Presenter: Alexey Savelyev

Topic:
Molecular Dynamics III

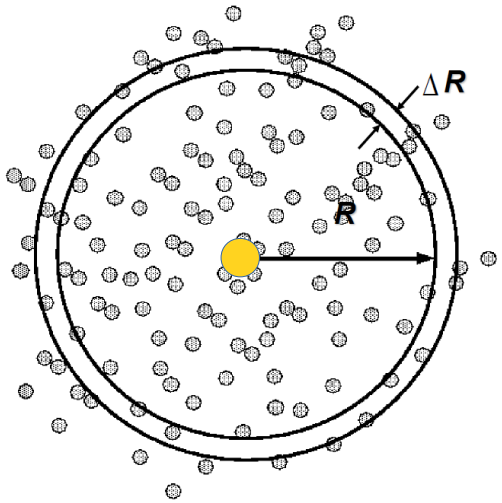
Copy of Lecture at:

https://demeler.uleth.ca/biophysics/archive/Savelyev

Last Lecture… (#2)

- **Connection to Statistical Mechanics:**
  - Thermodynamic ensembles in MD
  - Ergodicity
  - Computing space and time correlation functions
    - **Partial Specific Volume**
    - **Osmotic Pressure**

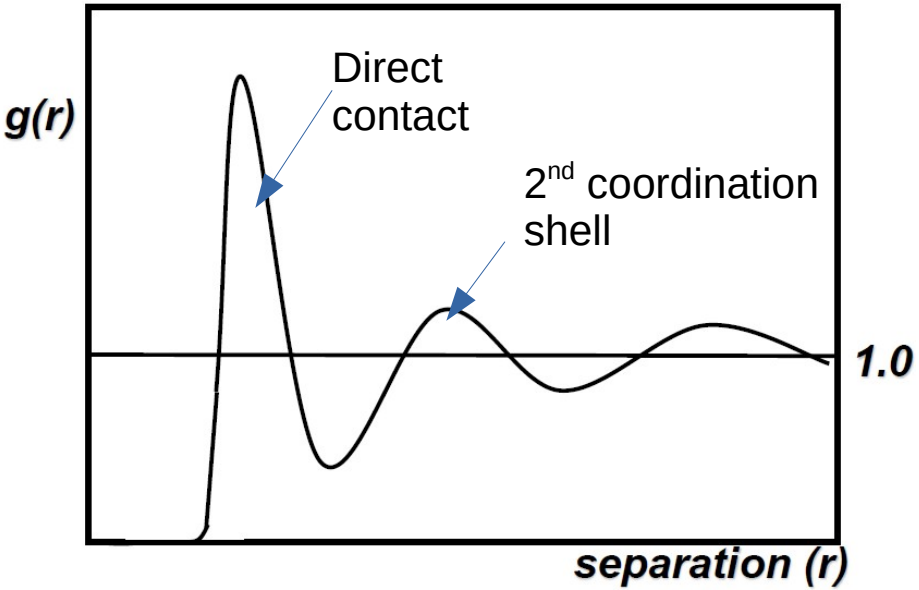
$$\rho g(\mathbf{r}) = \frac{1}{N}\left\langle \sum_{i=1}^{N}\sum_{j\neq i}^{N} \delta(\mathbf{r}-\mathbf{r}_{ij}) \right\rangle$$



$$g(r) = \frac{V}{N}\cdot\frac{\overline{N(r)}}{J(r)}$$

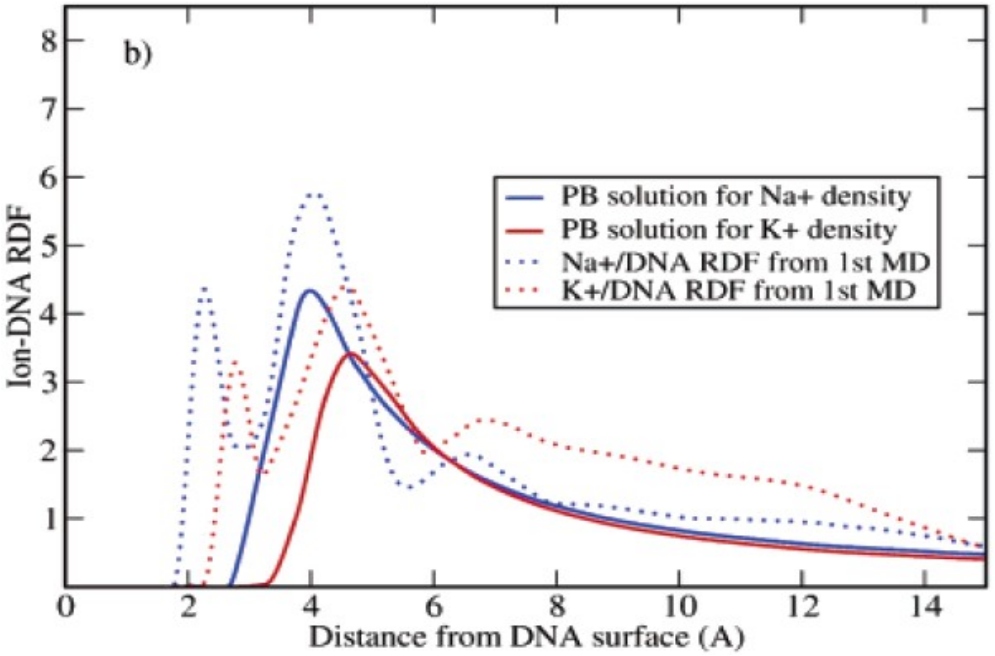
$$n(r) = \rho \int_0^r g(r)J(r)\,\mathrm{d}r$$

$$J(r) = 4\pi\, r^2$$

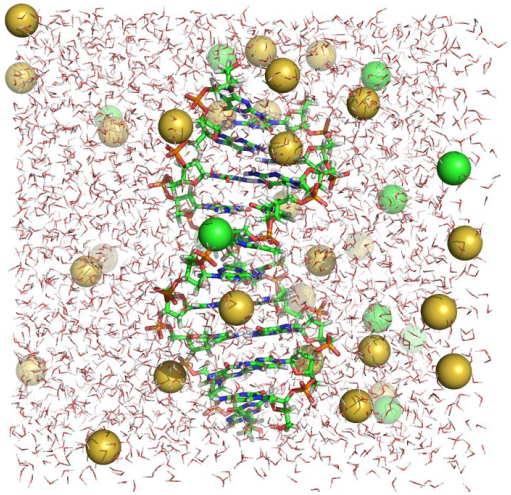


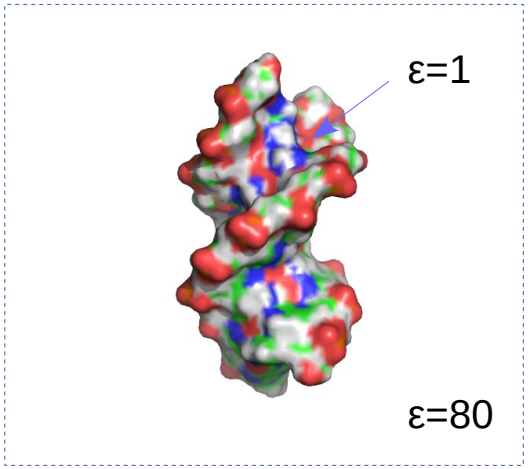
MD Simulations: cont. lecture #2

- **Ionic distribution around DNA from AA MD simulation & continuum theory**



- Oscillations in RDF from explicit solvent MD simulations come from discrete nature of the solvent and ions
- Smooth shape of Na+ & K+ density from PB calculations is caused by continuum mean-field treatment of electrostatics





Poisson-Boltznam eq.:

- Non-linear 2-order PDE for electrostatic potential given solute's charge density, ionic buffer strength and solvent (epsilon)

Savelyev & Papoian, JACS, 2006, 128

MD Simulations: cont. lecture #2

- **Partial specific volume (PSV):**

$$\Delta V = \int_{|\vec{r}|<\lambda} d\vec{r}\ \rho(\vec{r})\left(\frac{1}{\rho(\vec{r})} - \frac{1}{\rho_0}\right)$$

$$= -\int_{|\vec{r}|<\lambda} d\vec{r}\ (g_{uv}(\vec{r}) - 1) \overset{\text{def}}{=} \Delta V(\lambda)$$

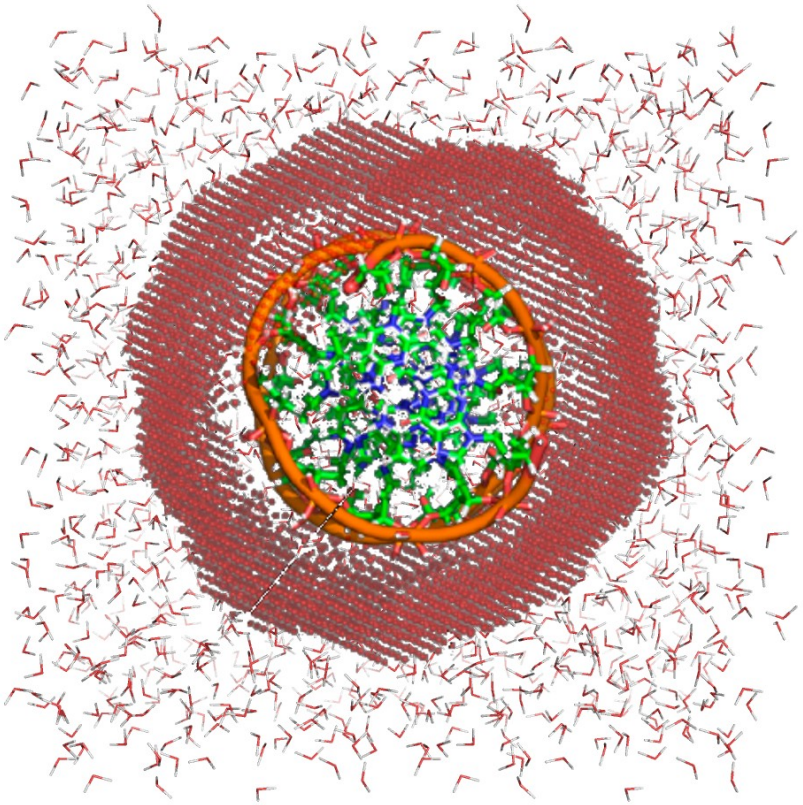
$$\bar{v}_2 = v_2 + \delta_1(v_1 - v_1^o)$$

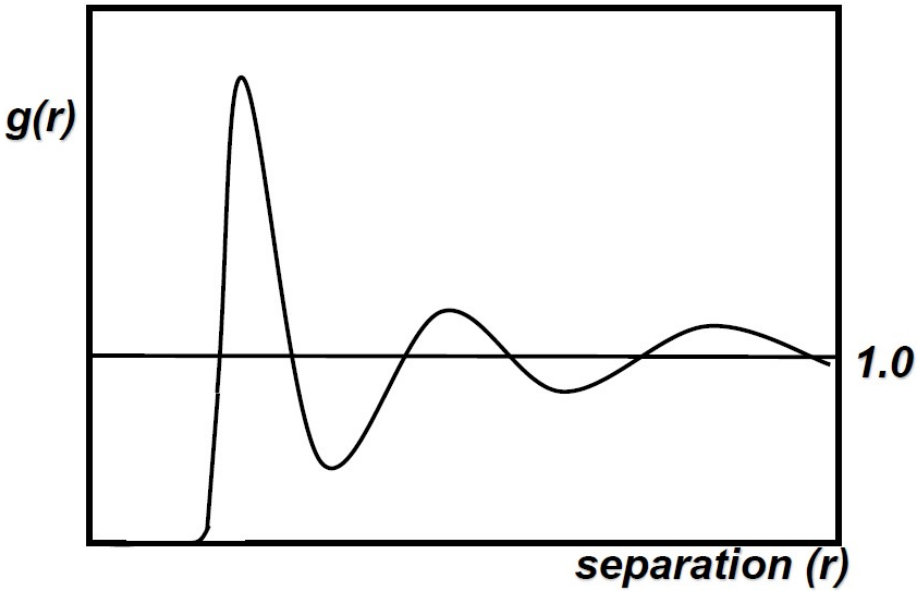
v_2 – "intrinsic" solute volume

δ_1 - # of waters in the hydration layer

v_1 - PSV of the water in hydration layer

v_0 - PSV of the water in the bulk



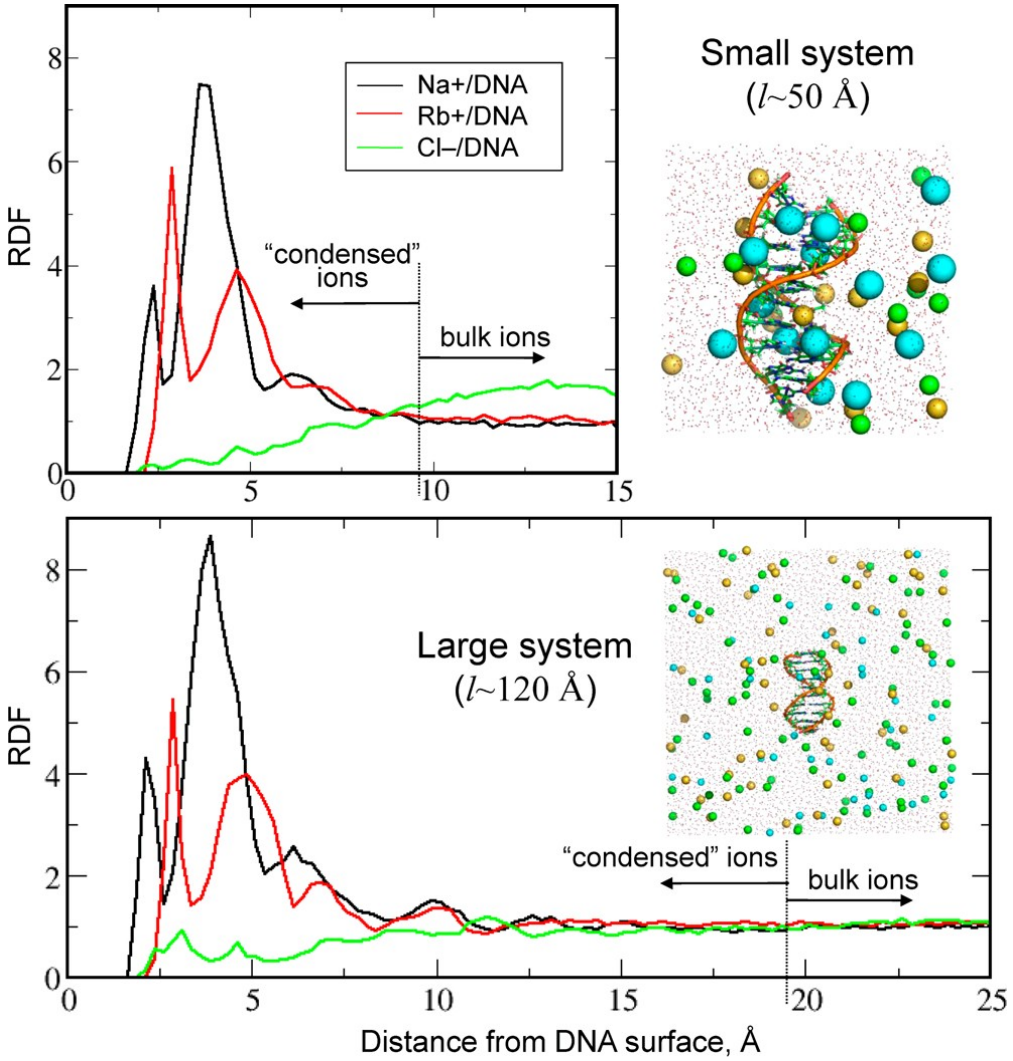


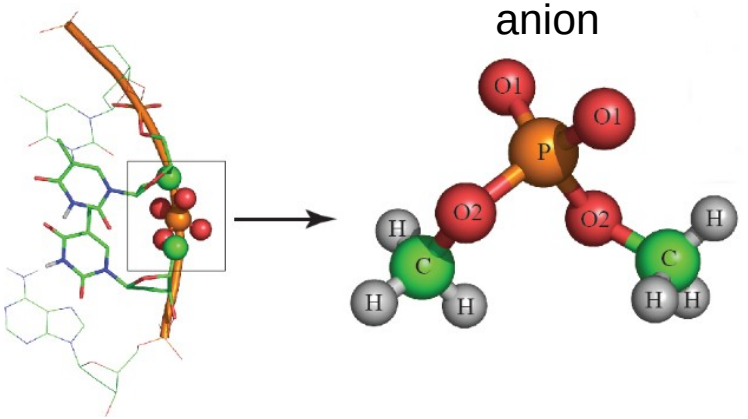
MD Simulations: cont. lecture #2

- **Osmotic pressure calculations**

Problem:

