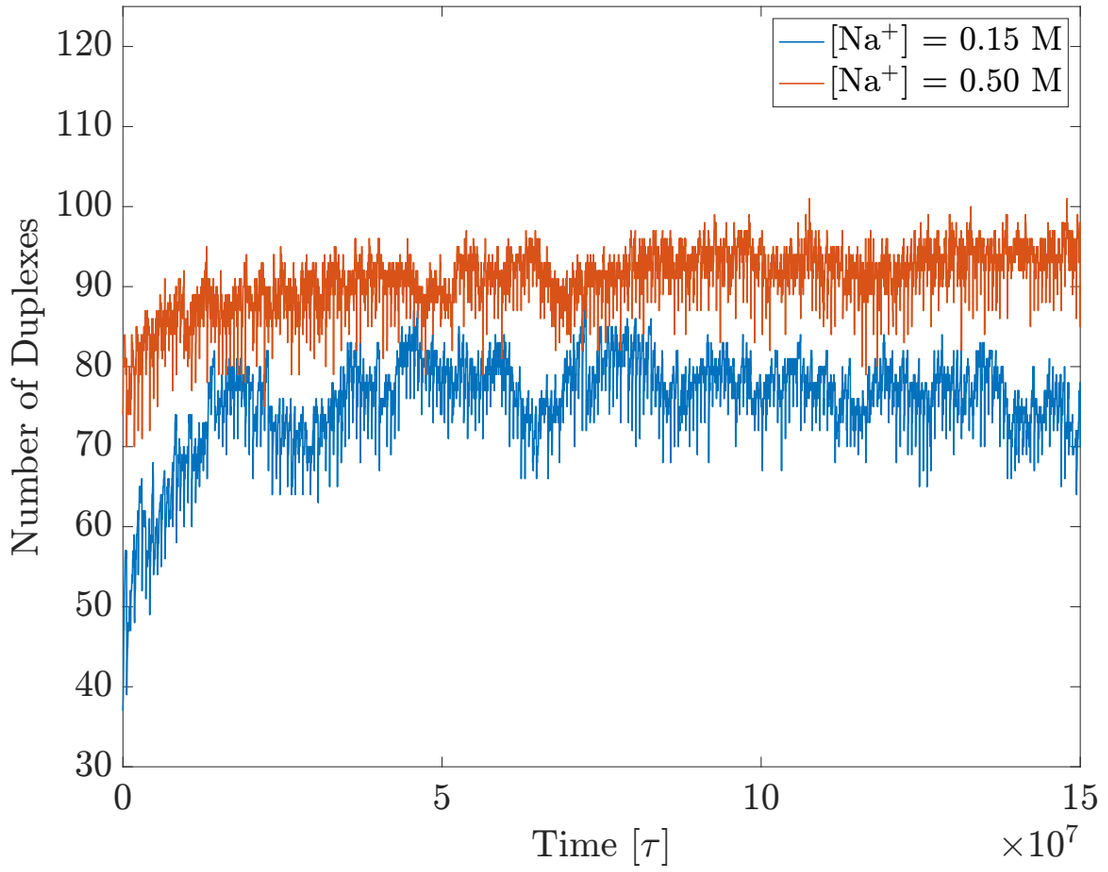
- Values stabilized in the equilibrated part
- Substantial difference (factor 2) between the 2 salt concentrations
- Far from being compatible with free space in capsid
- In vitro with 1M there is no CP-RNA interaction [5] ⇒ no self-assembly

