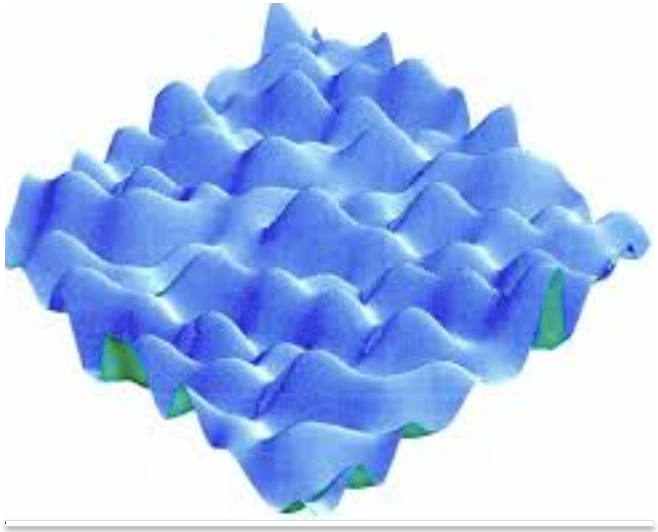
Topics in Enhanced Sampling

R-RESPA and FEP/REST



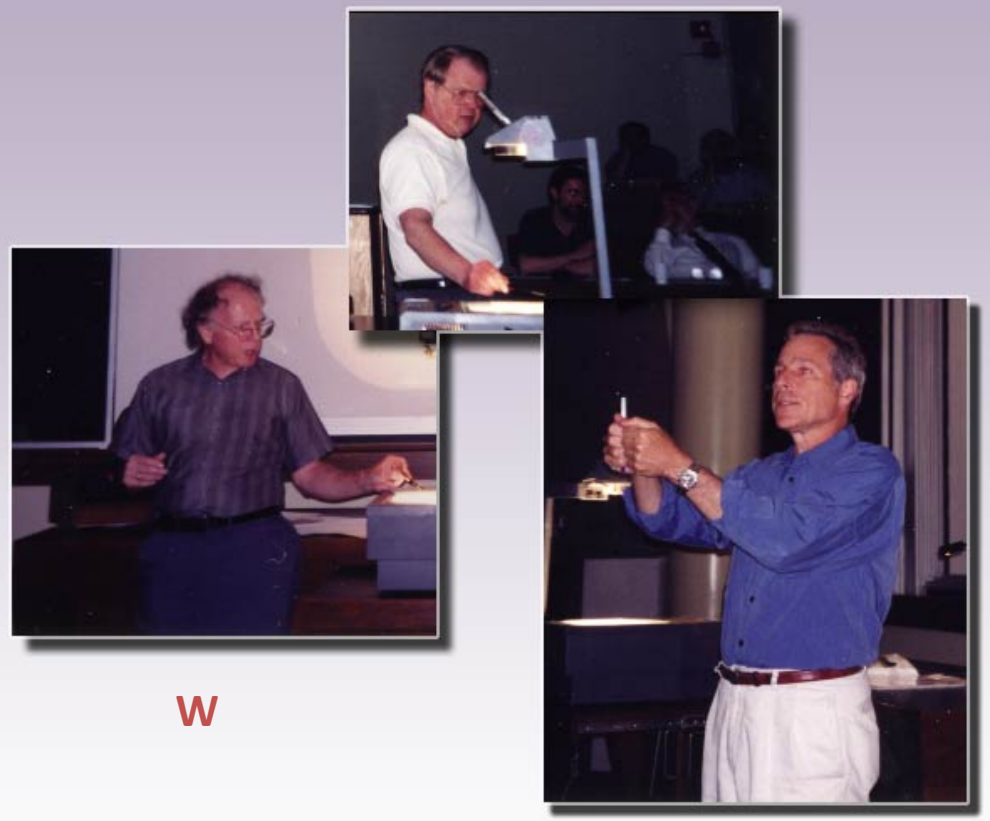
M. E. Tuckerman, G. J. Martyna and B. J. Berne, Reversible multiple time scale molecular dynamics, *J. Chem. Phys. 97, 1990-2001 (1992).*