Validate energetics of interactions btw. DNA and mobile ions (Na+, K+ etc.)
Competitive ionic binding to DNA

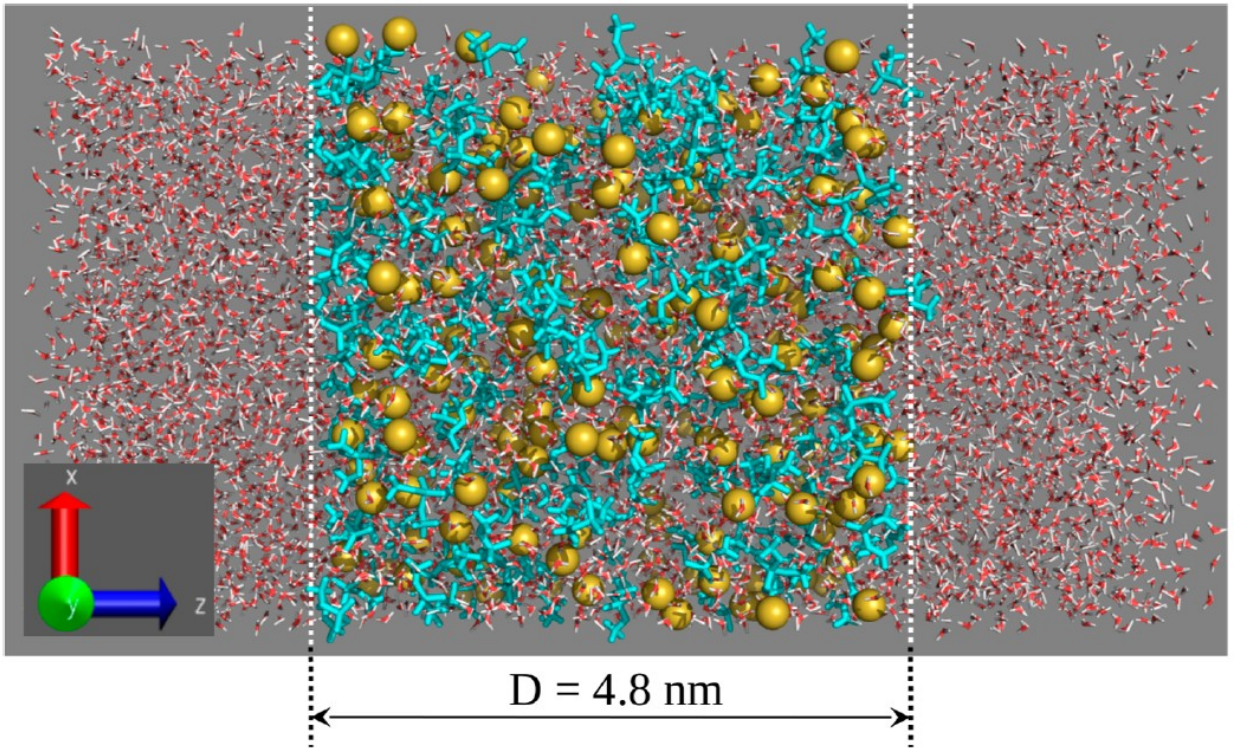




Savelyev & MacKerell, J. Phys. Chem. B, 2015, 119

MD Simulations: cont. lecture #2

- **Osmotic pressure calculations**



Experimentally available value:

- Osmotic pressure coeff.

$$\varphi = \pi / \pi_{\mathrm{id}}$$

$$\pi_{\mathrm{id}}(m) = \frac{RT}{V_{\mathrm{m}}} \nu \cdot m$$

$$\pi = F_{\mathrm{memb}} / A$$

$$F_{\mathrm{memb}} = -\sum_i k(z_i \pm D/2), \quad |z_i| > D/2$$

- Ensemble simulated: constrained NVT
  - Ratio of the unit cell in *xy*-plane is constant
  - Fluctuations along *z* axis is allowed

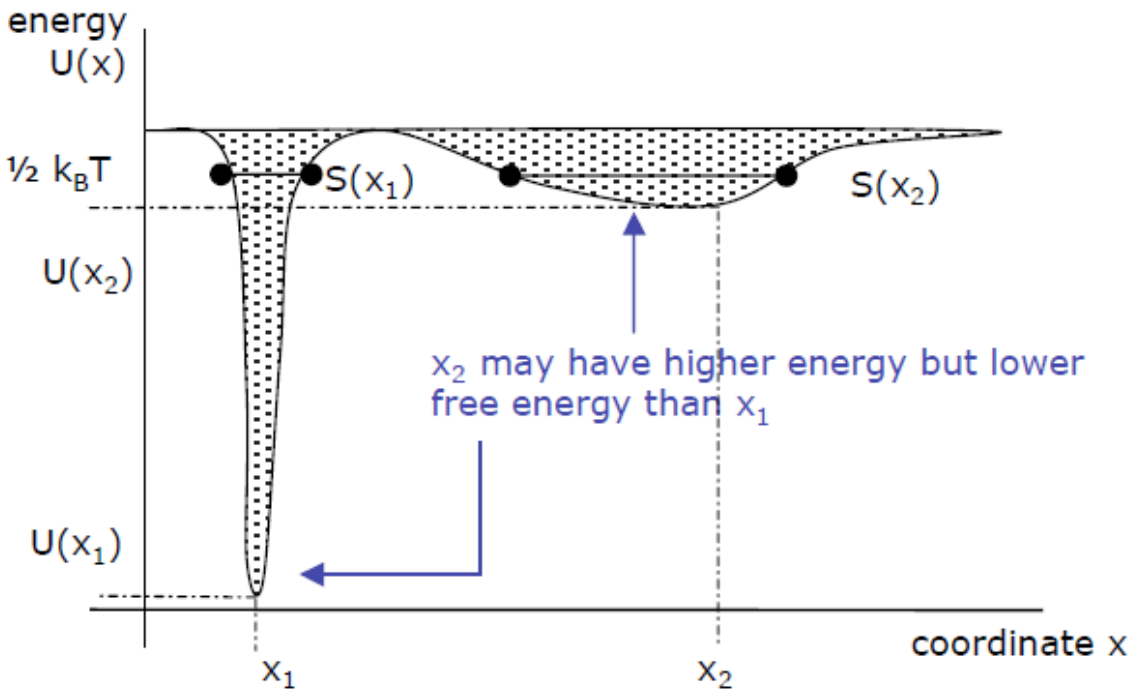
- Measure osmotic pressure of different electrolyte solutions (Na-DMP, K-DMP, Na-Cl, K-Cl,…) at different molar concentrations (~1M and ~3M) to optimize vdW interactions
- MD generates correct counter-ionic distributions around DNA
- Correct affinities of different ions towards binding to DNA

Savelyev & MacKerell, J. Phys. Chem. B, 2015, 119
Savelyev & MacKerell, J. Phys. Chem. B, 2014, 118

MD Simulations: Free Energy

- **Mechanics:** A state is characterised by one minimum energy structure (global min.)

- **Statistical mechanics:** A state is characterized by an ensemble of structures
  - Very small energy differences between microstates (~$k_B T$ = 2.5 kJ/mol) resulting from summation over very many contributions
  - Entropic effects : Not only energy minima are of importance but whole range of x-values with energies ~$k_B T$



The free energy (A) governs the system!!!

$$A = U - TS$$

MD Simulations: Free Energy

- Implication of the Free Energy

In thermodynamic equilibrium free energy is at its minimum:

- **N,V,T → Helmholtz: A = U -TS**
- **N,P,T → Gibbs: G = F + PV**

Example:

Q: which of 2 phases (A,B) is more stable at given T and density?

A: compare Helmholtz free energies: A(A) vs. A(B)

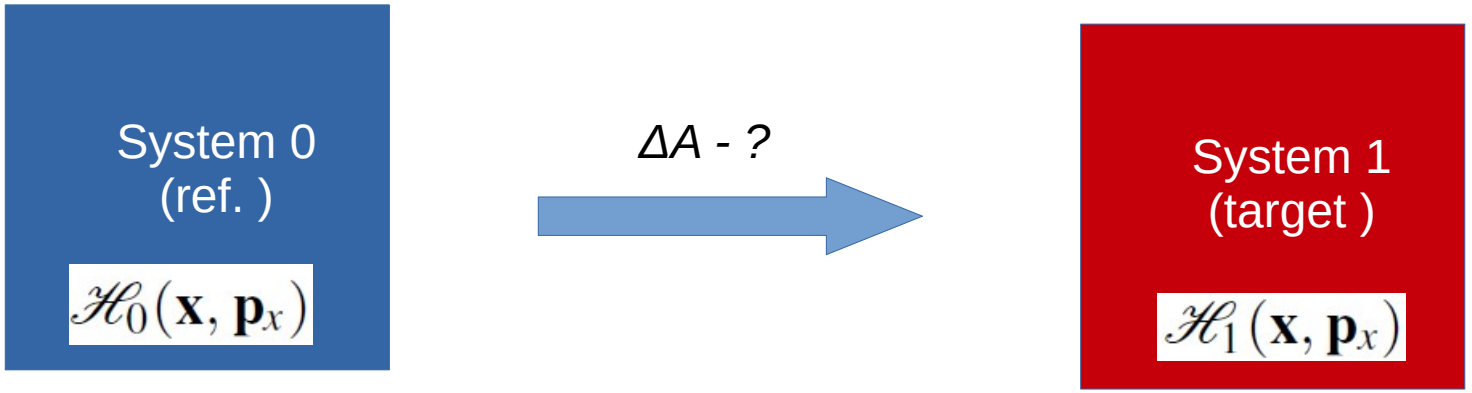
$$A = -k_B T \ln Q_{NVT} \qquad Q_{NVT} = \frac{1}{h^{3N} N!} \int \int \exp[-\frac{1}{k_B T} H(x, p_x)] dx dp_x$$

- Free energies can NOT be expressed as averages of functions of space coordinates; rather they directly depend on the available volume in phase space that is available to the system at given T.
- A, S, G (“thermal quantities”) can NOT be directly measured form MD simulations

BUT free energy differences CAN !!!

MD Simulations: Free Energy

- Free Energy Perturbation (FEP)



$$\mathcal{H}_1(\mathbf{x}, \mathbf{p}_x) = \mathcal{H}_0(\mathbf{x}, \mathbf{p}_x) + \Delta\mathcal{H}(\mathbf{x}, \mathbf{p}_x)$$

$$\Delta A = -\frac{1}{\beta} \ln \frac{Q_1}{Q_0}$$

$$\Delta A = -\frac{1}{\beta} \ln \frac{\int\int \exp\left[-\beta\mathcal{H}_1(\mathbf{x}, \mathbf{p}_x)\right] \, d\mathbf{x} \, d\mathbf{p}_x}{\int\int \exp\left[-\beta\mathcal{H}_0(\mathbf{x}, \mathbf{p}_x)\right] \, d\mathbf{x} \, d\mathbf{p}_x}$$

$$= -\frac{1}{\beta} \ln \frac{\int\int \exp\left[-\beta\Delta\mathcal{H}(\mathbf{x}, \mathbf{p}_x)\right] \exp\left[-\beta\mathcal{H}_0(\mathbf{x}, \mathbf{p}_x)\right] \, d\mathbf{x} \, d\mathbf{p}_x}{\int\int \exp\left[-\beta\mathcal{H}_0(\mathbf{x}, \mathbf{p}_x)\right] \, d\mathbf{x} \, d\mathbf{p}_x}$$

$$\Delta A = -\frac{1}{\beta} \ln\langle \exp\left[-\beta\Delta\mathcal{H}(\mathbf{x}, \mathbf{p}_x)\right]\rangle_0$$

MD Simulations: Free Energy

- Free Energy Perturbation (FEP)

$$\Delta A = -\frac{1}{\beta} \ln\langle \exp\left[-\beta \Delta \mathcal{H}(\mathbf{x}, \mathbf{p}_x)\right]\rangle_0$$

- In FEP the instantaneous change from one state to another is sampled over a canonical ensemble.
- FEP corresponds to fast growth with the constraint *immediately* moved to the target value.
- The *term 'perturbation' is misleading* because the method is exact and does not correspond to a perturbation theory in the usual sense

MD Simulations: Free Energy

- Free Energy Perturbation (FEP)

$$\Delta A = -\frac{1}{\beta} \ln\langle \exp\left[-\beta \Delta \mathcal{H}(\mathbf{x}, \mathbf{p}_x)\right]\rangle_0$$

- Small ΔH does NOT imply that the free energy difference between the reference and the target states must be small.
- Small free energy differences do NOT imply successful application of the direct FEP technique
- FEP will only provide accurate estimates of free energy differences under the condition that the target system be "*sufficiently similar*" to the reference system.

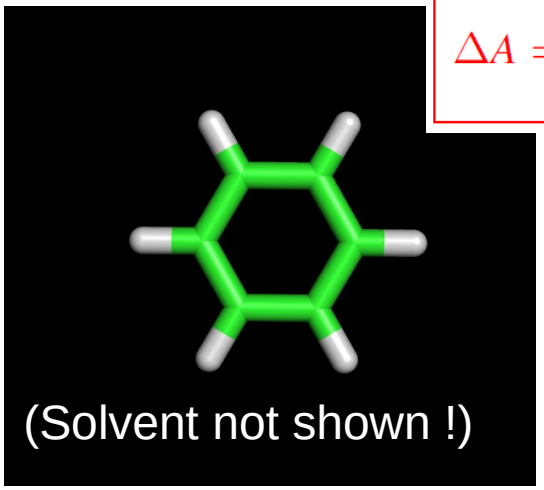
In practice: while simulating system 0, we also assess the energy of state 1

- This means that partition functions, or density of states of 0 and 1 must overlap
- Is NOT always the case

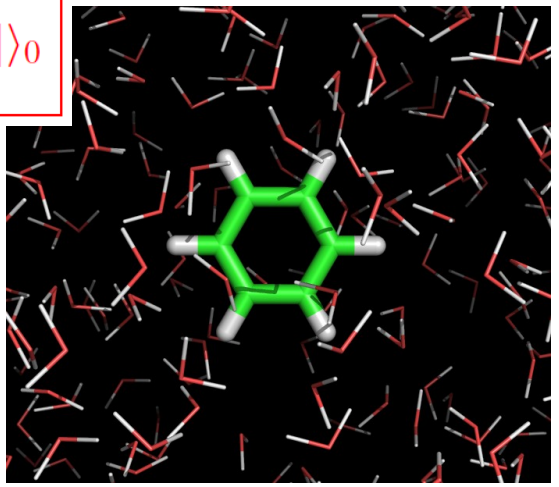
MD Simulations: Free Energy

- Free Energy Perturbation (FEP): example – hydration of benzene

$$\Delta A = -\frac{1}{\beta} \ln\langle \exp\left[-\beta \Delta \mathscr{H}(\mathbf{x}, \mathbf{p}_x)\right]\rangle_0$$



ΔA - ?



$$\mathscr{H}_1(\mathbf{x}, \mathbf{p}_x) = \mathscr{H}_0(\mathbf{x}, \mathbf{p}_x) + \Delta \mathscr{H}(\mathbf{x}, \mathbf{p}_x)$$

Solvent-solute interactions
are turned ON (LJ, Coulomb)

- **Example**: Although the hydration free energy of benzene is only **~0.767 kcal/mol at 298K**, this quantity cannot be successfully calculated by direct application of the FEP equation to a simulation of a reasonable length, because low–energy configurations in the target ensemble, which do not suffer from the overlap between the solute and solvent molecules, are not sampled in simulations of the reference state.

MD Simulations: Free Energy

- **Dealing with large perturbations / bad or insufficient sampling**

  - The difficulty in applying FEP theory can be circumvented through a *stratification strategy*, or *staging*.