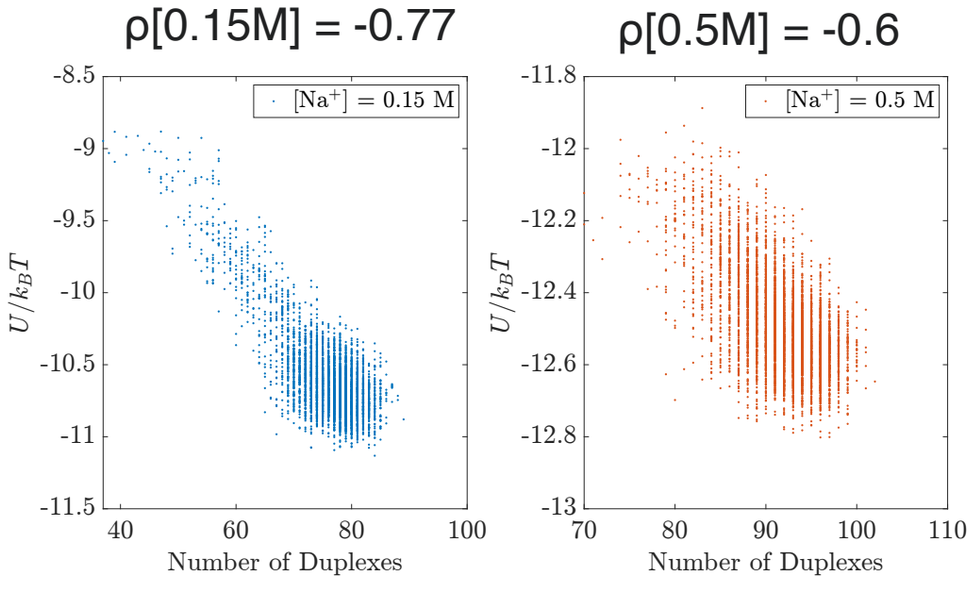


[5] *J Mol Biol. 2014 March 06; 426(5): 1050-1060*

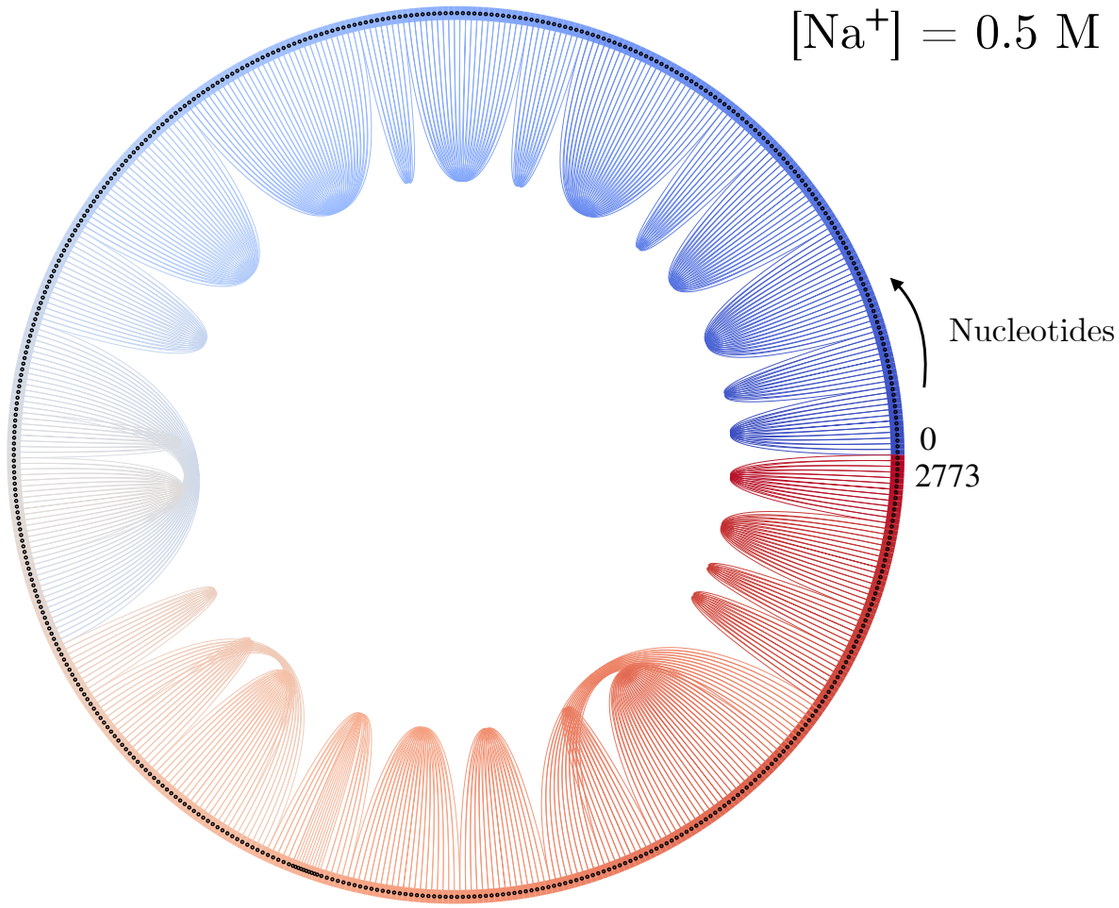
Duplexes in time

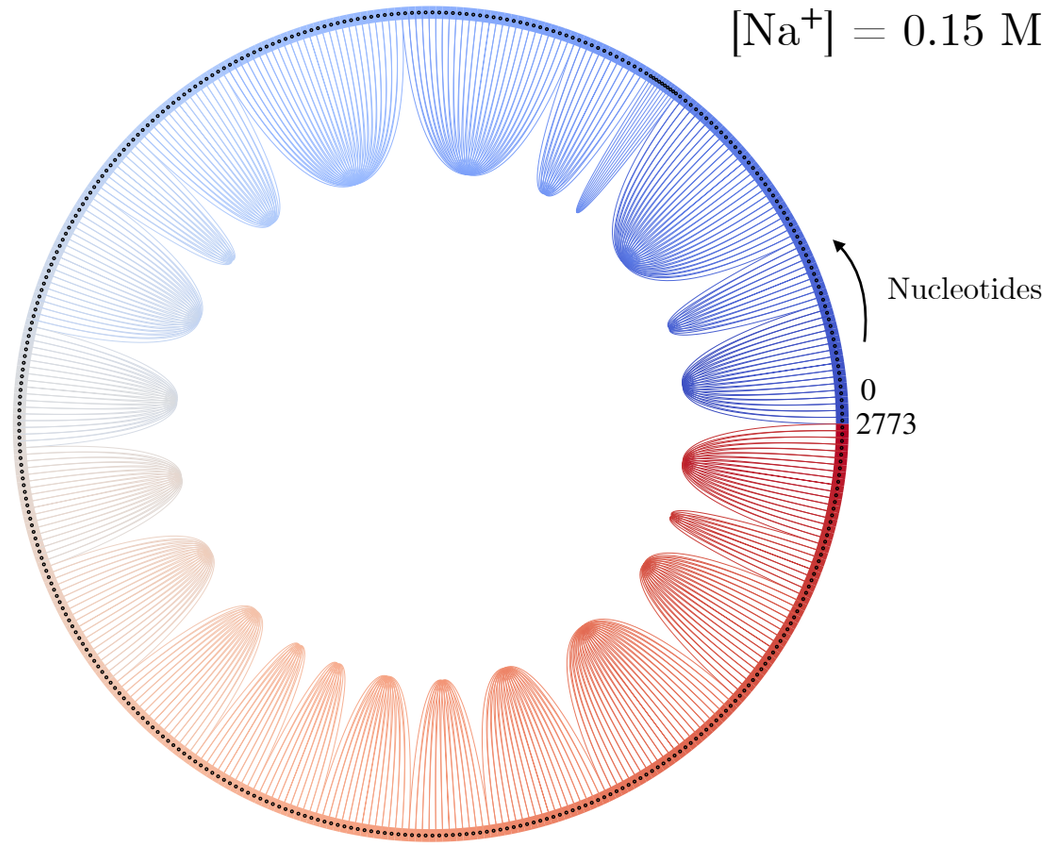




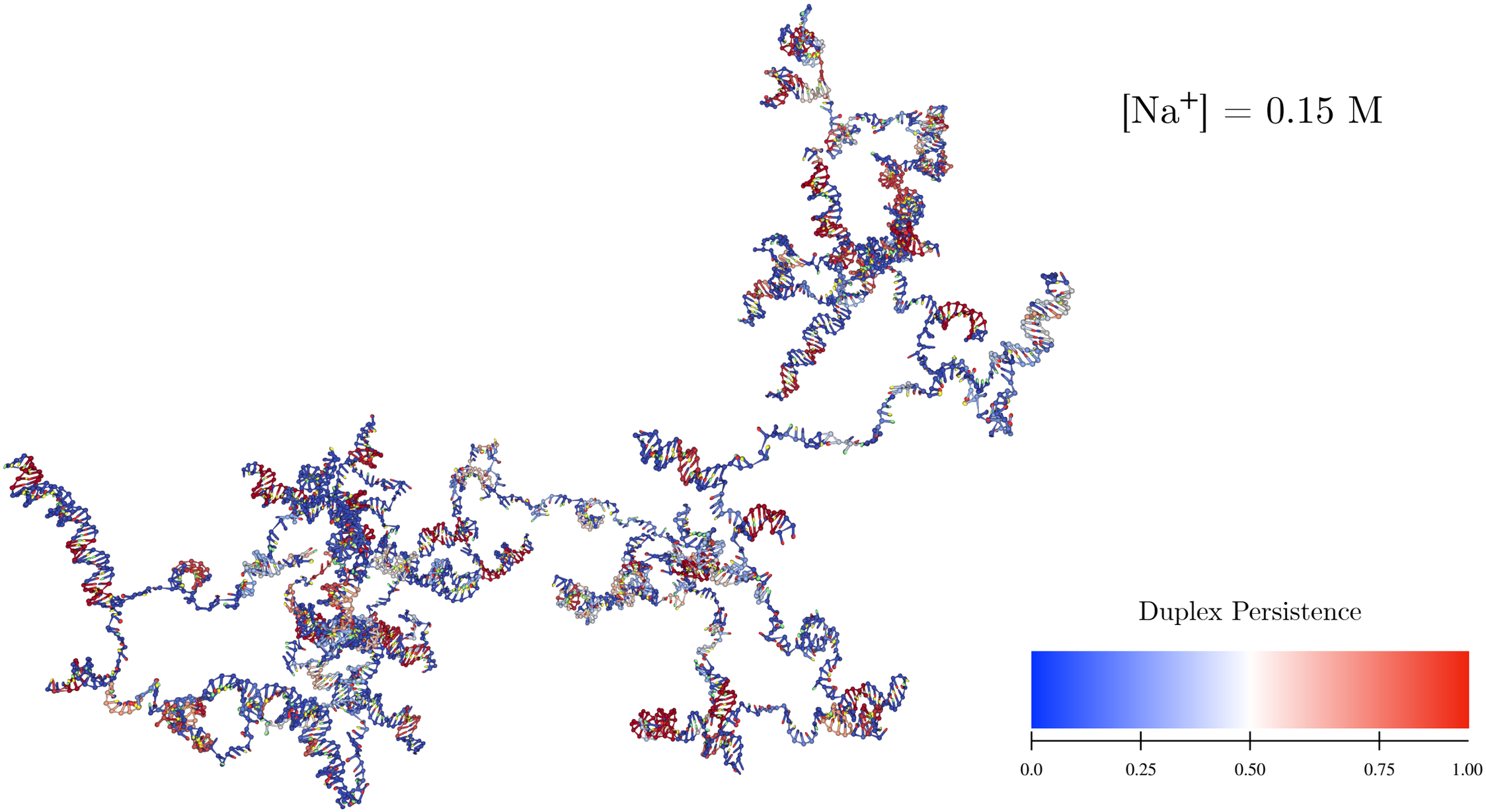


Duplexes persistence

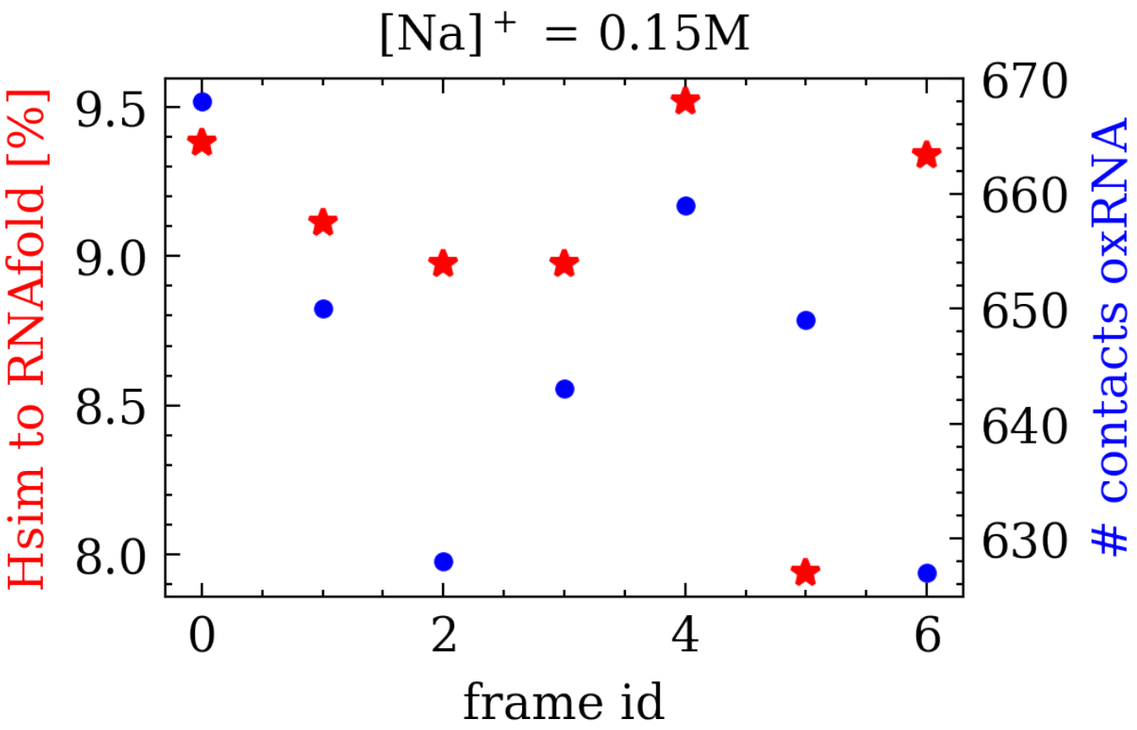