Ruhong Zhou, Edward Harder, Huafeng Xu and B. J. Berne, Efficient multiple time step method for use with Ewald and particle mesh Ewald for large biomolecular systems, *J. Chem. Phys. 115, 2348-2358 (2001).*

J. Morrone, Ruhong Zhou and B. J. BerneMolecular dynamics with multiple time scales: How to avoid pitfalls, J. Theoretical and Computational Chemistry, 6, 1798-1804, (2010)

WCA
Congratulations to W & C

A



W

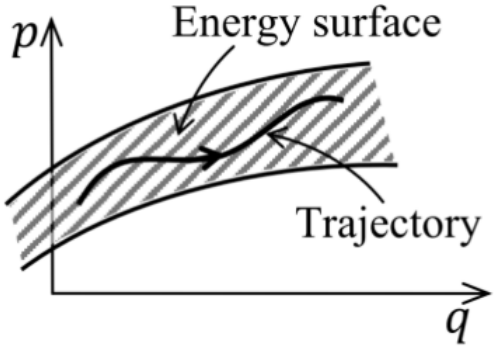
C

Liouville Propagator

Instantaneous state $\mathbf{\Gamma} = (q_1, q_2, \ldots, q_f, p_1, p_2, \ldots, p_f)$

The Liouvillian: $iL \equiv \{\ldots, H\} = \dot{\mathbf{\Gamma}} \cdot \nabla_{\mathbf{\Gamma}}(\ldots) = \nabla_{\mathbf{\Gamma}} \cdot (\dot{\mathbf{\Gamma}} \ldots)$

Evolution of mechanical state: $\mathbf{\Gamma}(t) = e^{iL(t-t_0)}\mathbf{\Gamma}(t_0)$



Example: $iL = \dot{x}\partial/\partial x + F(x)\partial/\partial p$

Propagator: $e^{iLt} = e^{t(\dot{x}\partial/\partial x + F(x)\partial/\partial p)}$

Some Properties of Liouville Operator

$$L = -i\{\ldots, H\} \text{ is Hermitian}$$

$$e^{iLt} \text{ is Unitary}$$

Time Reversible

$$e^{-Lt}e^{iLt} = e^{iLt}e^{-iLt} = \hat{I}$$

Liouville Theorem

$$\frac{df(\mathbf{\Gamma}, t)}{dt} = \frac{\partial f(\mathbf{\Gamma}, t)}{\partial t} + iLf(\mathbf{\Gamma}, t) = 0$$

$$f(\mathbf{\Gamma}, t) = f(\mathbf{\Gamma}, 0)$$

$$\delta\mathbf{\Gamma}_t = \delta\mathbf{\Gamma}_0$$

Reversible and Symplectic Integrators

$$iL = \dot{x}\partial/\partial x + F(x)\partial/\partial p = iL_1 + iL_2$$

Trotter Factorization of Propagator:

$$e^{i(L_1+L_2)t} = (e^{iL_1 t/2P} e^{iL_2 t/P} e^{iL_1 t/2P})^P + O(t^3/P^2)$$

Single Time Step Propagator Propagator: $\Delta t = t/P$

$$e^{i(L_1+L_2)\Delta t} = e^{iL_1 \Delta t/2} e^{iL_2 \Delta t} e^{iL_1 \Delta t/2} + O(\Delta t^3)$$

$$\begin{bmatrix} x(\Delta t) \\ p(\Delta t) \end{bmatrix} = e^{iL_1 \Delta t/2} e^{iL_2 \Delta t} e^{iL_1 \Delta t/2} \begin{bmatrix} x(0) \\ p(0) \end{bmatrix}$$

Time Reversibility:

$$e^{-iL_1 \Delta t/2} e^{-iL_2 \Delta t} e^{-iL_1 \Delta t/2} \cdot e^{iL_1 \Delta t/2} e^{iL_2 \Delta t} e^{iL_1 \Delta t/2} = \hat{I}$$

Preserves measure in phase space.