  - It relies on constructing several **intermediate states** between the reference and the target state such that the direct evaluation of the free energy difference between 2 consecutive states, $A_{i;i+1}$, is reliable.

$$\Delta A = \sum_{i=1}^{N-1} \Delta A_{i,i+1} = -\frac{1}{\beta} \sum_{i=1}^{N-1} \ln \left\langle \exp\left(-\beta \Delta U_{i,i+1}\right) \right\rangle_i$$

MD Simulations: Free Energy

- Intermediate states do not need to be physically meaningful, i.e. they do not have to correspond to systems that actually exist.
- More generally, the Hamiltonian can be considered to be a function of some parameter, λ, ***an order parameter***.
- Without loss of generality λ can be defined between 0 and 1, such that $\lambda = 0$ and $\lambda = 1$ for the reference (U_0) and target (U_1) states, respectively.
- A simple choice for the dependence of the Hamiltonian on λ, the coupling parameter:

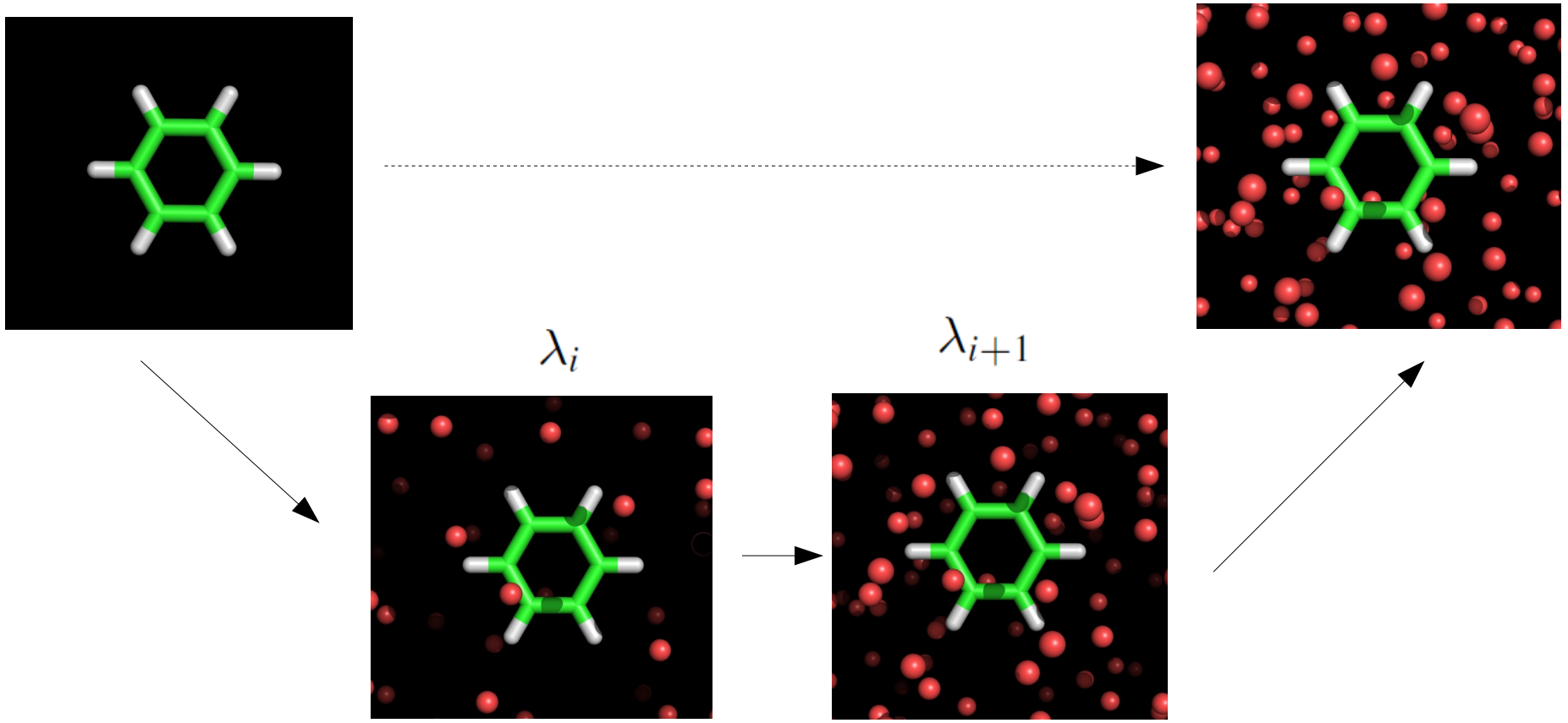
$$\mathscr{H}(\lambda_i) = \lambda_i \mathscr{H}_1 + (1 - \lambda_i)\mathscr{H}_0 = \mathscr{H}_0 + \lambda_i \Delta\mathscr{H}$$

$$\Delta A = -\frac{1}{\beta} \sum_{i=1}^{N-1} \ln \langle \exp(-\beta \Delta\lambda_i \Delta\mathscr{H}) \rangle_{\lambda_i}$$

$$\Delta\lambda_i = \lambda_{i+1} - \lambda_i$$

$$\lambda_1 = 0 \text{ and } \lambda_N = 1$$

MD Simulations: Free Energy



Graduate switching ON solvent-solute interactions

There are alternative (more commonly used) thermodynamic cycles for solvation free energies

MD Simulations: Free Energy

- **Thermodynamic Integration**

Free energy difference is calculated by defining a **thermodynamic path** between the states and integrating over ensemble-averaged enthalpy/internal energy changes along the path.

$$\Delta A_{a\to b} = A(\lambda_b) - A(\lambda_a) = \int_{\lambda_a}^{\lambda_b} \frac{dA(\lambda)}{d\lambda} d\lambda$$

$$\Delta A_{a\to b} = \int_{\lambda_a}^{\lambda_b} \left\langle \frac{\partial H(x, p_x; \lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda$$

Thermodynamic paths can either be **real chemical** processes or **alchemical** processes.

- Free energies of solvation
- Free energies of binding
- **Free energy differences associated with chemical transformations of species**

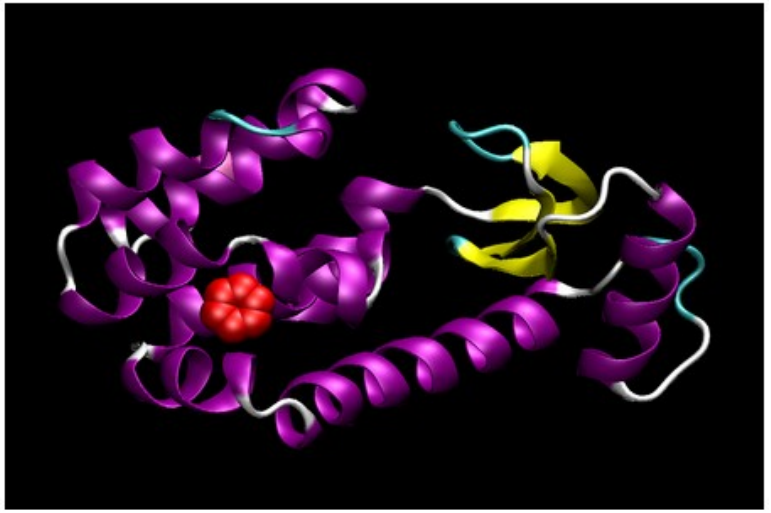
MD Simulations: Free Energy

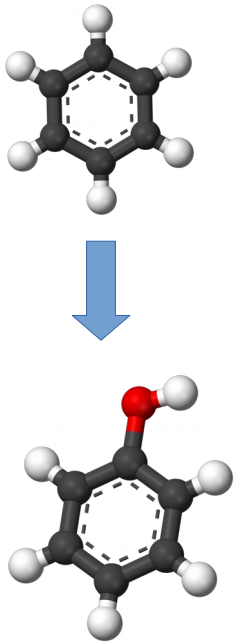
- **Thermodynamic Integration**

Modern alchemistry, done computationally, can turn structural information into “gold” of free energies of binding, mutations, or other chemical modifications...

Example:

Relative binding free energies of benzene and phenol to lysozyme

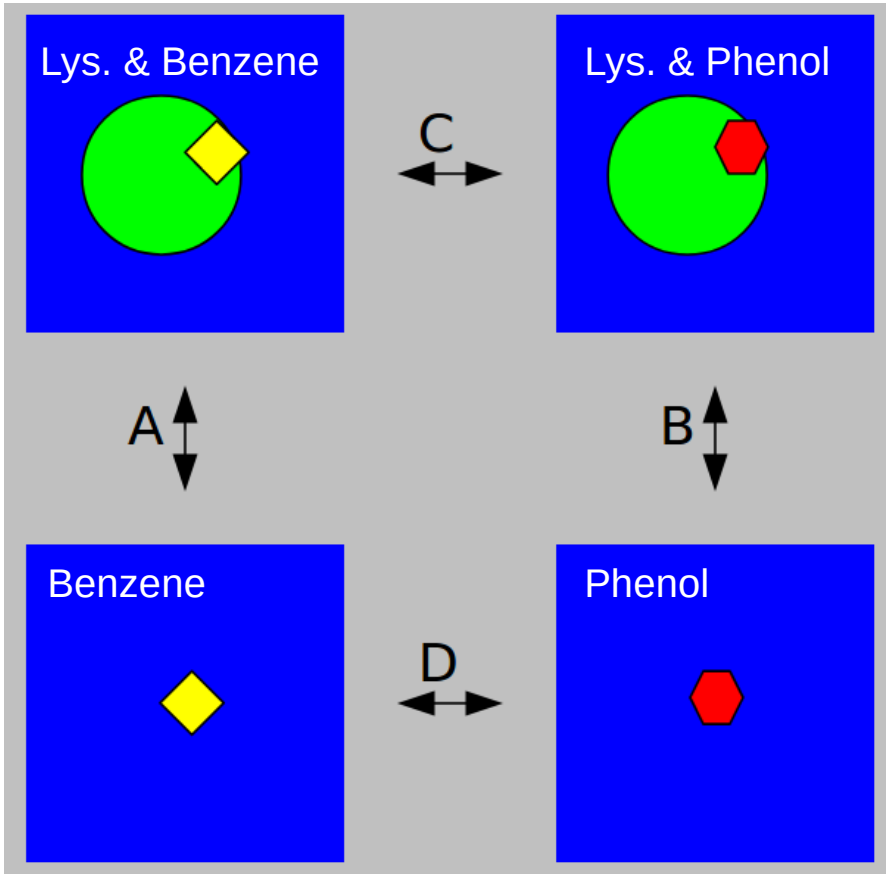




http://ambermd.org/tutorials/advanced/tutorial9/

MD Simulations: Free Energy

- Example of alchemical calculations:
  Relative binding free energies of benzene and phenol to lysozyme



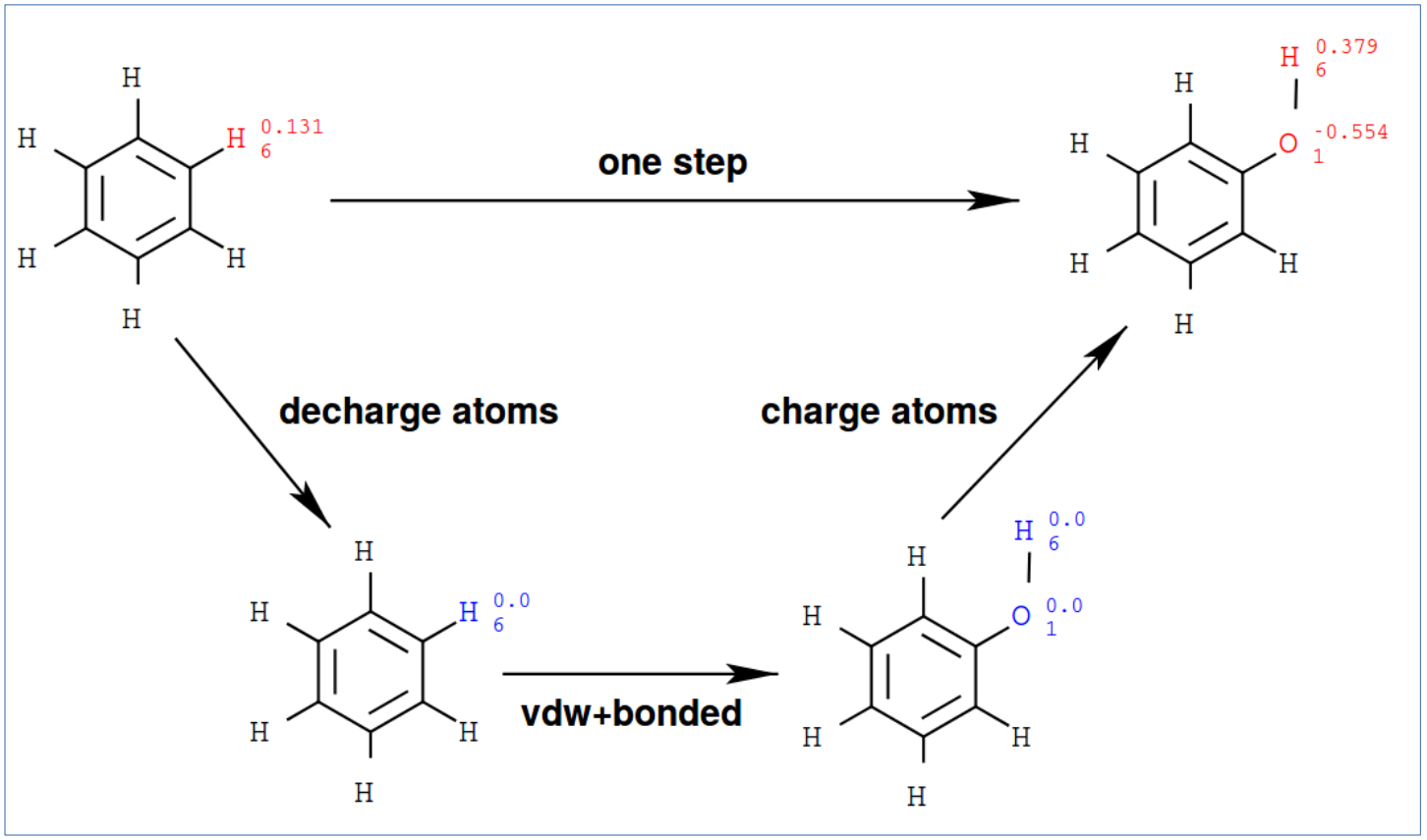
- Processes A & B are real (binding)
- Processes C & D are alchemical (B → P)

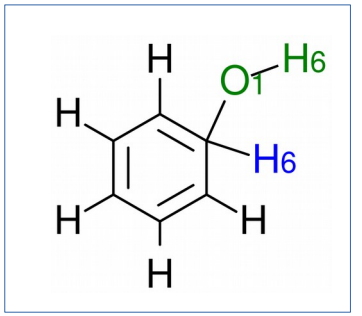
$$\Delta G(C) - \Delta G(D) = \Delta G(A) - \Delta G(B)$$

Free energy difference of the binding of benzene vs. phenol to Lysozyme is equal to the free energy difference of the (alchemical) transformation of the benzene to phenol in the bound and free state.

MD Simulations: Free Energy

- Example of alchemical calculations:
  - Relative binding free energies of benzene and phenol to lysozyme
    - Morphing benzene to phenol





"single–topology" of the benzene-phenol-in-one

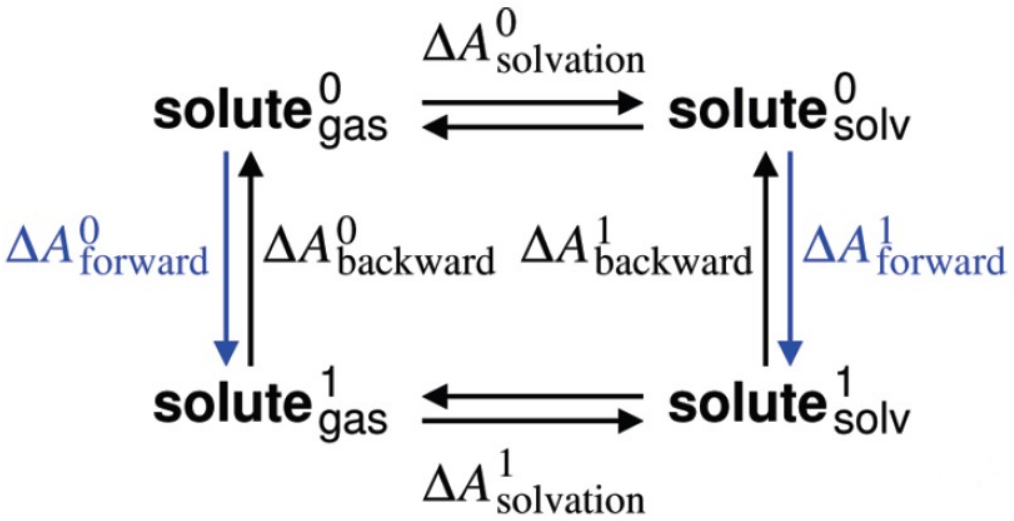
Overall Result:
Benzene binds to lysozyme stronger than phenol
(in fact, phenol does not at all)

WHY?