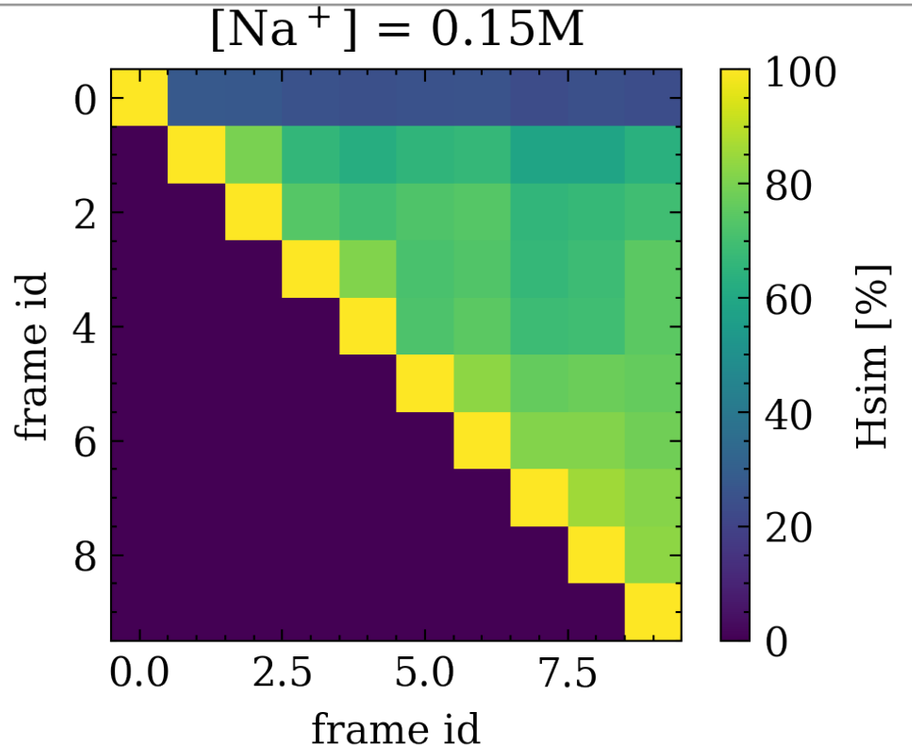


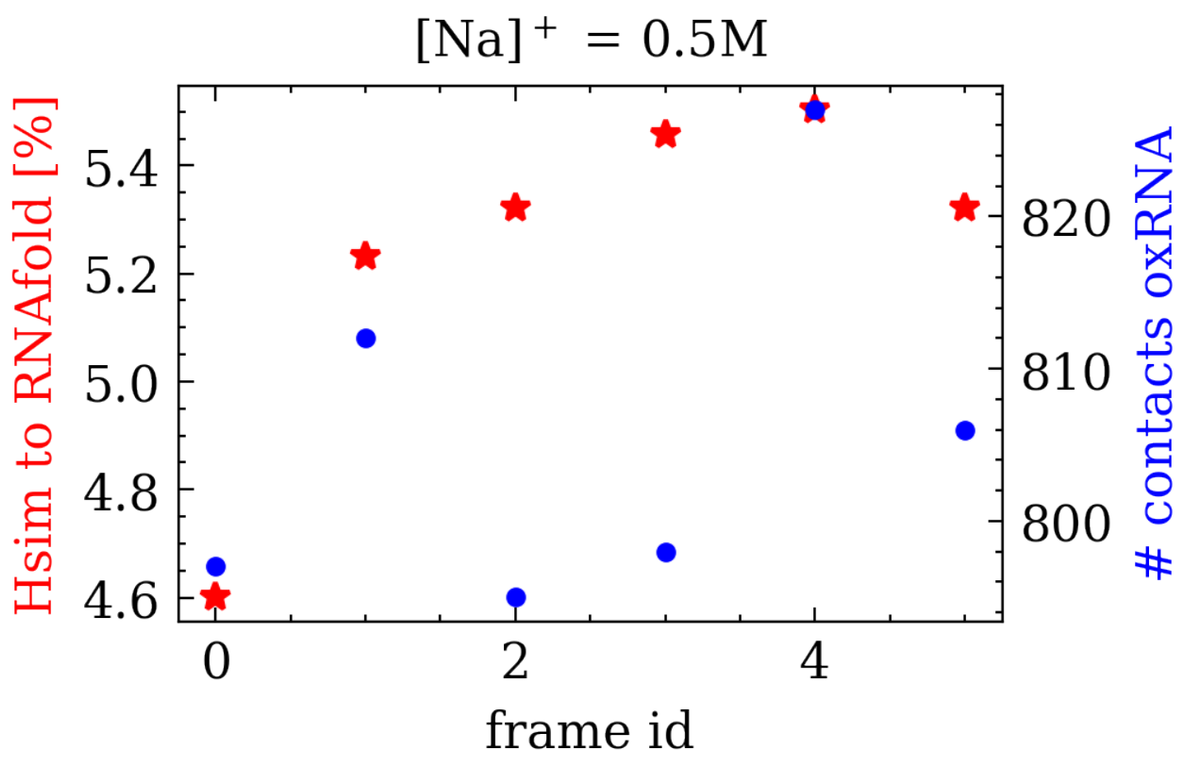
Duplexes persistence

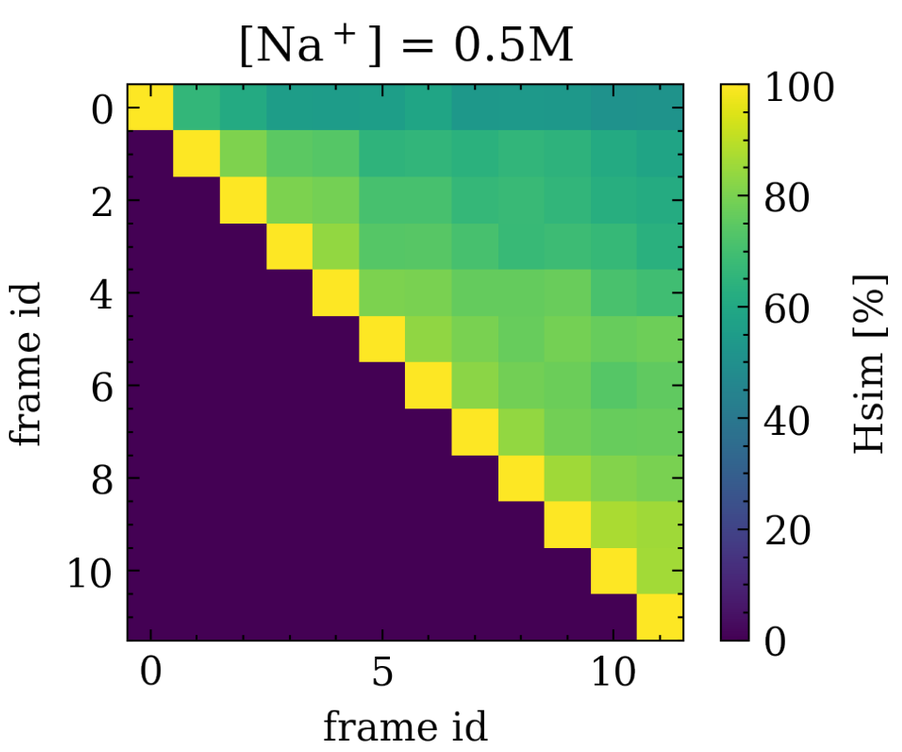


Hydrogen Bonds Analysis









[6] *Lorenz et al. ViennaRNA Package 2.0. Algorithms Mol Biol **6**, 26 (2011)*

Big thanks to my collaborators!