A Time Reversible MD Integrator

$$iL = \dot{x}\partial/\partial x + F(x)\partial/\partial p = iL_1 + iL_2$$

$$\begin{bmatrix} x(\Delta t) \\ p(\Delta t) \end{bmatrix} = e^{\frac{\Delta t}{2} F \cdot \frac{\partial}{\partial p}} e^{\Delta t \dot{x} \cdot \frac{\partial}{\partial x}} e^{\frac{\Delta t}{2} F \cdot \frac{\partial}{\partial p}} \begin{bmatrix} x(0) \\ p(0) \end{bmatrix}$$

$$e^{a\partial/\partial y} f(y) = f(y + a)$$

$$\begin{bmatrix} x(\Delta t) \\ \dot{x}(\Delta t) \end{bmatrix} = \begin{bmatrix} x(0) + \Delta t \dot{x}(0) + \frac{\Delta t^2}{2m} F(x(0) \\ \dot{x}(0) + \frac{\Delta t}{2m} [F(x(0)) + F(x(\Delta t)] \end{bmatrix}$$

$$\begin{bmatrix} v \leftarrow v + F(x)\frac{\Delta t}{2m} \\ x \leftarrow x + v\Delta t \\ v \leftarrow v + F(x)\frac{\Delta t}{2m} \end{bmatrix}$$

This is the Velocity Verlet Integrator (Andersen)

Another Time Reversible MD Integrator

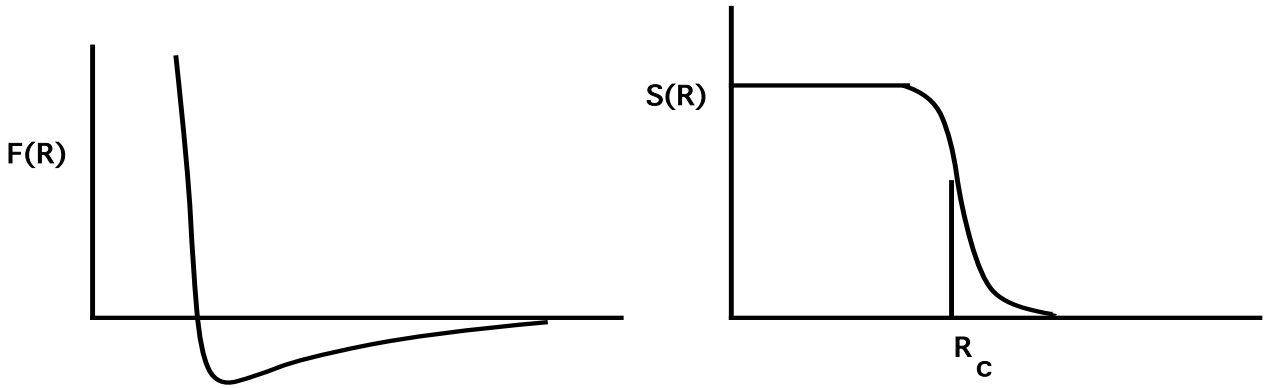
$$\begin{bmatrix} x(\Delta t) \\ p(\Delta t) \end{bmatrix} = e^{\frac{\Delta t}{2}\dot{x}\cdot\frac{\partial}{\partial x}} e^{\Delta t F\cdot\frac{\partial}{\partial p}} e^{\frac{\Delta t}{2}\dot{x}\cdot\frac{\partial}{\partial x}} \begin{bmatrix} x(0) \\ p(0) \end{bmatrix}$$

$$\begin{bmatrix} x(\Delta t) \\ \dot{x}(\Delta t) \end{bmatrix} = \begin{bmatrix} x(0) + \frac{\Delta t}{2}[\dot{x}(0) + \dot{x}(\Delta t)] \\ \dot{x}(0) + \frac{\Delta t}{m} F(x(0) + \frac{\Delta t}{2}\dot{x}(0)) \end{bmatrix}$$

$$\begin{bmatrix} x \leftarrow x + v\frac{\Delta t}{2} \\ v \leftarrow v + F\Delta t \\ x \leftarrow x + v\frac{\Delta t}{2} \end{bmatrix}$$

We call this the position Verlet integrator (Berne et al 1992)