- Calculations are done both in solution and in a bound (to lysozyme) state
- Each step (1-3) is broken into the “windows” (λ_i); All procedure is turned to many independent (parallel) jobs
- Assumptions: no major conformational changes to the protein;

MD Simulations: Free Energy

- Solvation free energies: **relative (solvent 0 and solvent 1)**



NEED: ranking a series of ligands according to their affinity toward a given protein

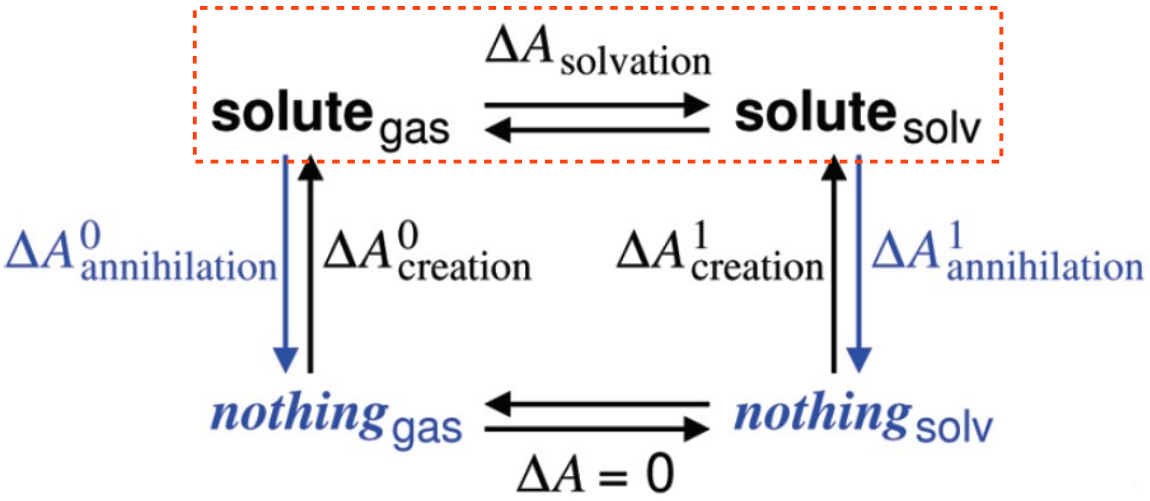
- solvation processes are described by the upper and lower horizontal legs, which corresponds to the transfer of the solute from the gas phase to the bulk solvent
- relative solvation free energies of two solutes can be determined by transforming one into another (alchemical transformations) in both the gas phase and in solution

$$\Delta A^{1}_{\text{solvation}} - \Delta A^{0}_{\text{solvation}} = \Delta A^{1}_{\text{forward}} - \Delta A^{0}_{\text{forward}}$$

Pohorille et al, J. Phys. Chem. B, 2010, 114, *Good Practices in Free-Energy Calculations*

MD Simulations: Free Energy

- Solvation free energies: **absolute**



- to avoid singularities that might arise when interatomic distances approach zero during particle creation or annihilation at the end points of the reaction pathway - Lennard-Jones potential is scaled and shifted

$$U^{\text{vdW}}(r_{ij};\lambda) = 4\varepsilon_{ij}(1-\lambda)\left[\left(\frac{\sigma_{ij}^2}{r_{ij}^2+\alpha\lambda}\right)^6 - \left(\frac{\sigma_{ij}^2}{r_{ij}^2+\alpha\lambda}\right)^3\right]$$

- solvation free energy can be measured by coupling the solute to its environment (gas or solution)
- *annihilation* (transformation to nothing) should not be taken literally – inter- and intra-molecular solute interactions are turned OFF

$$\Delta A_{\text{solvation}} = \Delta A^{1}_{\text{creation}} - \Delta A^{0}_{\text{creation}}$$

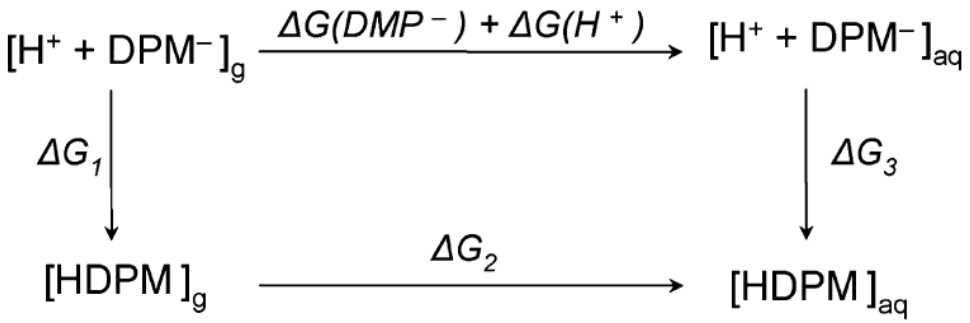
Pohorille et al, J. Phys. Chem. B, 2010, 114, *Good Practices in Free-Energy Calculations*

MD Simulations: Free Energy

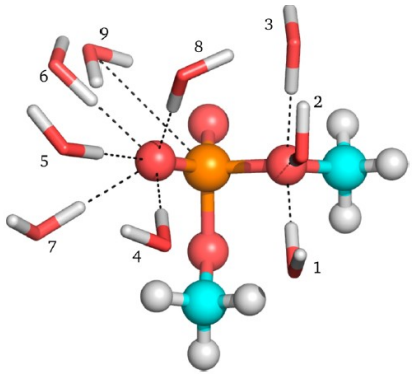
- Solvation free energies: **estimate for DMP**

What if we need to develop computational model (all-atom force field)?

- We do not have experimental solvation free energy
- We need non-bonded force field parameters to reproduce experimental free energy of solvation...



EXAMPLE:
Force-field parametrization



Solvation of Dimethyl phosphate, part of the DNA backbone

Savelyev et al, J. Phys. Chem. B, 2014

ΔG_1 – gas acidity data

ΔG_3 – from pK_a of a protonated DMP

ΔG_2 – start from free en. hydr. of related TMP, then compute relative solv. free en. btw TMP and HDMP (with QM AMSOL)

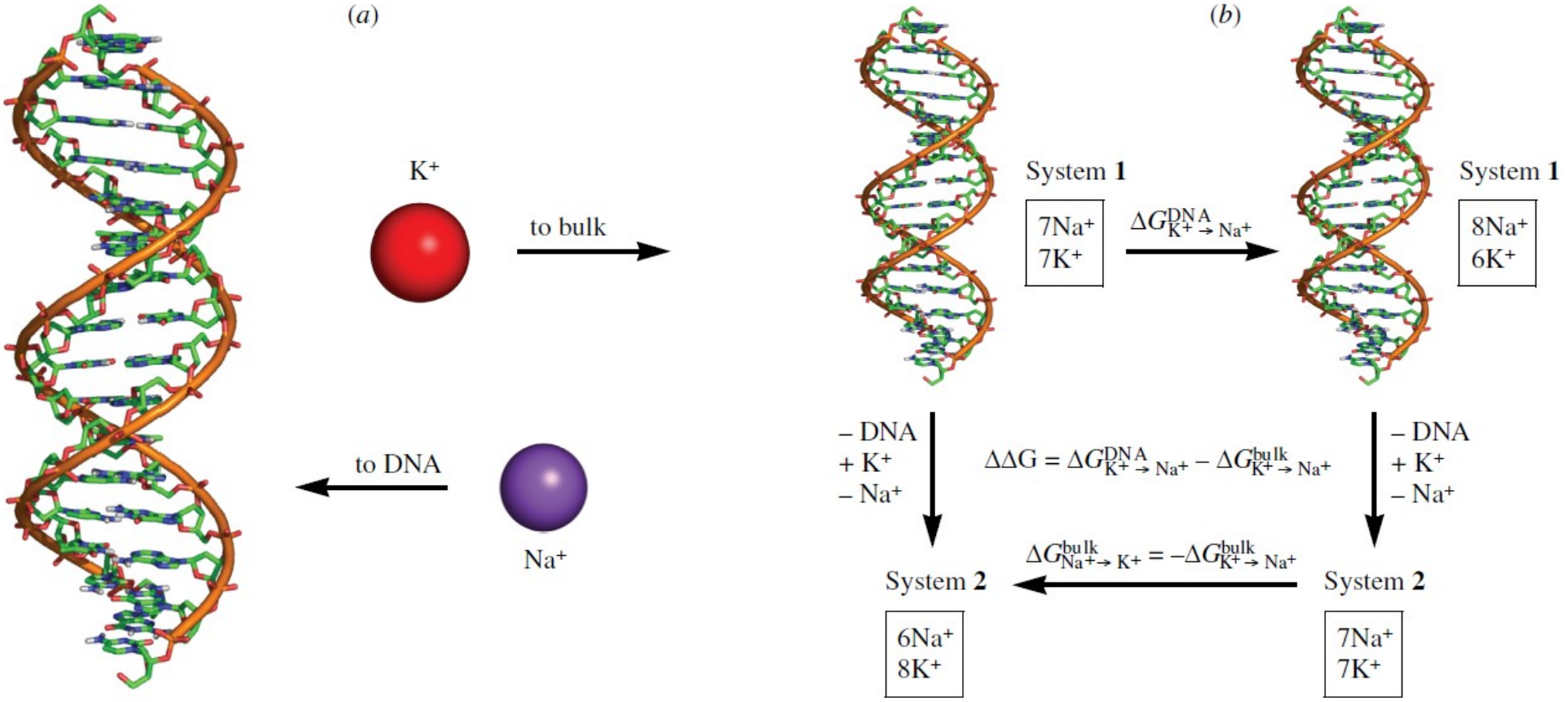
$\Delta G(H^+)$ - hydration free energy of the proton

We use the above thermodynamic cycle to estimate experimental free energy of solvation of DMP to tune interaction parameters (vdW, Coulomb)

MD Simulations: Free Energy

- Example of alchemical calculations:

  Free energy difference of counterion partitioning btw. DNA and bulk (chloride solution)



$$\Delta G \equiv G(\lambda = 1) - G(\lambda = 0) = \int_0^1 d\lambda \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle$$

Savelyev and Papoian, Mendeleev Commun., 2007, 17, 97–99

MD Simulations: Free Energy

- **Summary: why do we use thermodynamic cycles?**

  - The “alchemical” transformations require two set of simulations instead of one, one of them involving only the solute in the gas phase and is much less computationally intensive.

  - Discrepancies between the forward and the reverse transformations yield the hysteresis of the reaction, which constitutes a measure of the error in the free energy calculation

  - If the hysteresis is markedly larger than the estimated statistical errors, it is usually indicative of ergodicity issues during the transformations

MD Simulations: Free Energy

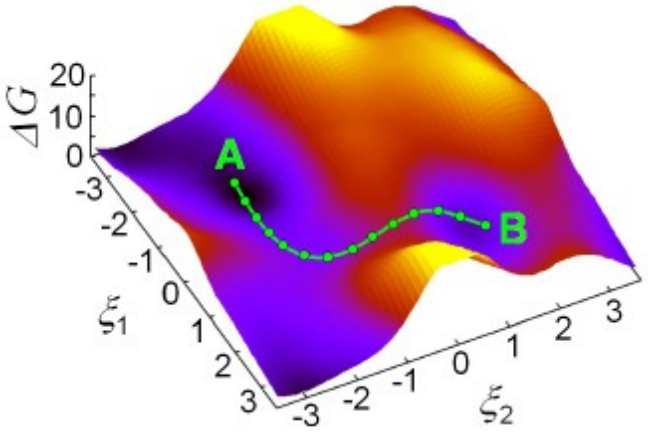
- **Potential of Mean Force (PMF)**
  - Concept of the reaction coordinate, or order parameter (ξ) which is used to distinguish between thermodynamic states
  - Often ξ is defined on geometric grounds: distance, (pseudo-) torsion angle, RMSD etc.
  - But ξ can also be more exotic quantity such as principal components or normal modes of the molecule

$$Q(\xi) = \frac{\int \delta[\xi(r) - \xi] \ \exp[(-\beta E) d^N r]}{\int \exp[(-\beta E) d^N r]}$$

$$A(\xi) = -1/\beta \ \ln \ Q(\xi)$$

  - **Potential of mean force:** the rest of degrees of freedom are effectively integrated out

  - PMF can be 1D or of higher dimensions



MD Simulations: Free Energy

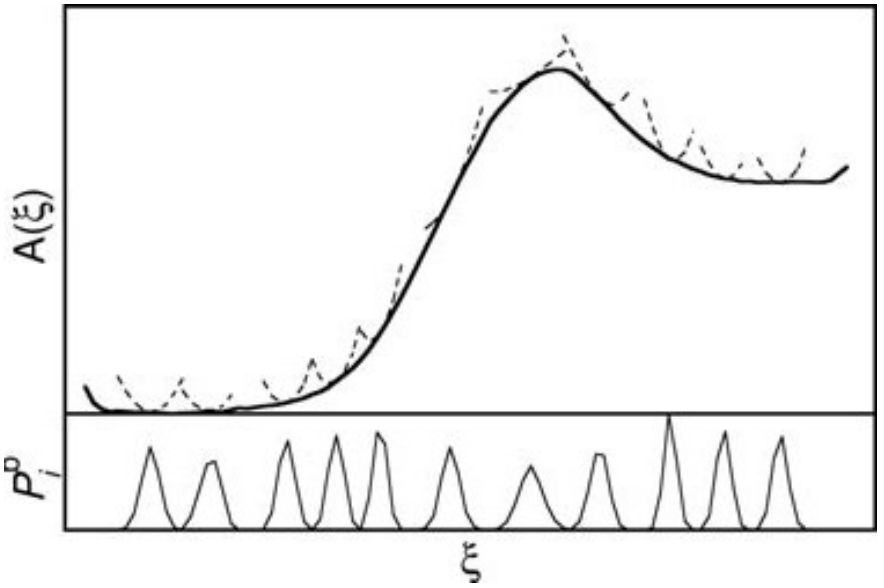
- **Means of effective sampling rare events (events of interest):**

  - Constrained MD
  - Non-equilibrium MD
  - Adaptive umbrella sampling
  - Local elevation/flooding
  - Adaptive biasing force
  - Metadynamics
  - Adaptively biased MD
  - Replica exchange MD
  - ……

Umbrella Sampling

- A bias, an additional energy term, is applied to the system to ensure efficient sampling along the whole reaction coordinate.

$$E^{b}(r) = E^{u}(r) + w_i(\xi) \qquad \omega_i(\xi) = K/2\left(\xi - \xi_i^{ref}\right)^2$$

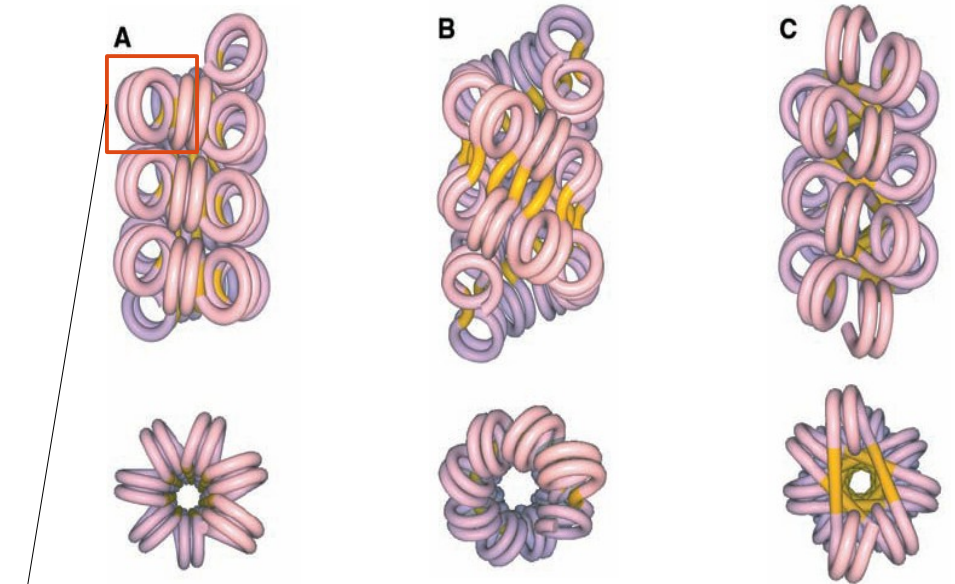


$$A_i(\xi) = -(1/\beta)\ \ln P_i^{b}(\xi) - w_i(\xi) + F_i$$

MD Simulations: Free Energy

- **Umbrella Sampling**
  - All-atom MD simulations of approaching 2 DNAs in NaCl and KCl buffers