PoteLab:

- Prof. Raffaello Potestio
- Dr. Lorenzo Petrolli
- Manuel Micheloni
- Dr. Thomas Tarenzi
- Dr. Roberto Menichetti
- Dr. Raffaele Fiorentini
- Margherita Mele

Former PhD students:

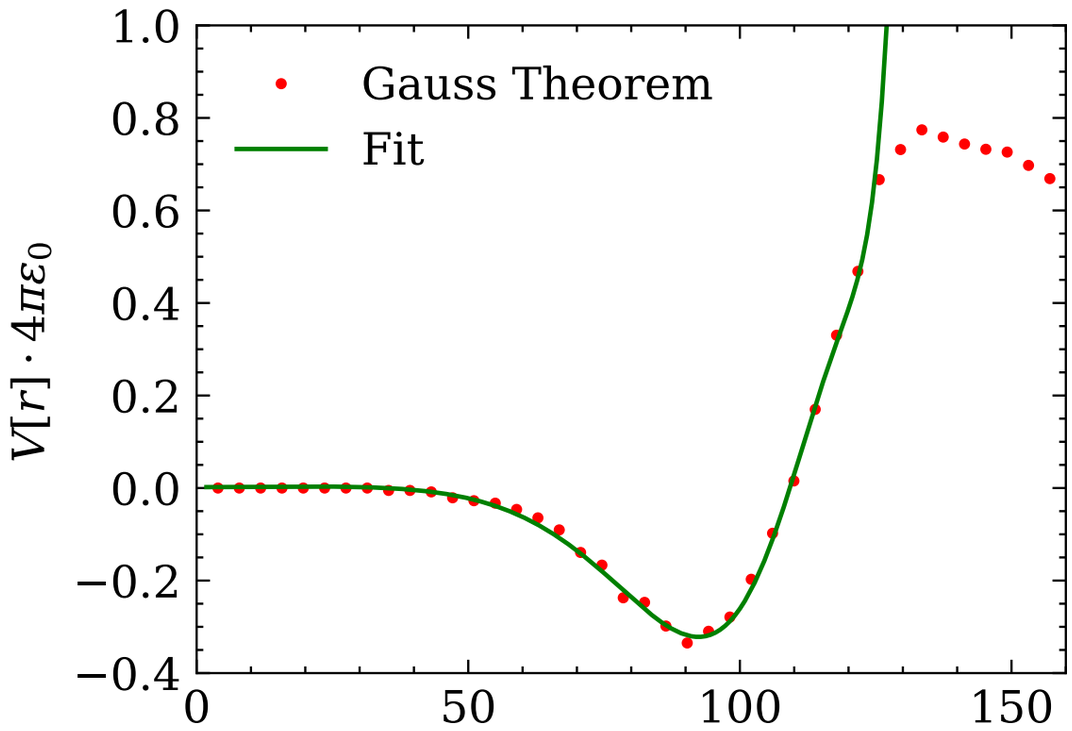
- Dr. Marco Giulini
- Dr. Marta Rigoli

EXTRAS

Perspective: spherically constrained simulations

- “Mean field” radial Coulomb potential calculated from charge distribution of atomistic capsid
- Fit to extract analytic function implemented in oxDNA
- Simulation of kinetic of spherical enclosure to see the effect on secondary/tertiary structures formation w.r.t. free filament

Capsid and ions radial electric potential



$$\hat{V}(r) = \hat{V}_0 + \sum_{i=1}^{N} \hat{V}_i(r)$$

$$\hat{V}_i(r) = \begin{cases} 0, & \text{if } r < R_i \\ \dfrac{Q_i}{4\pi\epsilon_0 r}, & \text{if } r > R_i \end{cases}$$

$$V(r,t) = e^{A[r-(r_0+kt)]}(B + Cr + Dr^2 + Er^3 + Fr^4)$$

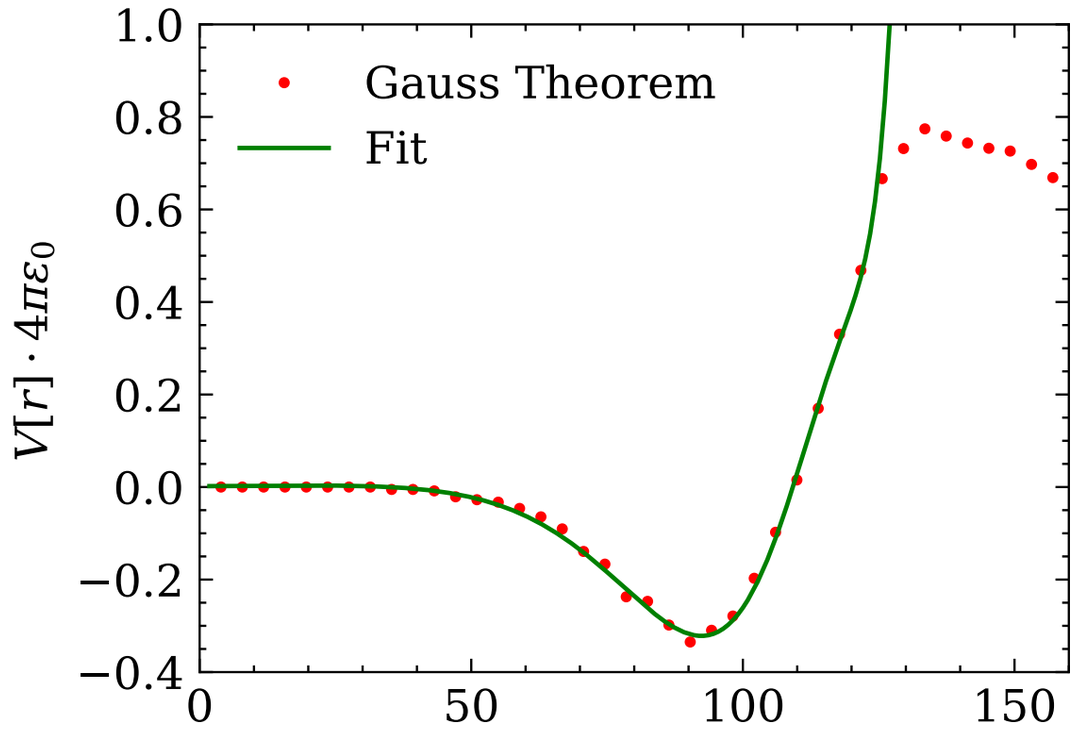
Perspective: spherically constrained simulations

- “Mean field” radial Coulomb potential calculated from charge distribution of atomistic capsid
- Fit to extract analytic function implemented in oxDNA

- Simulation of kinetic of spherical enclosure to see the effect on secondary/tertiary structures formation w.r.t. free filament

To be done…

Capsid and ions radial electric potential



$$\hat{V}(r) = \hat{V}_0 + \sum_{i=1}^{N} \hat{V}_i(r)$$

$$\hat{V}_i(r) = \begin{cases} 0, & \text{if } r < R_i \\ \dfrac{Q_i}{4\pi\epsilon_0 r}, & \text{if } r > R_i \end{cases}$$

$$V(r,t) = e^{A[r-(r_0+kt)]}(B + Cr + Dr^2 + Er^3 + Fr^4)$$

Multidimensional Scaling Analysis

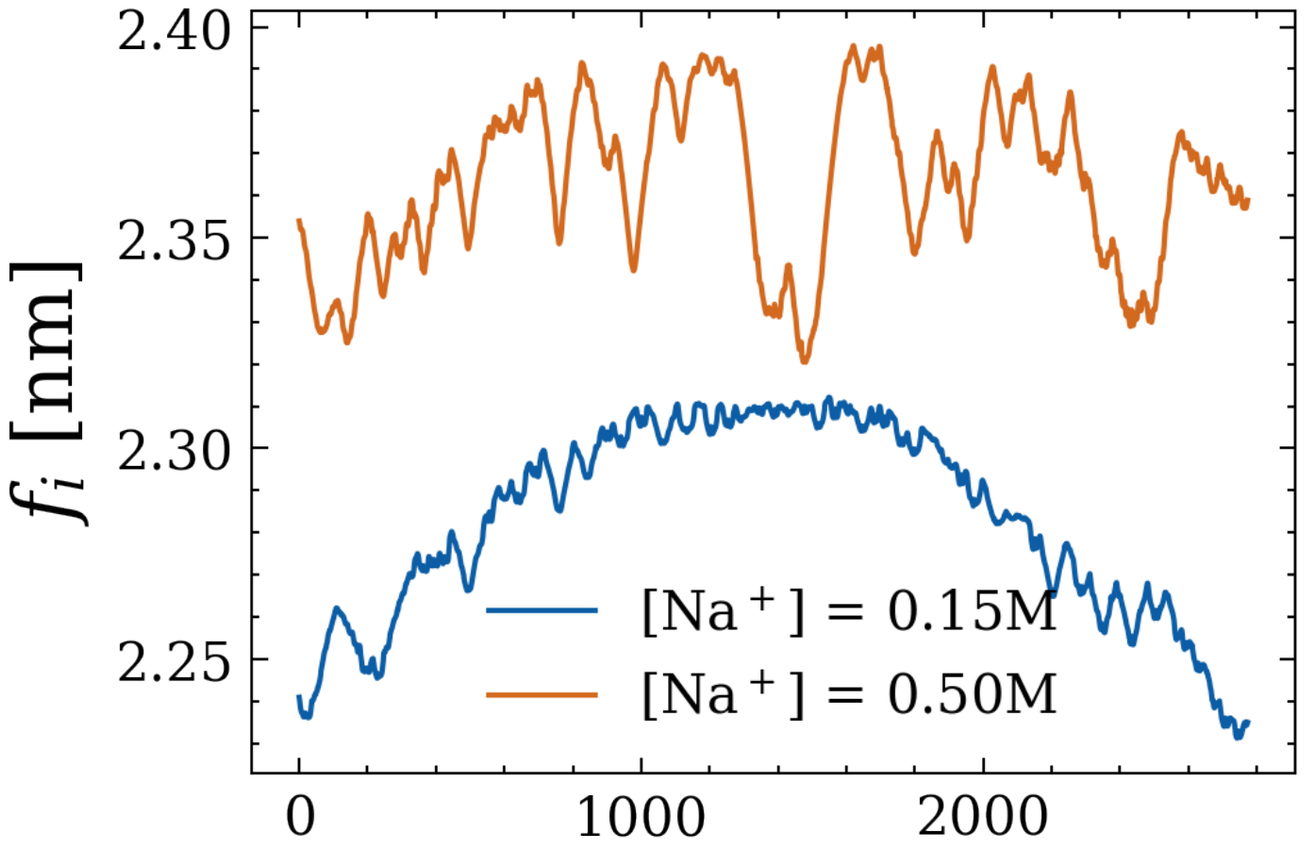
- Fluctuations per nucleotide hard to track on very mobile and disordered molecules (*e.g.* via RMSF)

- Alternative provided by MDS [CIT], alignment based on filtered distance maps

- Average variability of distances between a nucleotides and its neighborhood

- At higher salt concentration more internal degree of flexibility

- Possible artifacts of alignment due to long filament



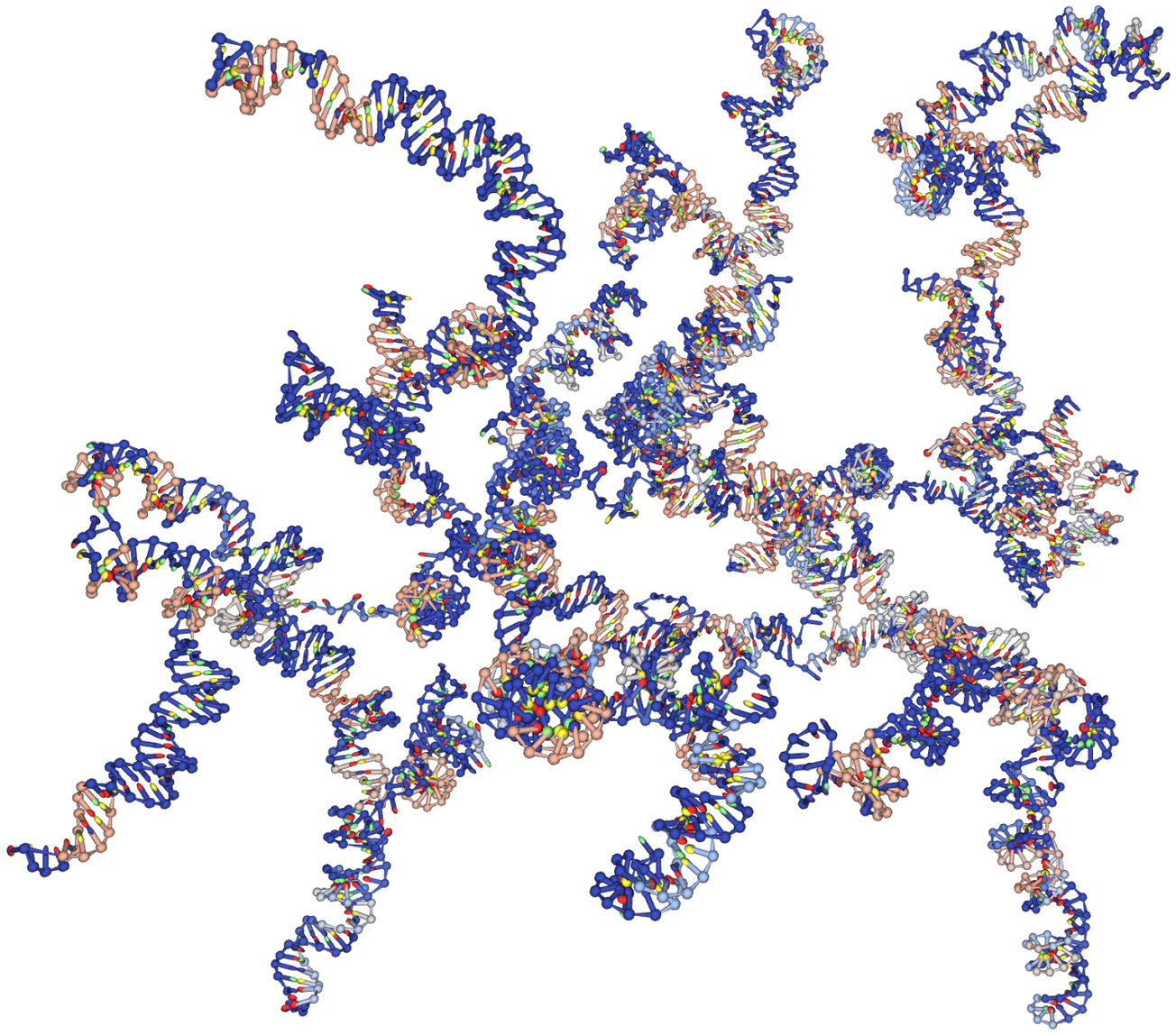
$$D_{ij}(t) \to \delta_{ij} := \langle D_{ij}(t) \rangle_t$$

$$\to S(\mathbf{r}_1, \ldots, \mathbf{r}_N) := \sum_{i \neq j} (\delta_{ij} - ||\mathbf{r}_i - \mathbf{r}_j||)^2$$

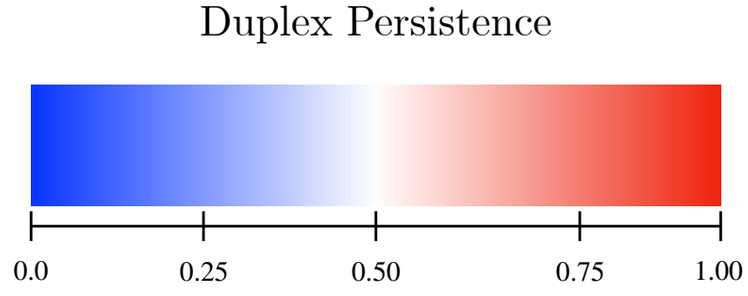
$$f_i := \langle \delta_{ij} \rangle_j = \frac{1}{N_i} \sum_{j=1}^{N_i} \delta_{ij}$$