Models *(in vitro)* for the DNA path in the chromatin fiber



one-start solenoidal: bent linker DNA

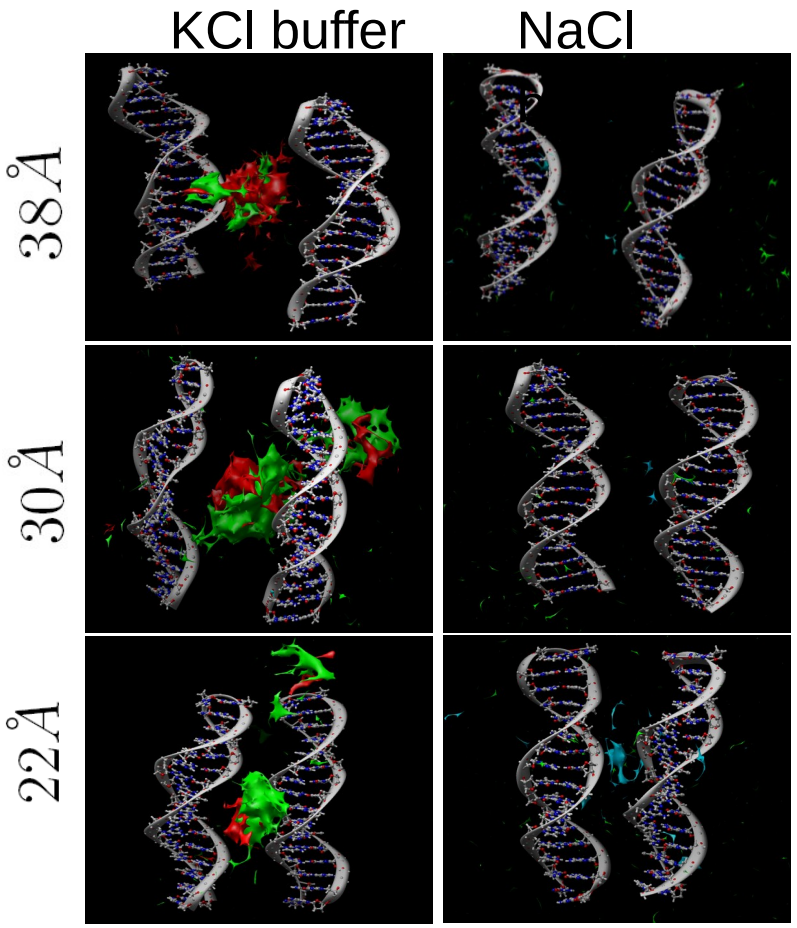
two-start supercoiled: straight linker DNA

two-start twisted: straight linker DNA

Nucleosome

~11 nm

B. Dorigo et al, Science 306, 1571 (2004)

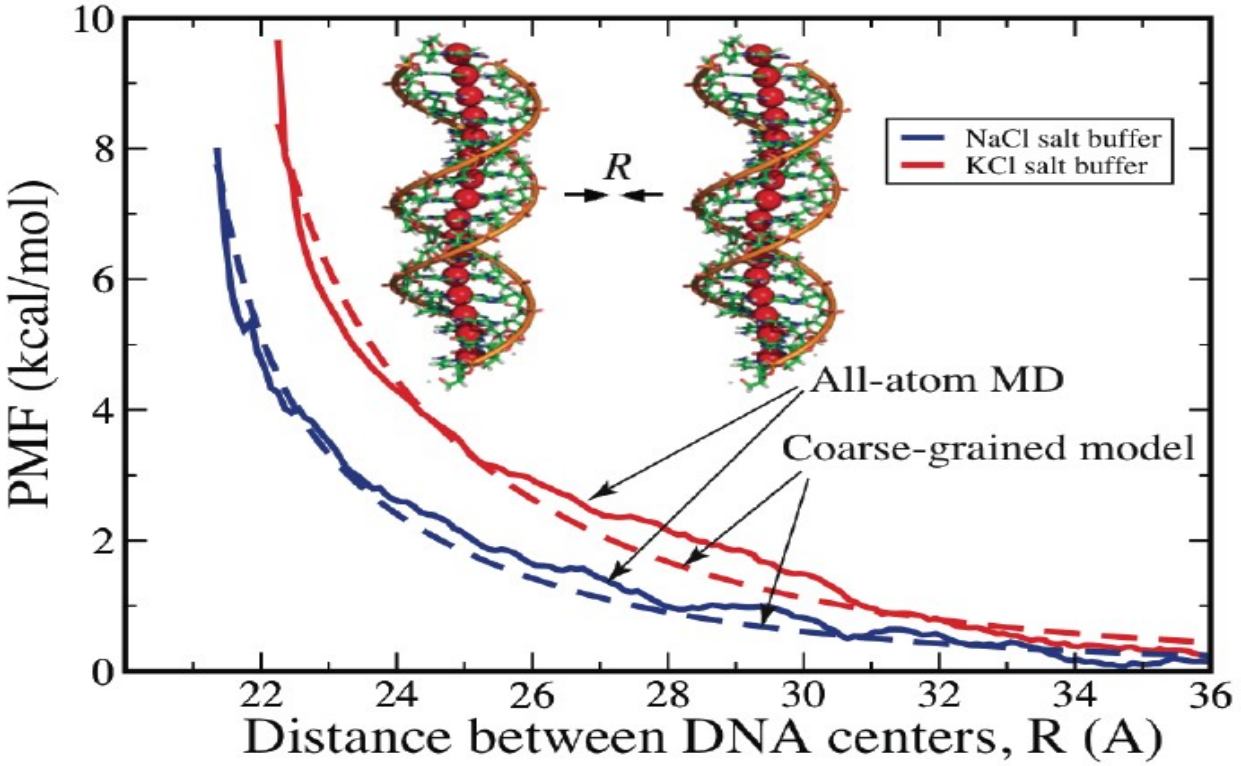


MD simulations of 2 DNA oligomers approaching in-paraller in different ionic buffers (NaCl & KCl)

Savelyev A. and Papoian G. JACS (2007) 129

MD Simulations: Free Energy

- WHAM is used for reconstruction of the PMF from AA MD



- ***Distance R*** *is broken into "windows"*
- *Biasing harmonic potential is applied to each window to keep 2 DNA segments at certain distance and in-parallel orientation*

Savelyev A. and Papoian G. JACS (2007) 129

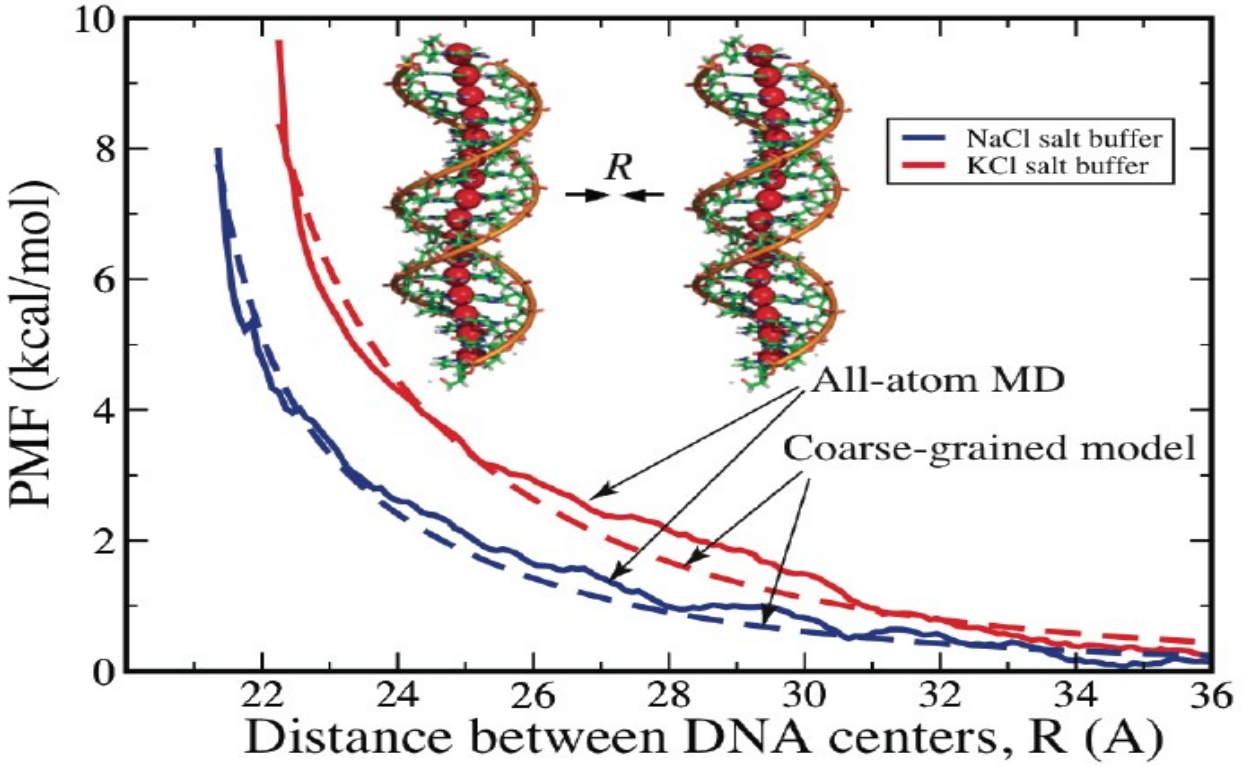
- ***WHAM*** *– weighted histogram analysis method is used to combine results from many simulations corresponding to different "windows" to compute the resulting free energy profile*
  - *MD at each window generates biased P'(R_i)*
  - *WHAM takes care of proper re-weighting of all P'(r) to generate unbiased P(r), or resulting free energy profile*
  - *It is also called Potential of mean force as the rest of degrees of freedom are effectively "integrated out"*

Developed:
Torrie, Valleau 1974,1977

Implemented in MD:
Alan Grossfield, 2003

MD Simulations: Free Energy

- WHAM is used for reconstruction of the PMF from AA MD



- *Fit of the AA PMF*

Effective (CG) potential

$$U_{el} = A\frac{e^{-\kappa r}}{r^4} + B\frac{e^{-\kappa(r-a)}}{r(1+a\kappa)}$$

a – DNA bead size (can be adjustable);

k_D – Debye length;

$$\mathcal{F}|_R = \sum_{i\neq j} U_{el} = A\sum_{i\neq j}\frac{e^{-\kappa r}}{r^4} + B\sum_{i\neq j}\frac{e^{-\kappa(r-a)}}{r(1+a\kappa)}$$

$$S_A = \sum_{i\neq j}\frac{e^{-\kappa r}}{r^4}, S_B = \sum_{i\neq j}\frac{e^{-\kappa(r-a)}}{r(1+a\kappa)}$$

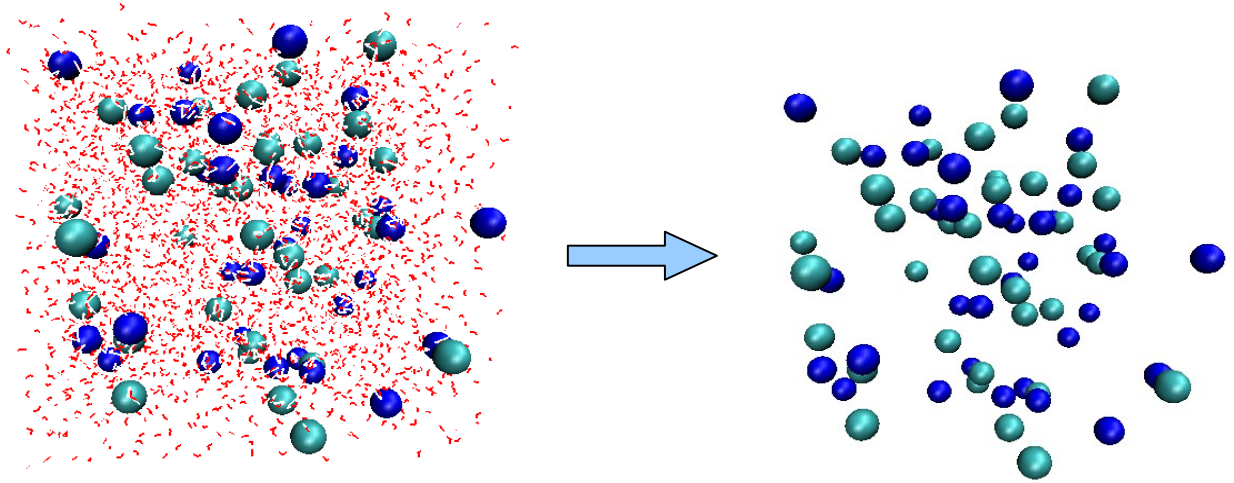
$$\mathcal{F}|_R = AS_A + BS_B$$

Linear in A, B !!!

Savelyev A. and Papoian G. JACS (2007) 129

MD Simulations: Free Energy

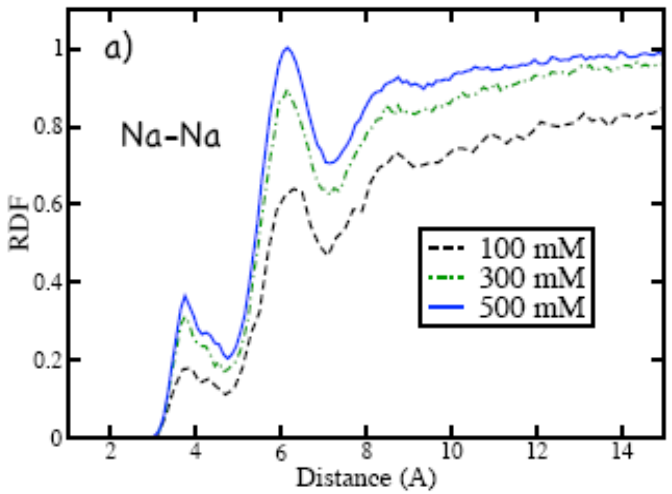
- **Potential of Mean Force (PMF): example of coarse-graining NaCl solutions**