[CIT] *J Mol Biol. 2014 March 06; 426(5): 1050-1060*

Duplexes persistence



$$[\mathrm{Na}^+] = 0.5\ \mathrm{M}$$



Multiscale Modeling

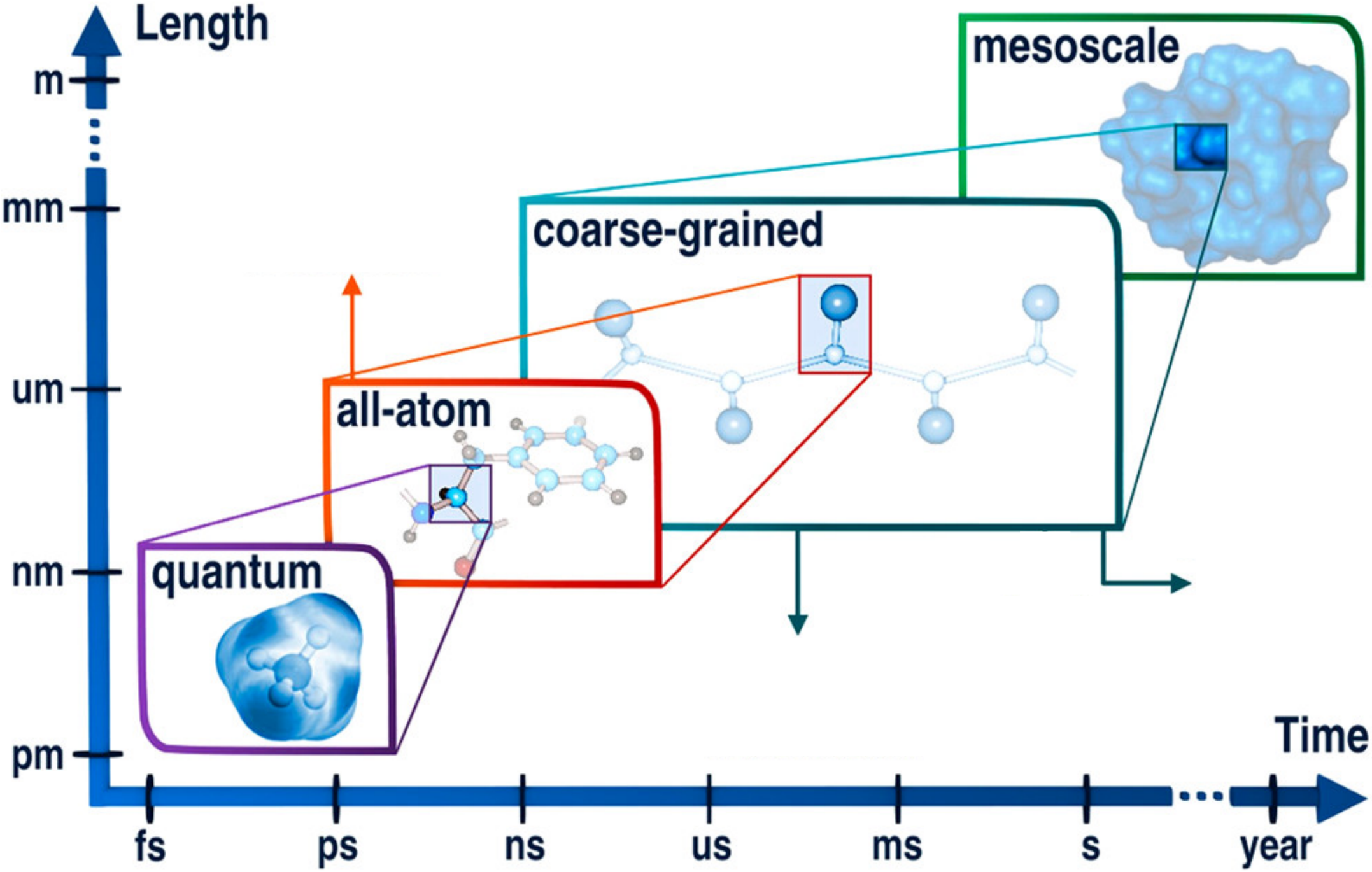


image from Chem. Rev. 2016, 116, 14, 7898–7936