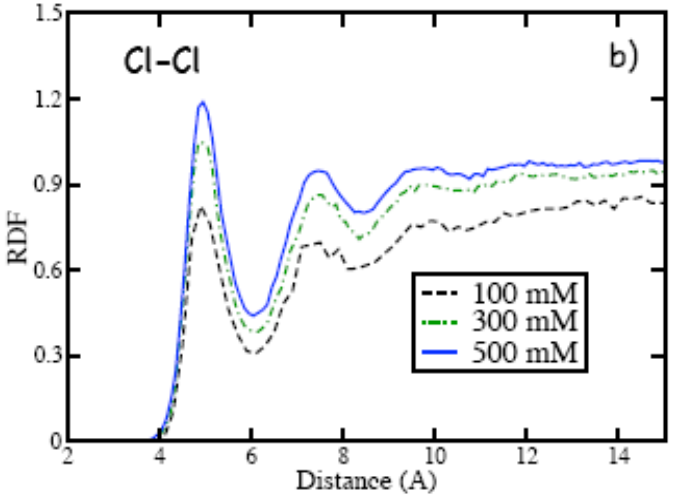


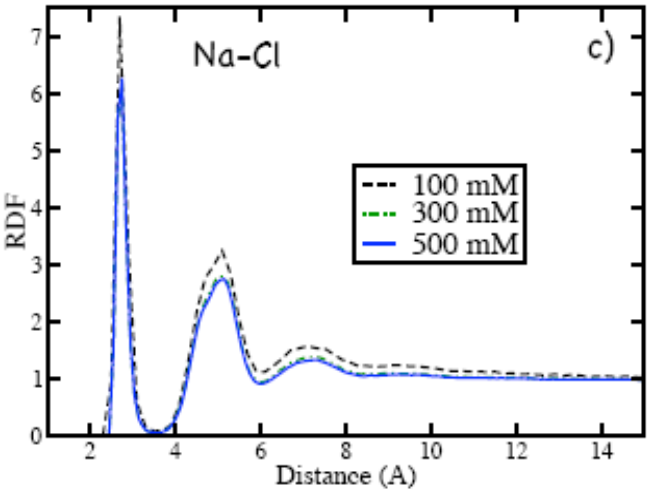
$g(r)$ – RDF

$$n(r) = \rho \int_0^r g(r) J(r)\, \mathrm{d}r$$

$$J(r)\, \mathrm{d}r = 4r\pi^2\, \mathrm{d}r$$





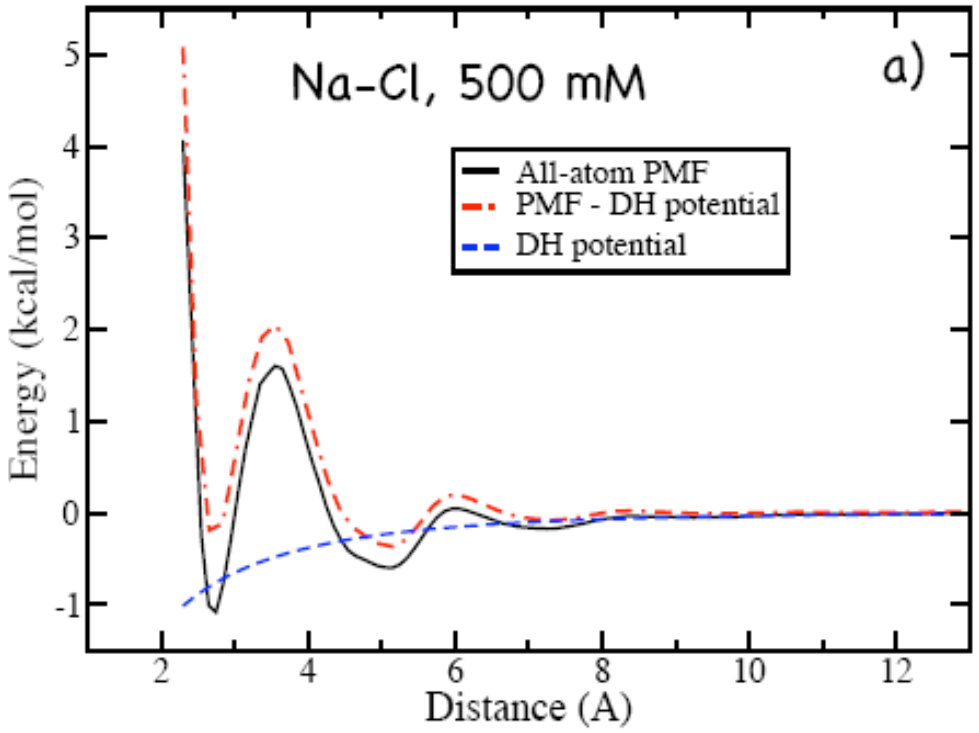


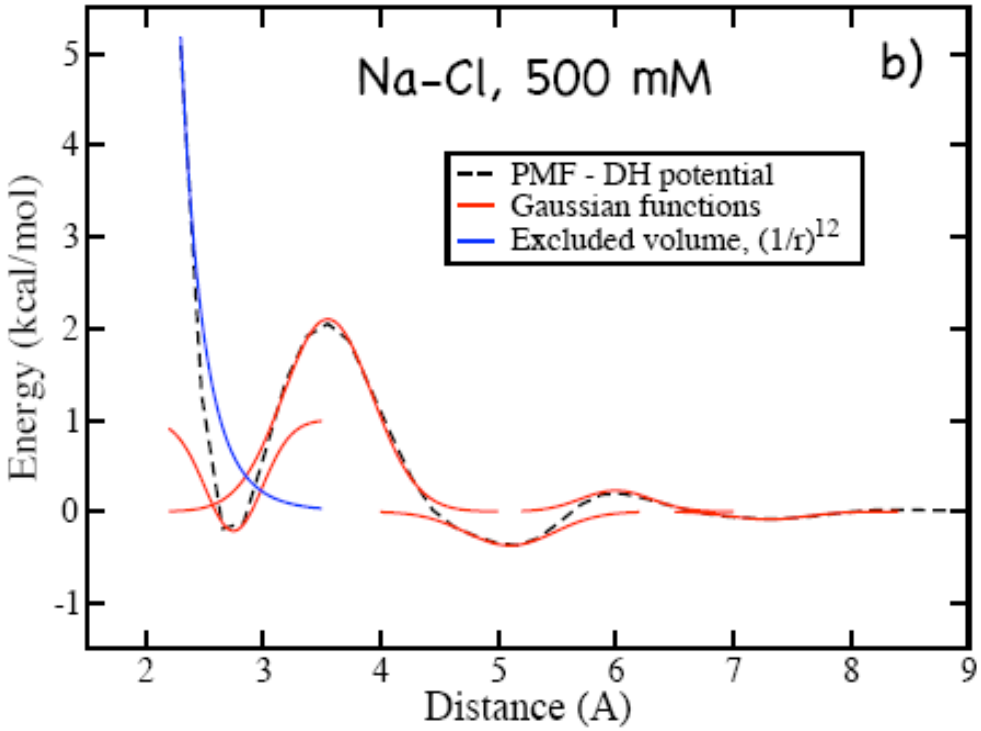
$$U_{\mathrm{PMF}} = -k_{\mathrm{B}} T \ln g(r)$$

Unbiased PMF

MD Simulations: Free Energy

- **Potential of Mean Force (PMF): example of coarse-graining NaCl solutions**





$$\mathscr{H} = \sum_{i>j} \left[\frac{A}{r_{ij}^{12}} + \sum_{k=1}^{5} B^{(k)} \mathrm{e}^{-C^{(k)}\left[r_{ij}-R^{(k)}\right]^2} + \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon r_{ij}} \right]$$

Savelyev A. and Papoian G. J Phys Chem B, 2009

MD Simulations: Free Energy

- **Potential of Mean Force (PMF): generalization**

  - Cartesian order parameter: conventional form

$$Q(\xi) = \frac{\int \delta[\xi(r) - \xi] \ \exp[(-\beta E) d^N r]}{\int \exp[(-\beta E) d^N r]}$$

  - Non-Cartesian order parameter: different set of coordinates $\boldsymbol{q}$ $\{\xi_i\}$

$$Q(\xi) = \int \exp(-\beta E) |J(q)| d^{N-1} q / Q$$

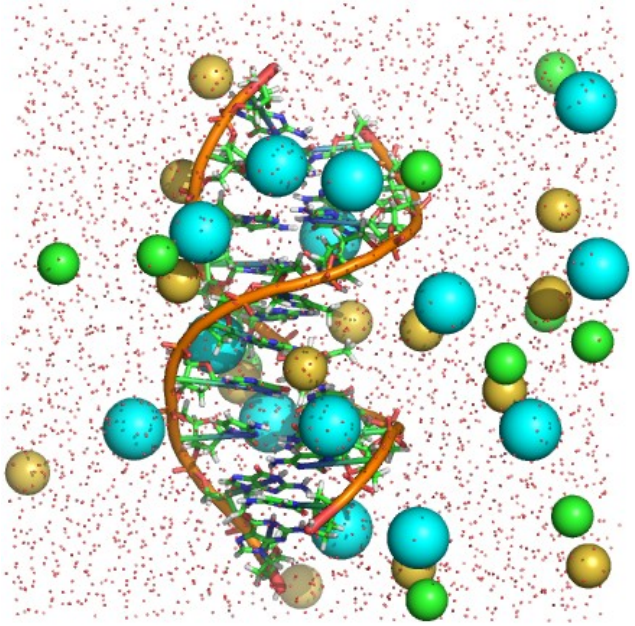
$$J_{ij} = \mathrm{d}q_i / \mathrm{d}r_j \qquad |J(q)|$$

    - Jacobian determinant

    - Integrate over all generalized coordinates but ξ

MD Simulations: Free Energy

- **Potential of Mean Force (PMF): generalization example, PCA**



CHARMM C36, Drude and AMBER FFs

- 1DCV (10 b.p.) [B form]
- Ecor1 (12 b.p.) [B form]

AMBER setup:

- Parmbsc0 FF for DNA;
- TIP3P water model;
- Cheatham&Joung monovalent ion parameters for Ewald and TIP3P water

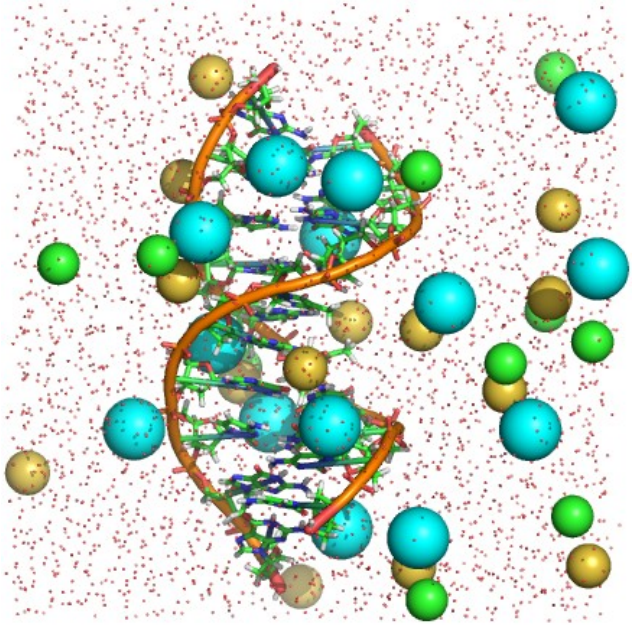
MD simulation protocol:

NPT ensemble;
P = 1 atm;
T = 300K; (additional thermostat at 1K for Drude particles)
Δt = 2 fs (non-polarizable); 1 fs (Drude polarizable)
Particle mesh Ewald summation for long-range electrostatics
Periodic boundary conditions (*l* ~ 50 Å)
atoms: ~15,000 / ~22,000 (Additive/Drude polarizable)

Ensemble: ~10^5 frames
(every 6ps: ~600 ns)

MD Simulations: Free Energy

- **Potential of Mean Force (PMF): generalization example, PCA**



Ensemble: ~10^5 frames
(every 6ps: ~600 ns)

- How do we uncover distinct conformational modes of DNA oligomer?
- How strength/content of ionic buffer affect these conformational modes ?

Savelyev & MacKerell, J.Phys. Chem. Lett. 6, 212, 2014

- **Numerical SAXS DNA profiles**

$$I(\mathbf{q}) = \sum_j^N \sum_k^N A_j A_k e^{i\mathbf{q}\cdot\mathbf{r}_{j,k}}$$

$$I(q) = \langle I(\mathbf{q}) \rangle_\Omega = \sum_j^N \sum_k^N A_j A_k \frac{\sin qr_{j,k}}{qr_{j,k}}$$

$$A_j = f_j(q) - \rho_0 g_j(q)$$

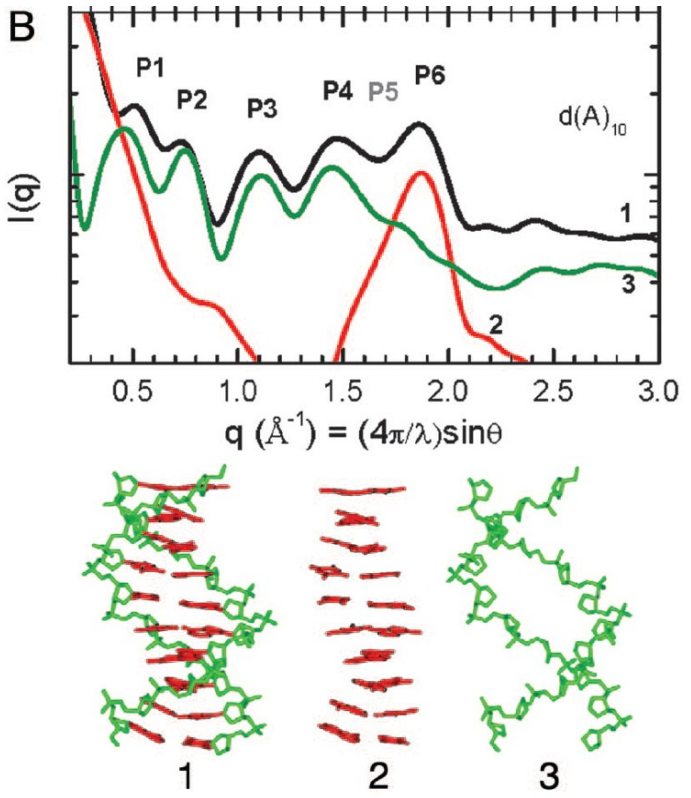
$$g_j(q) = G(q) V_j e^{-q^2 V_j^{2/3}/4\pi}$$

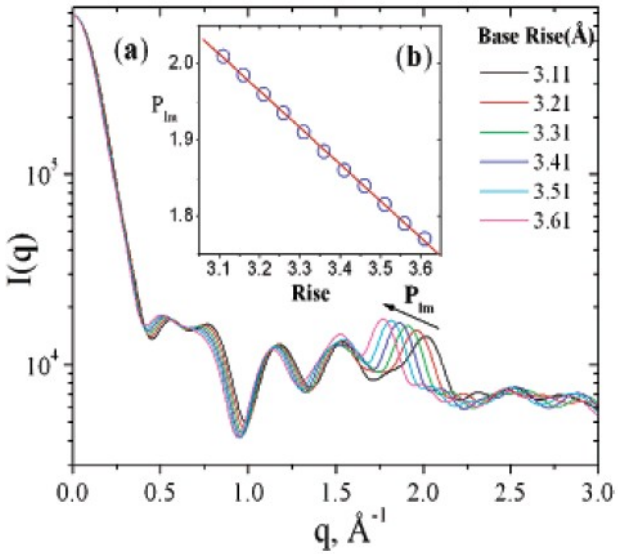
$$G(q) = \frac{V_o}{V_m} e^{-q^2 (V_o^{2/3} - V_m^{2/3})/4\pi}$$

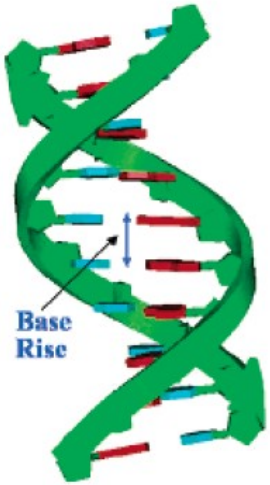
0< q < 25 A^{-1}

$$f(q) = \sum_{i=1}^{4} a_i \exp\left(-b_i \left(\frac{q}{4\pi}\right)^2\right) + c,$$

(NIST web site)

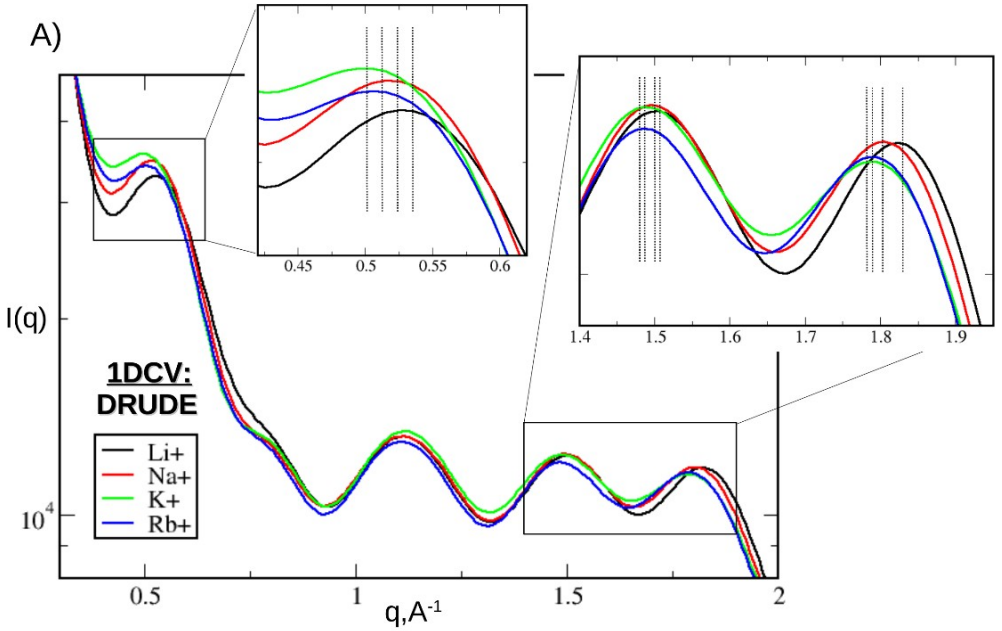






Tiede et al, JACS, 127, 16, 2005

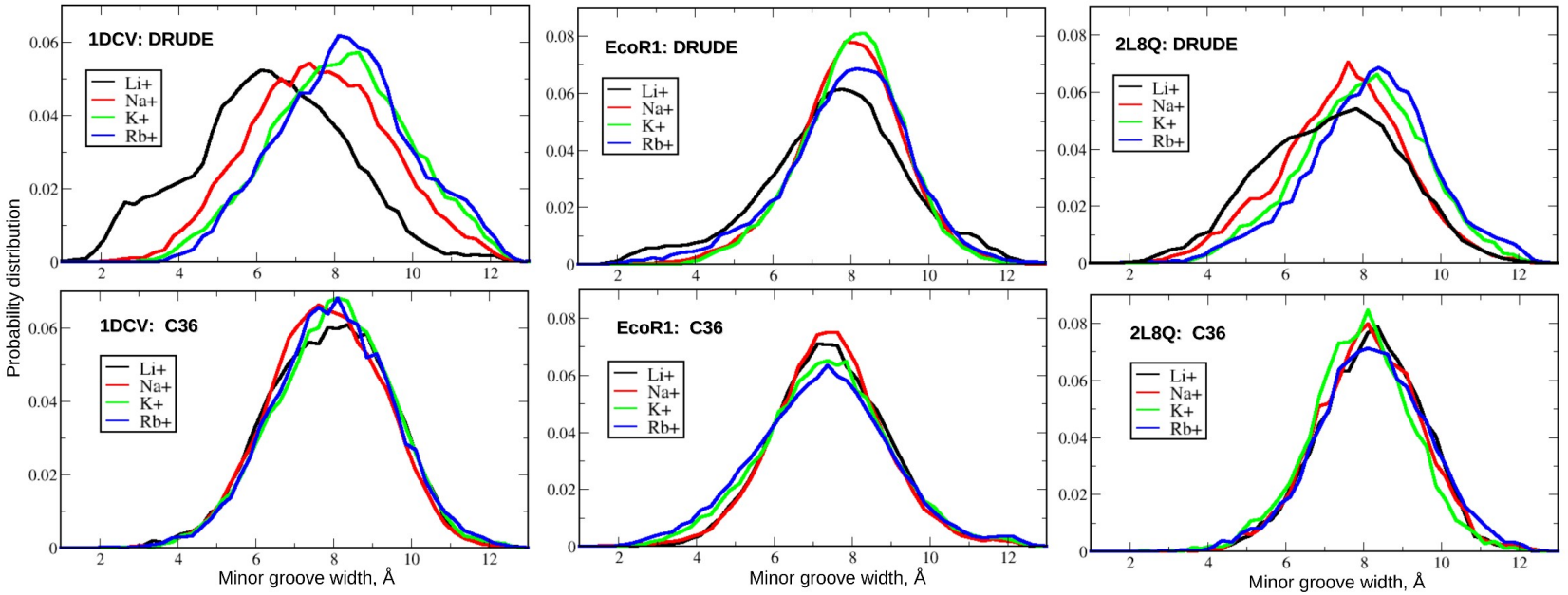
- **Numerical SAXS DNA profiles**



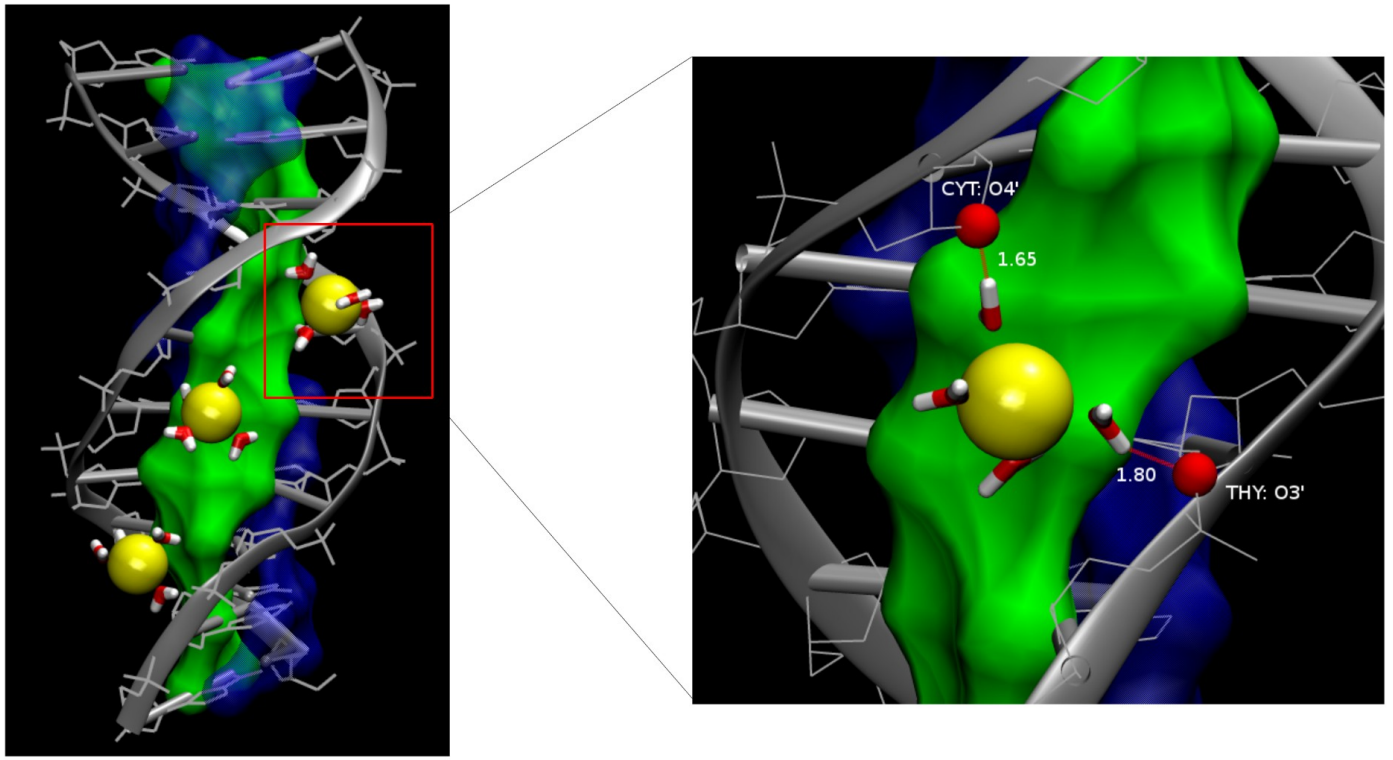
- **DNA minor groove** width correlates the most with scattering profiles changes as function of ion type

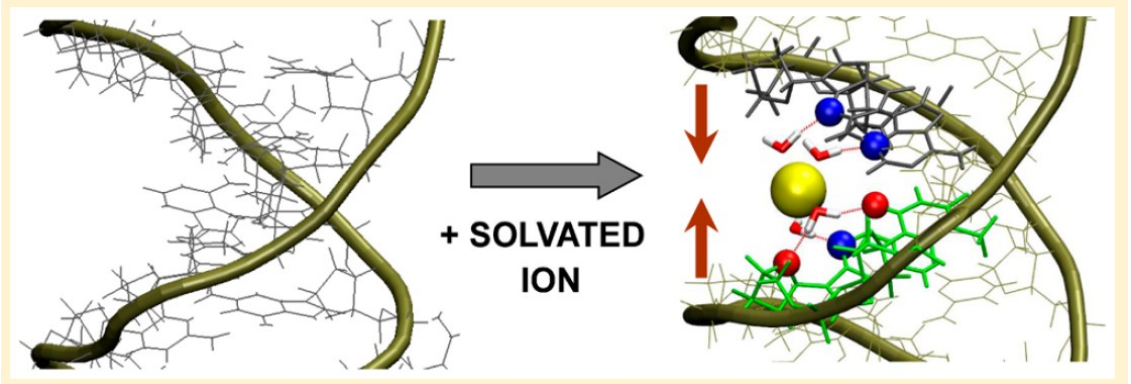
(among numerous other DNA geometric parameters: roll, twist, base-pait rise etc.)

Savelyev & MacKerell, J.Phys. Chem. Lett. 6, 212, 2014



Hydrated cations modulate minor groove via hydrogen bond formation between ion's water and DNA strands





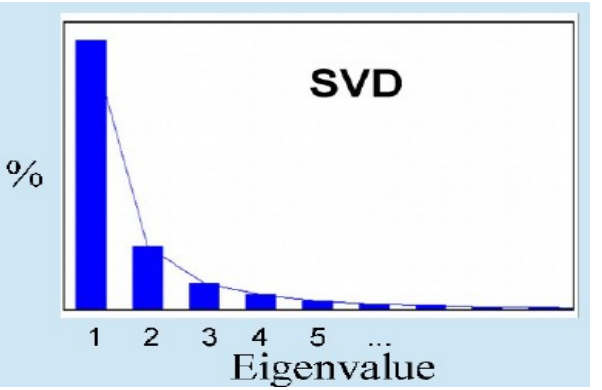
Savelyev & Mackerell, J. Chem. Theory Comput., 11, 4473, 2015

MD Simulations: Free Energy

- Construct covariance matrix in dihedral-angle phase space

$$C_{ij} = \left\langle (x_i(t) - \bar{x}_i)(x_j(t) - \bar{x}_j) \right\rangle$$

$$x_i = \{\cos\varphi_i,\ \sin\varphi_i\}, \quad \varphi = \{\alpha, \beta, \gamma, \delta, \varepsilon, \zeta, \chi_{PUR}, \chi_{PYR}\}$$



- Diagonalize covariance matrix

$$\mathbf{C} = \mathbf{T}\Lambda\mathbf{T}^T \qquad \{\lambda_{ii}\}$$

$$\{\mathbf{t}^{(i)}\}$$

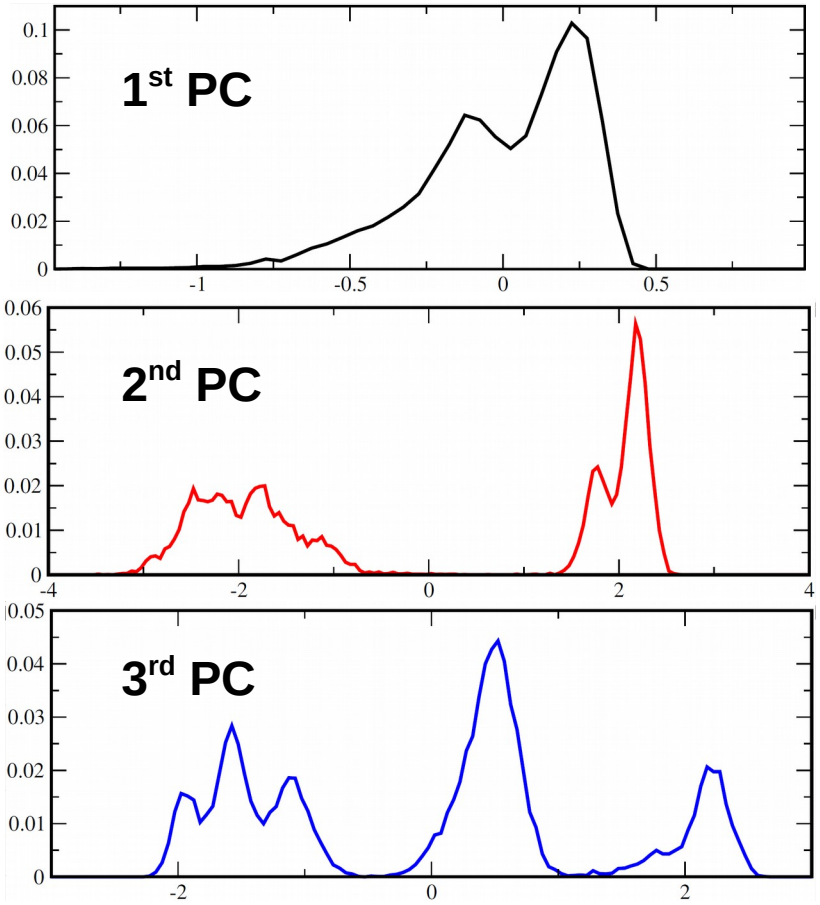
- Project entire MD trajectory onto few largest PCs

$$P_i(t) = \mathbf{t}^{(i)} \cdot (\mathbf{x}(t) - \bar{\mathbf{x}})$$

- Build 1D and 2D free energy landscapes

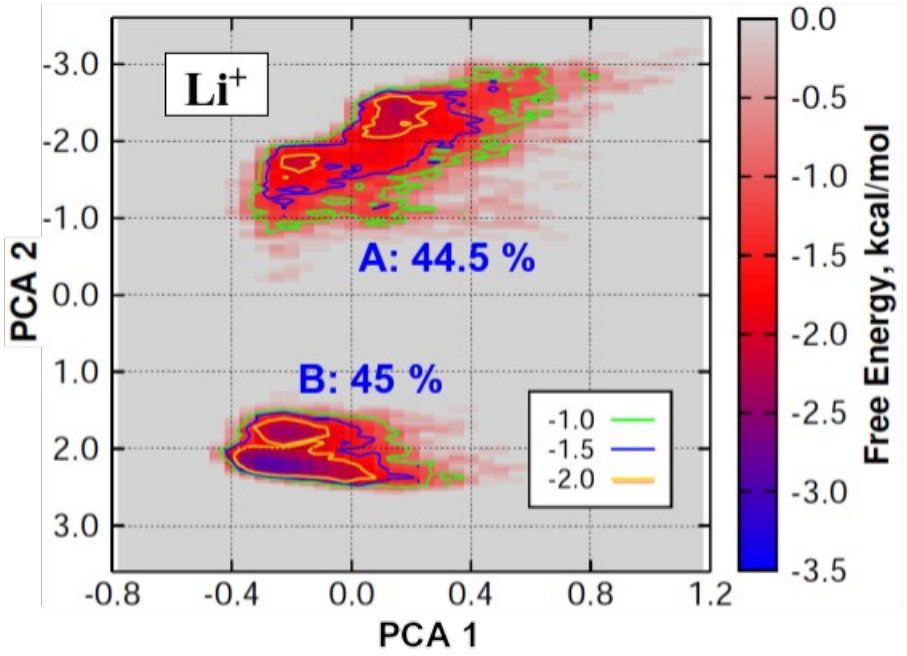
$$F(P_1) = -k_B T \ln\left[\rho_1(P_1)\right]$$

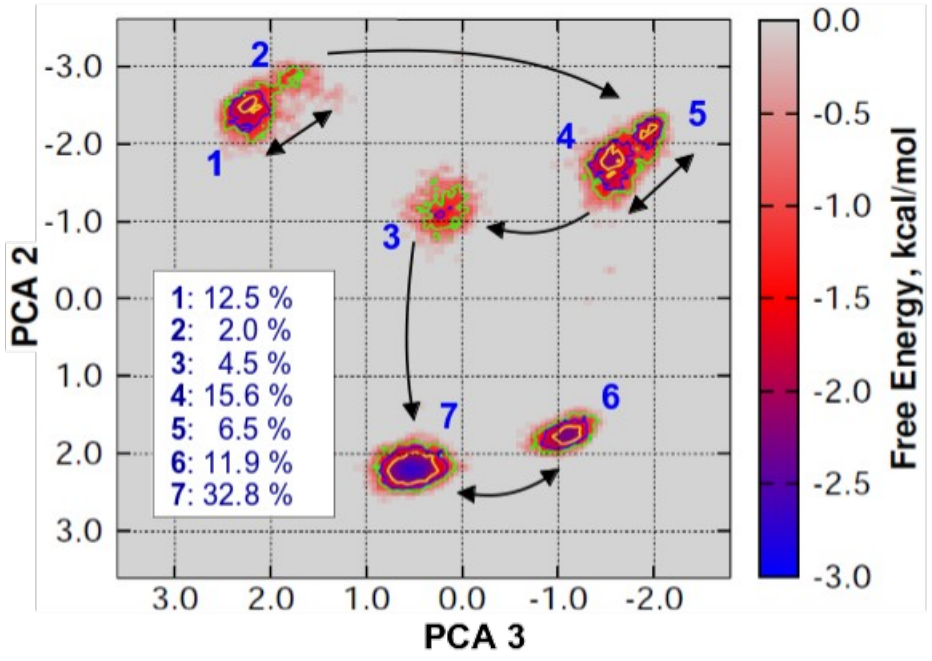
$$F(P_1, P_2) = -k_B T \ln\left[\rho_{12}(P_1, P_2)\right]$$