Multiscale Modeling

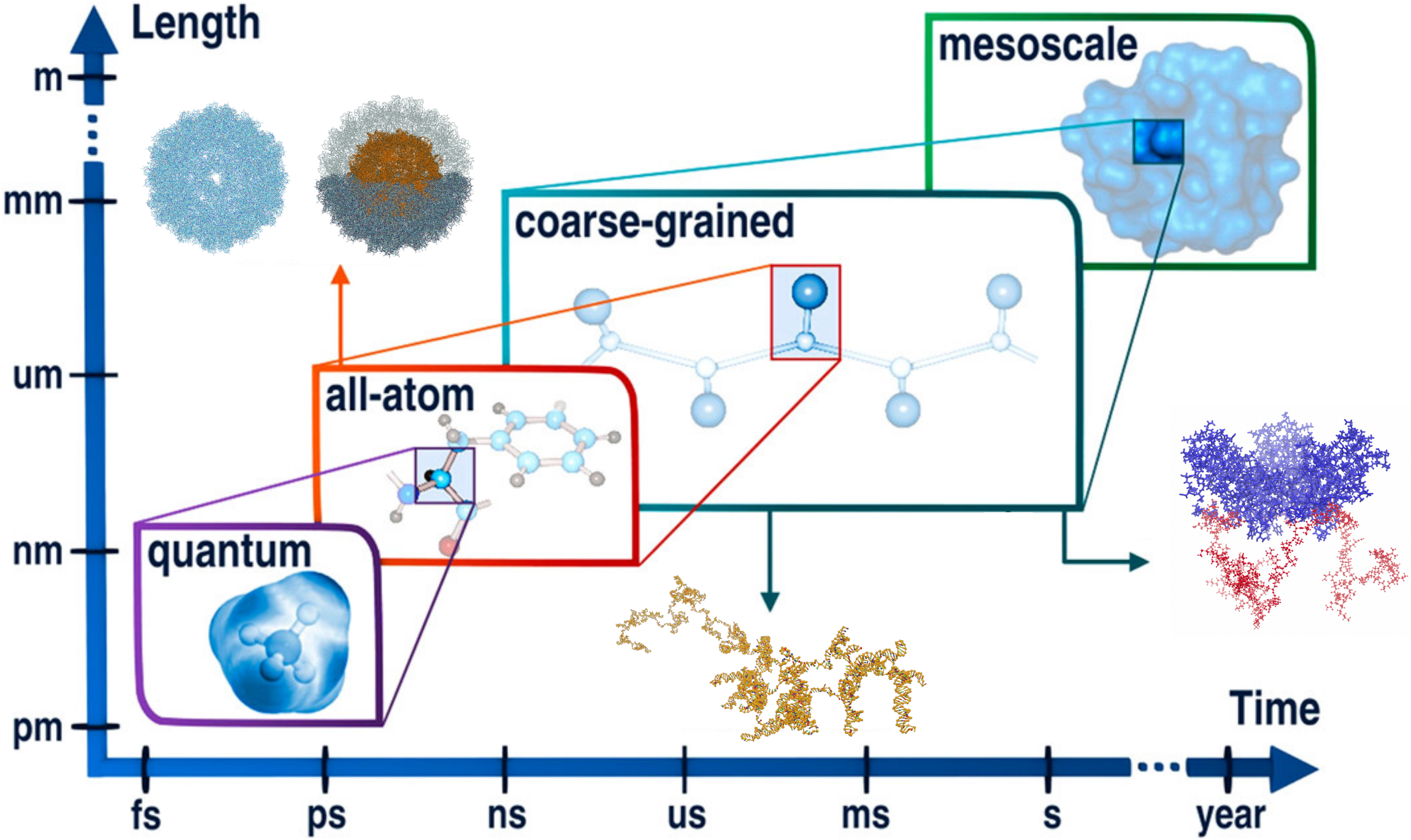
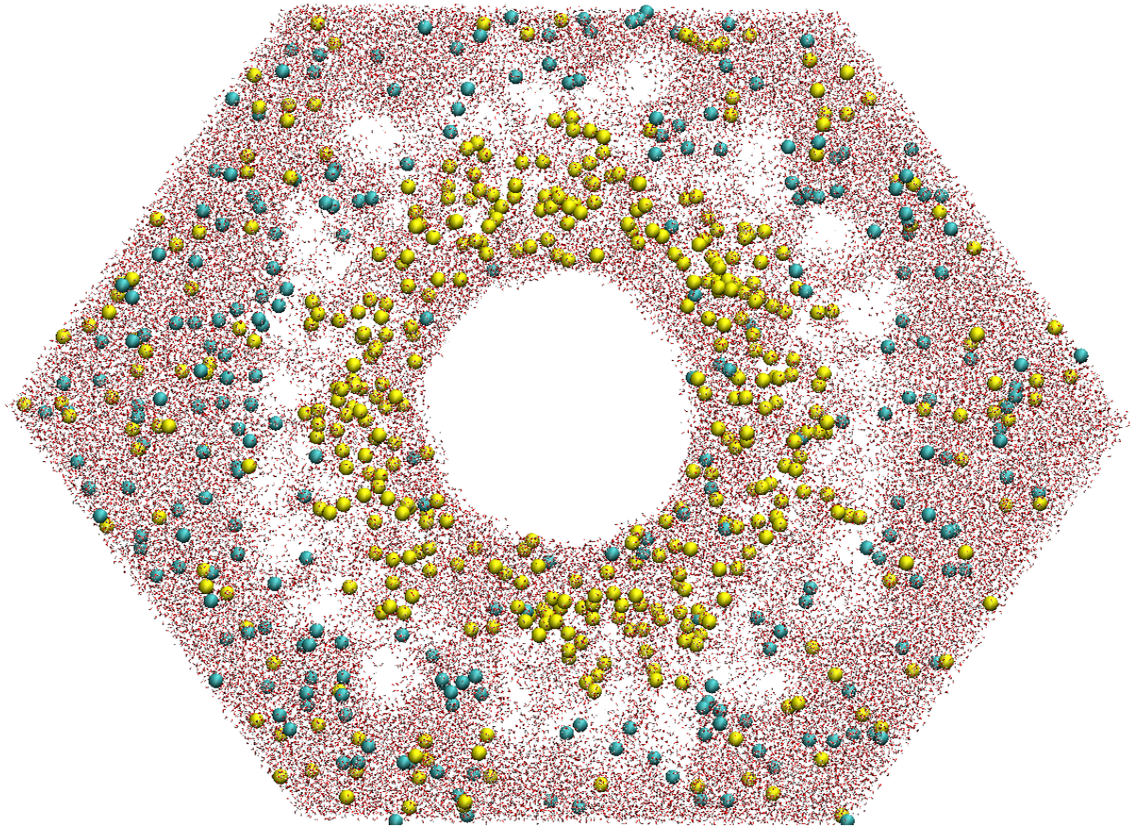
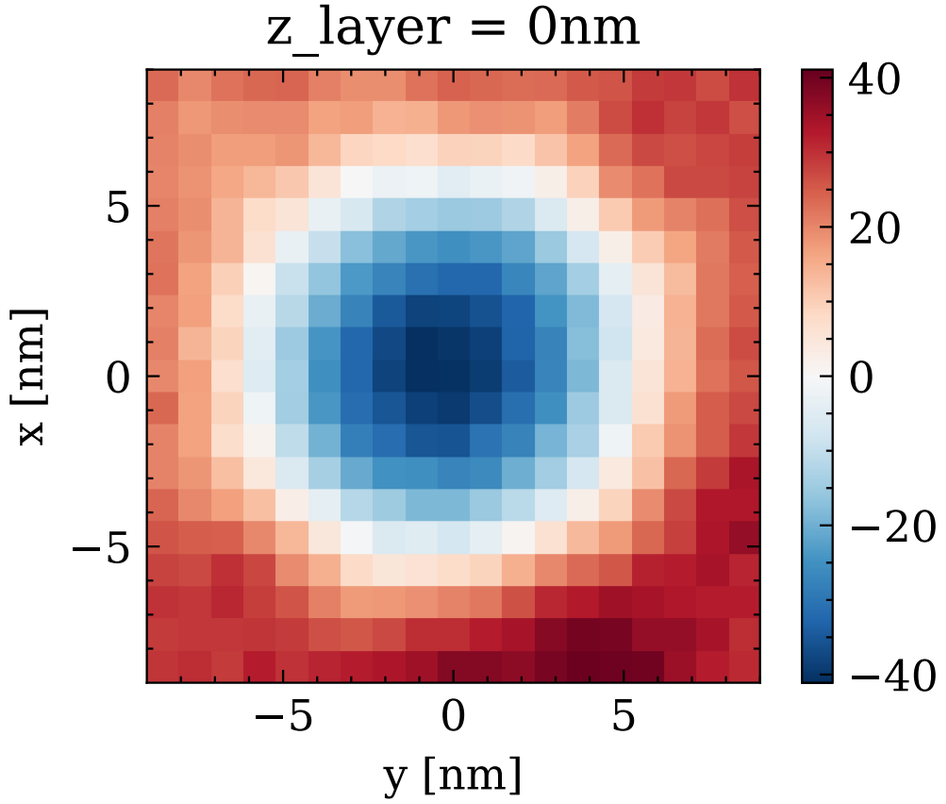


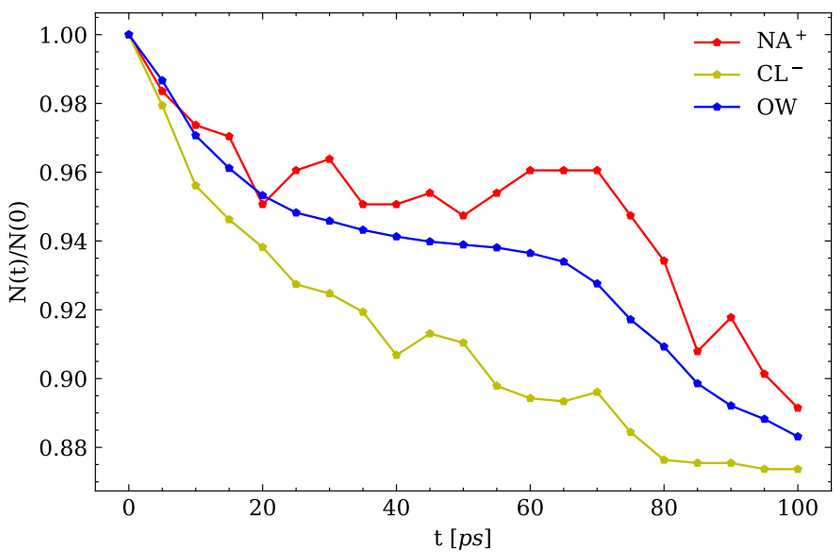
image from Chem. Rev. 2016, 116, 14, 7898–7936