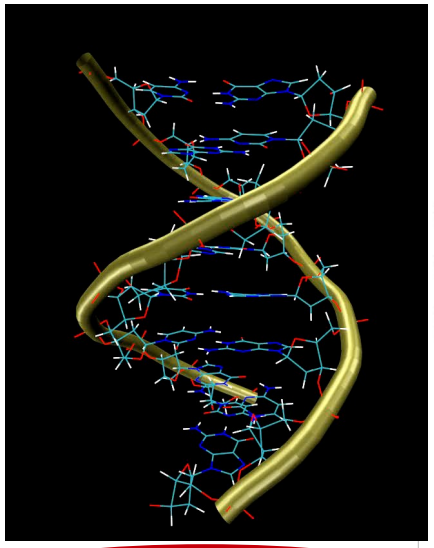


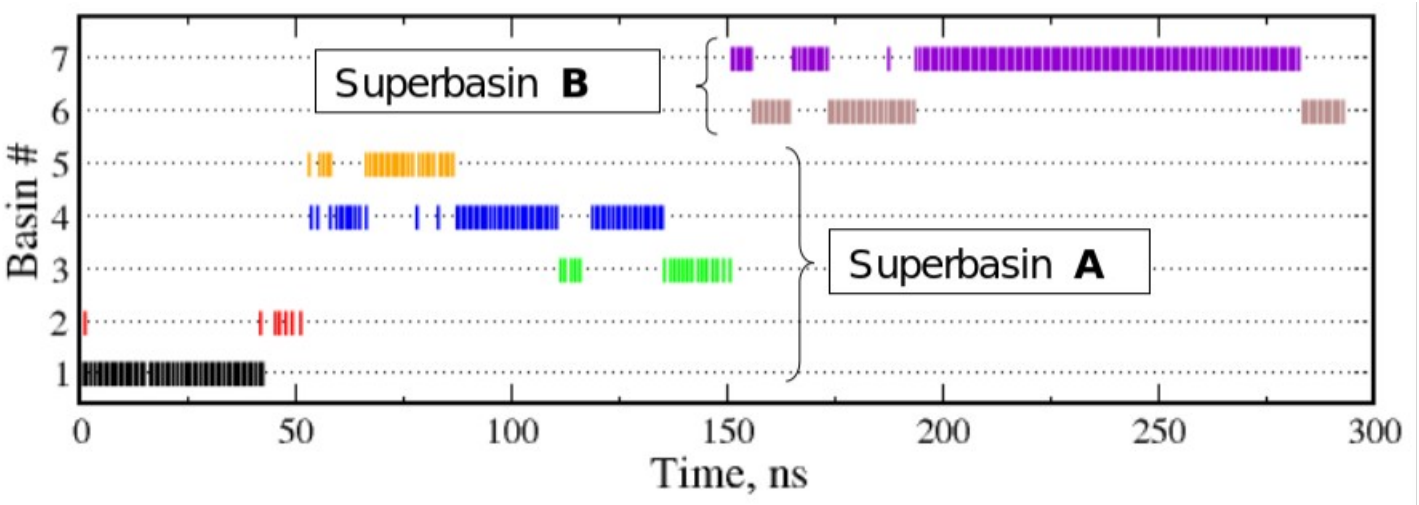
Savelyev A. *in preparation*

1DCV in LiCl [DRUDE]: Dynamics Inferred from dPCA





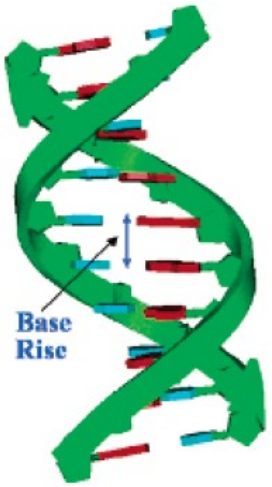


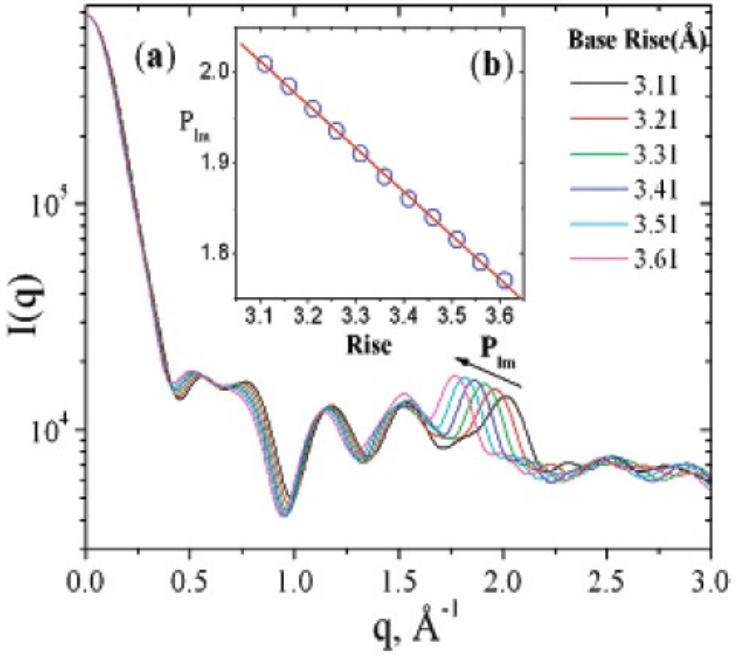


Tens of thousands of frames reduced to ~10 representative structural modes !!!

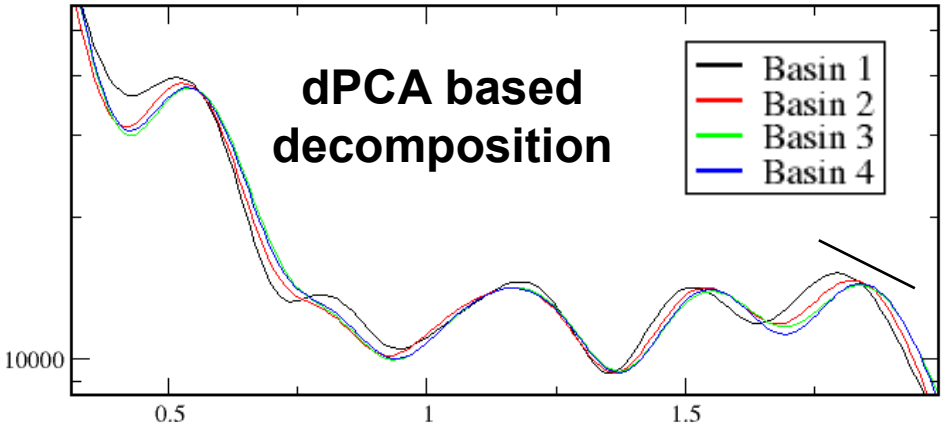
Savelyev A. *in preparation*

Use of dPCA to Characterize SAXS DNA Data





Tiede et al, JACS, 127, 16, 2005

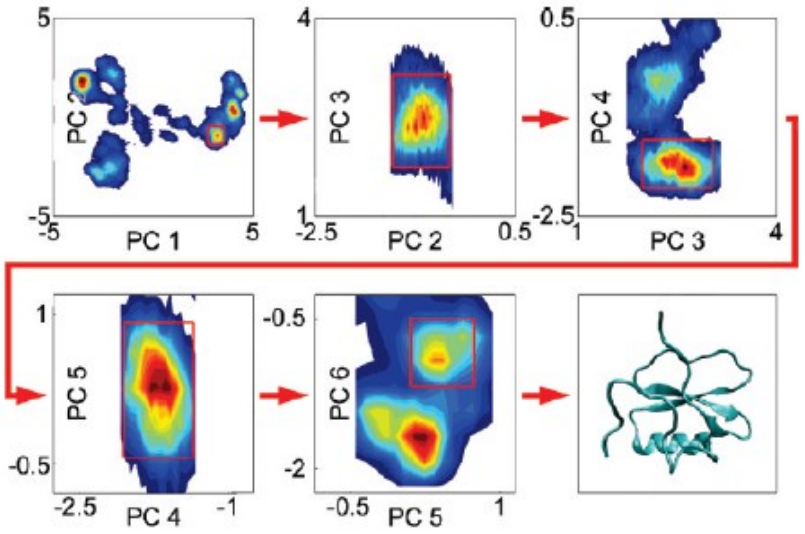


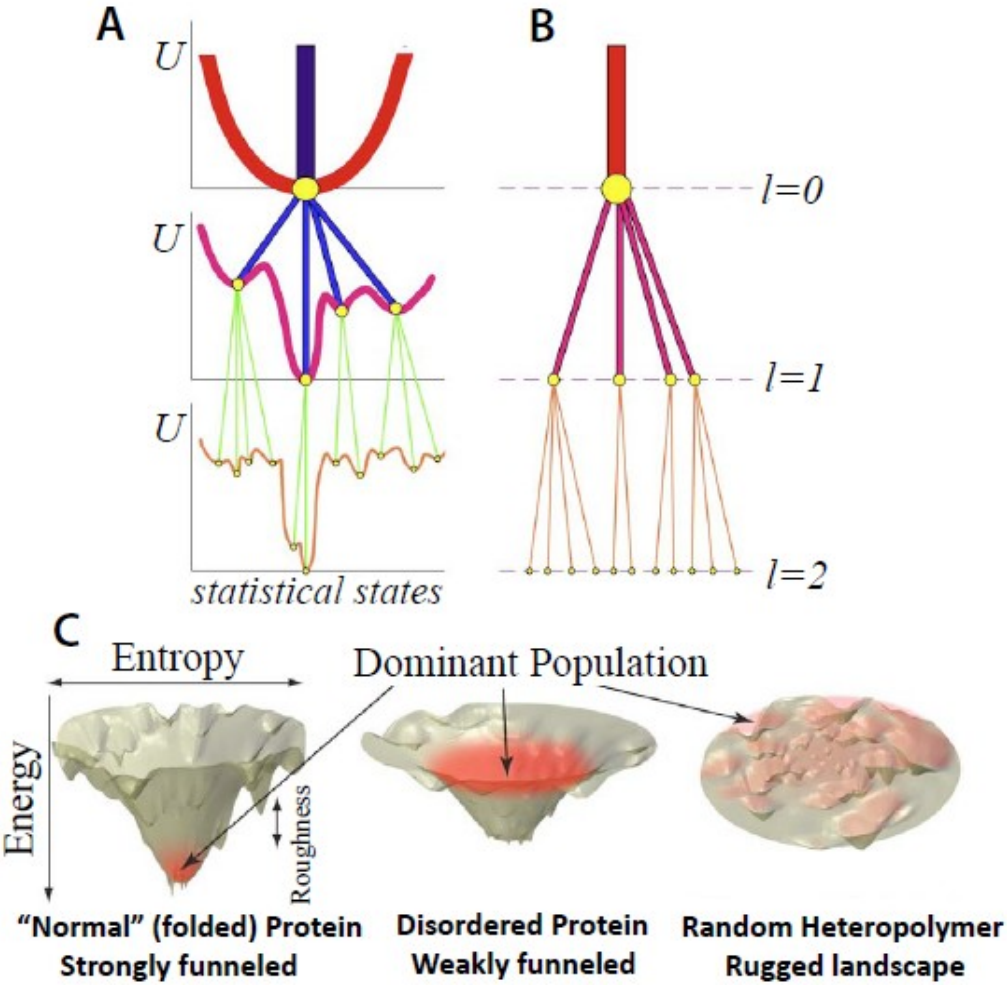
Savelyev A. *in preparation*

Other PCA applications: Essential Protein Dynamics

- Use of dPCA to describe proteins:

how sub-basins are split into smaller basins in a hierarchically constrained fashion





Zhuravlev et al, J Phys Chem B 113, 26, 2009

Other PCA applications: Polymorphism of G-Quadruplexes

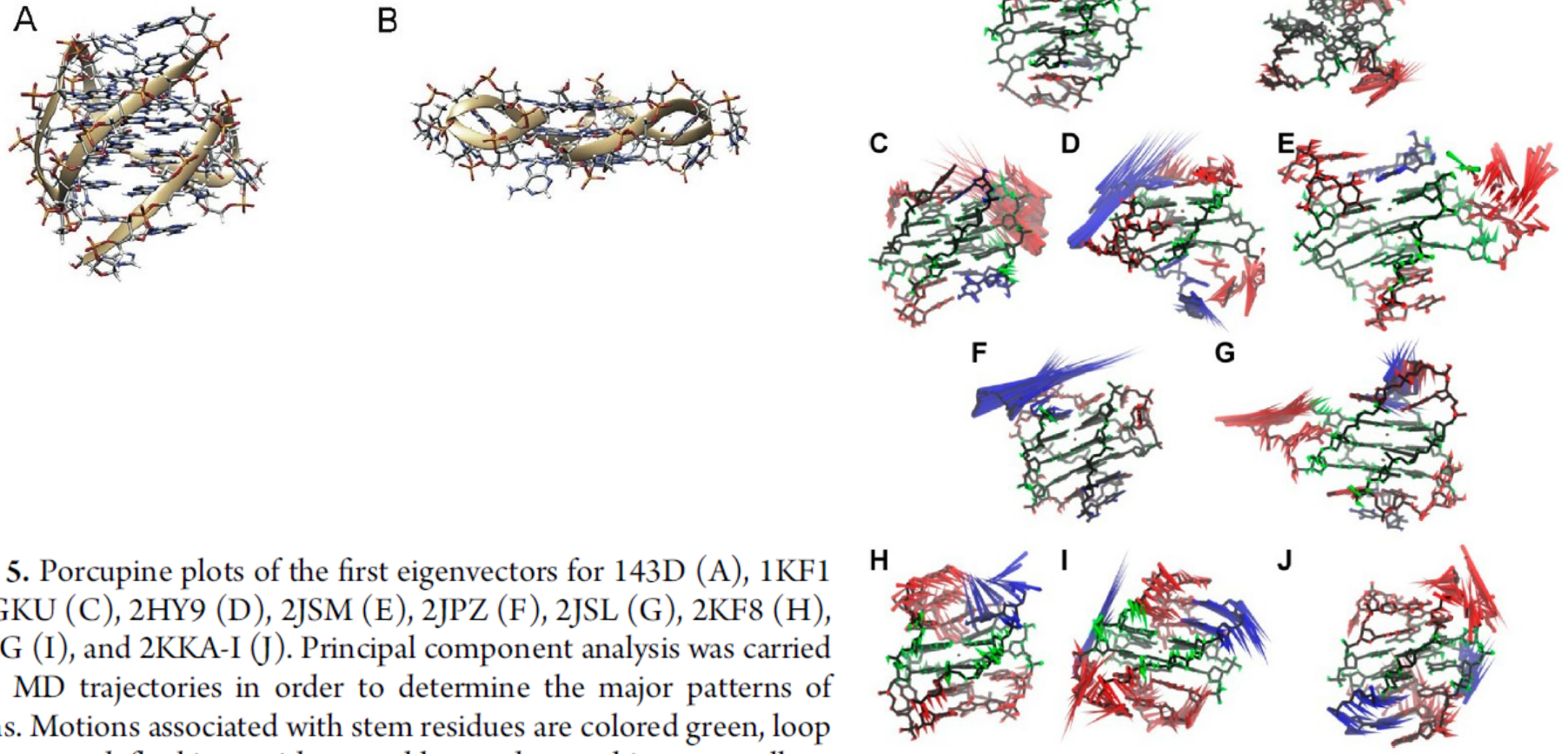


Figure 5. Porcupine plots of the first eigenvectors for 143D (A), 1KF1 (B), 2GKU (C), 2HY9 (D), 2JSM (E), 2JPZ (F), 2JSL (G), 2KF8 (H), 2KKA-G (I), and 2KKA-I (J). Principal component analysis was carried out on MD trajectories in order to determine the major patterns of motions. Motions associated with stem residues are colored green, loop residues are red, flanking residues are blue, and central ions are yellow.

Huy T. Le,William L. Dean,Robert Buscaglia, Jonathan B. Chaires, and John O. Trent, J Phys Chem B, 2014