Solvent behavior in AA CCMV capsid simulations

Solvent migration due to Coulomb repulsion





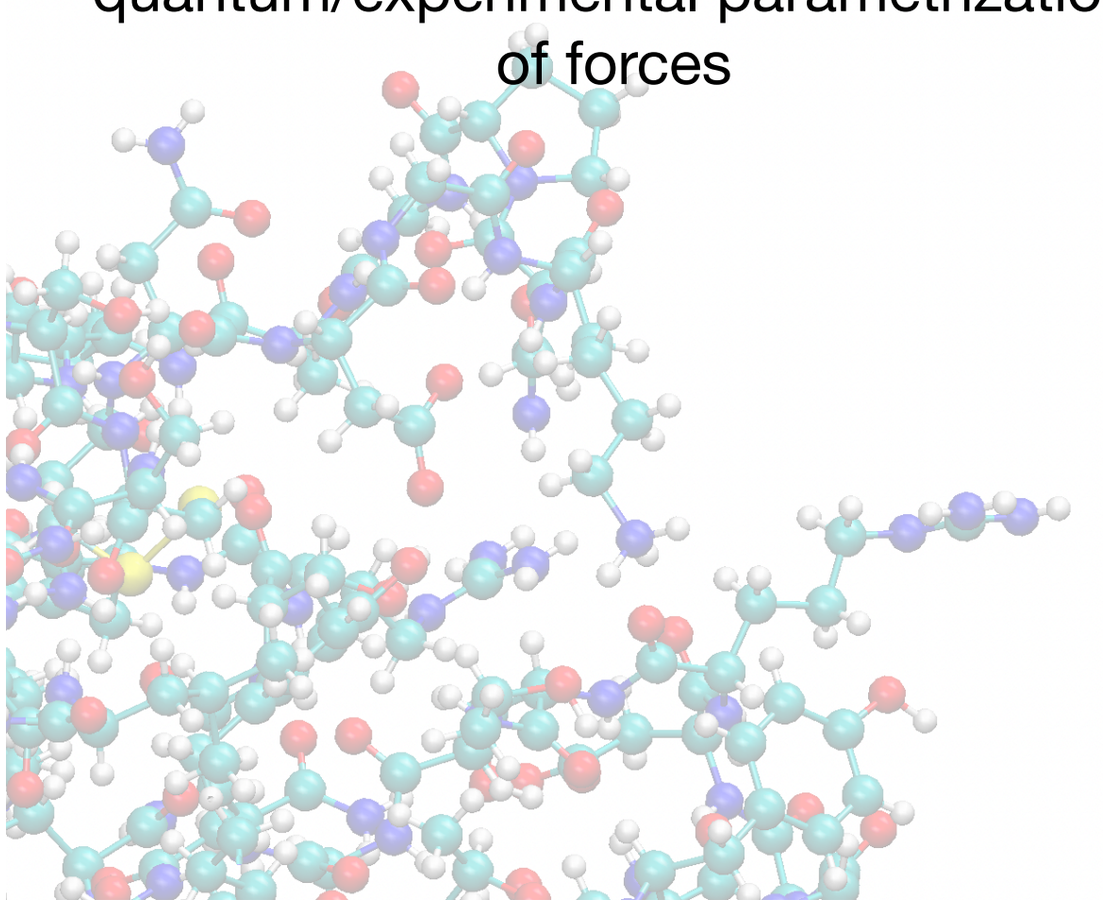


All-Atom vs Coarse-Grained

All-Atom Molecular Modeling

- Every atom's nuclei in the system is taken into account explicitly (in bio-molecules mainly H,C,O,N,S)

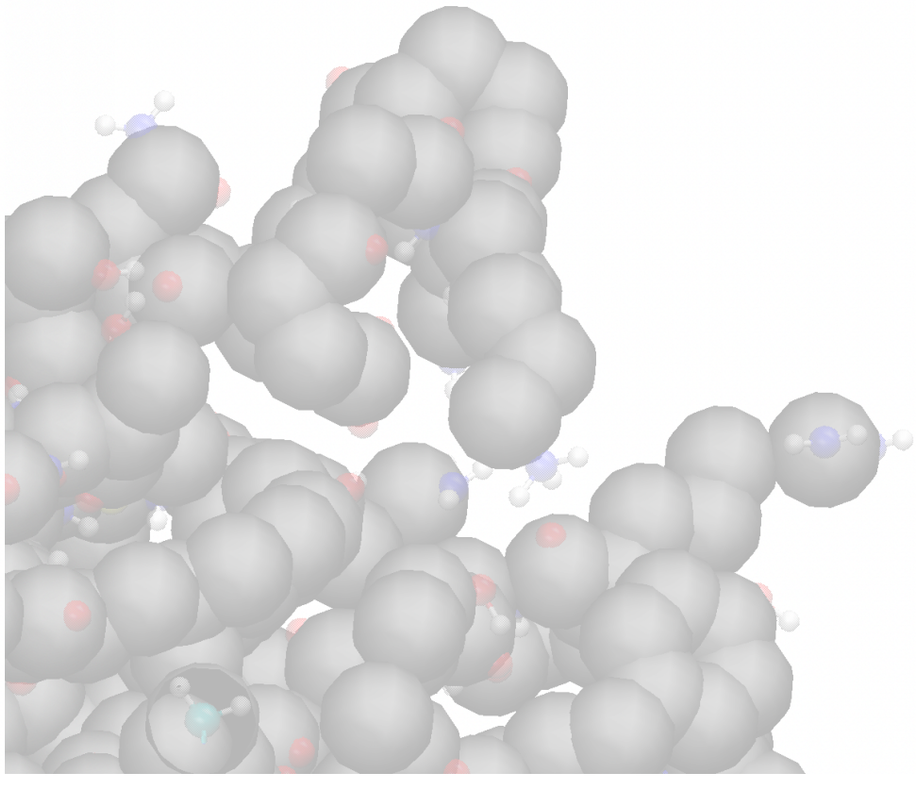
- Born-Oppenheimer approximation + quantum/experimental parametrization of forces



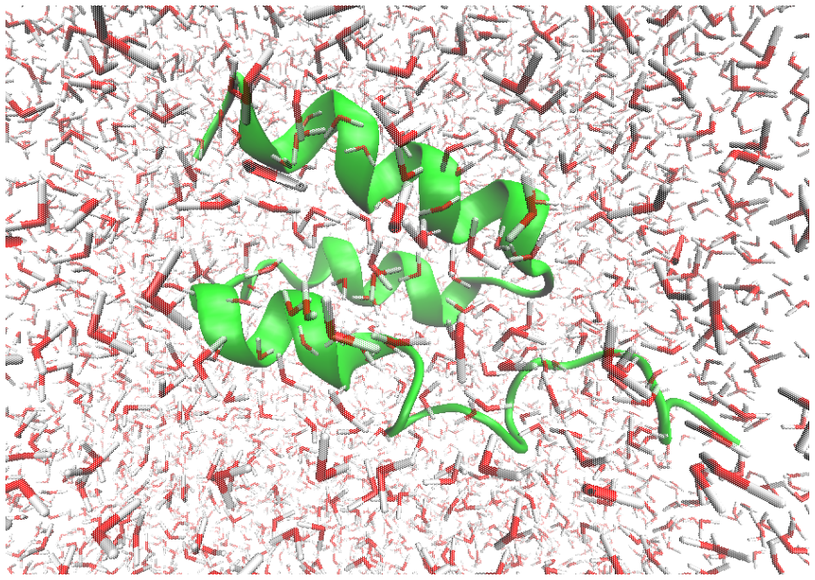
Coarse-Grained Molecular Modeling

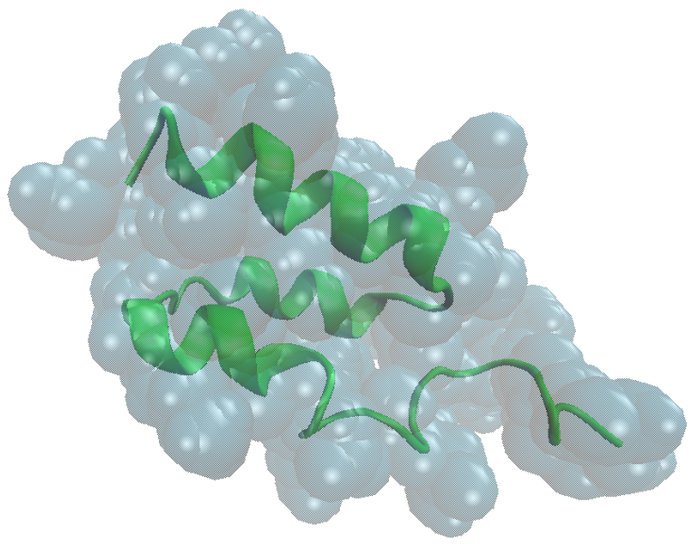
- Atoms grouped into **effective interaction sites** in a statistical mechanical based or "experimental based" approach

- Goal is twofold: reduce computational weight of simulations & simplify description in order to make it more human readable



Explicit vs Implicit Solvent Models





$$V[\{\mathbf{r}\}_S, \{\mathbf{r}\}_P] = V_{PP} + V_{SS} + V_{PS} \quad \longrightarrow \quad V_{eff}[\{\mathbf{r}\}_P] = V_{PP} + V_{eff}$$

$$e^{-\beta V_{eff}[\{\mathbf{r}\}_P]} := \int d\{\mathbf{r}\}_S \, e^{-\beta(V_{SS}[\{\mathbf{r}\}_S] + V_{PS}[\{\mathbf{r}\}_S, \{\mathbf{r}\}_P])}$$

- Higher chemical accuracy implies higher predictive power of simulations
- Super high computational cost: >90% of comp. power spent in simulating water/ions motion, then discarded in the analysis

- Limitations when effects of polarization/ entropic contribution of solvent molecules are important (conformational changes or ligand-protein binding)
- Very fast and cheap, potentially can close the gap between physical (μs-ms) and biological timescales (s)