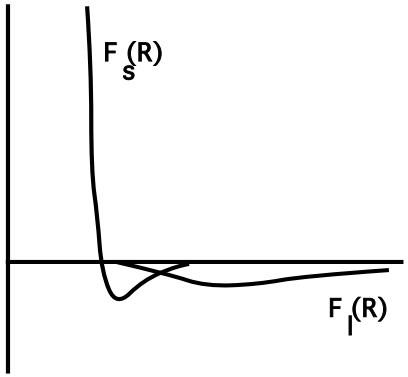
Long and Short Range Forces



$$F(R) = S(R)F(R) + [1 - S(R)]F(R)$$

$$F_s(R) = S(R)F(R)$$

$$F_l(R) = [1 - S(R)]F(R)$$



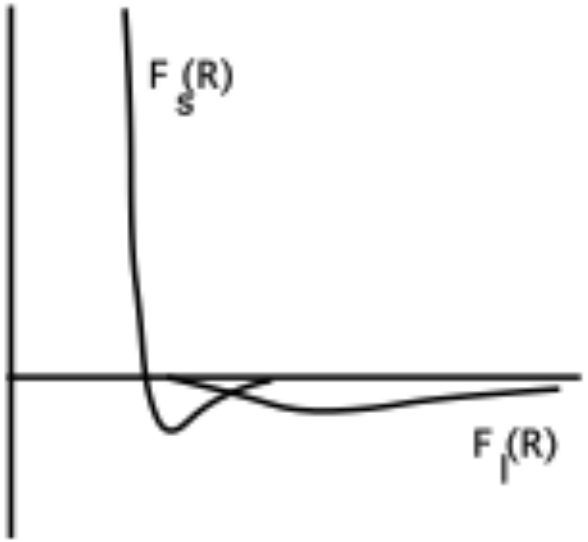
Reversible RESPA with long range force

Subdivision of Forces

$$F(x) = F_s(x) + F_l(x)$$

Subdivision of Liouvillian

$$iL = \dot{x}\frac{\partial}{\partial x} + F_s\frac{\partial}{\partial p} + F_l\frac{\partial}{\partial p} = iL_s + F_l\frac{\partial}{\partial p}$$



$$G_{sls}(\Delta t) = e^{\frac{\Delta t}{2}F_l\frac{\partial}{\partial p}}e^{\Delta t L_s}e^{\frac{\Delta t}{2}F_l\frac{\partial}{\partial p}}$$

where $\Delta t = n\delta t$

$$e^{iL_s\Delta t} = \left(e^{\frac{\delta t}{2}F_s\frac{\partial}{\partial p}}e^{\delta t\dot{x}\frac{\partial}{\partial x}}e^{\frac{\delta t}{2}F_s\frac{\partial}{\partial p}}\right)^n$$

Propagator

$$G_{sls}(\Delta t) = e^{\frac{\Delta t}{2}F_l\frac{\partial}{\partial p}}\left(e^{\frac{\delta t}{2}F_s\frac{\partial}{\partial p}}e^{\delta t\dot{x}\frac{\partial}{\partial x}}e^{\frac{\delta t}{2}F_s\frac{\partial}{\partial p}}\right)^n e^{\frac{\Delta t}{2}F_l\frac{\partial}{\partial p}}$$

Short and Long Range Force Breakup

$$
\begin{aligned}
& v \leftarrow v + F^{(l)} \frac{\Delta t}{2m} \\
& \text{do } i = 1, n \\
& \qquad v \leftarrow v + F^{(s)} \frac{\delta t}{2m} \\
& \qquad x \leftarrow x + v\, \delta t \\
& \qquad v \leftarrow v + F^{(s)} \frac{\delta t}{2m} \\
& \text{end do} \\
& v \leftarrow v + F^{(l)} \frac{\Delta t}{2m} .
\end{aligned}
$$

Ewald with RESPA

Separate the real-space and k-space parts ,

$$V_{ij}^{\rm el} = V_{ij}^{\rm rs} + V_{ij}^{\rm ks},$$

$$V_{ij}^{\rm rs} = (1-\delta_{ij})\, q_i q_j \frac{{\rm erfc}(\kappa r_{ij})}{r_{ij}}$$

$$V_{ij}^{\rm ks} = q_i q_j \left[\frac{1}{\pi L^3} \sum_{\mathbf{k}\neq\mathbf{0}} \frac{4\pi^2}{k^2} e^{-k^2/4\kappa^2} \cos \mathbf{k}\cdot\mathbf{r}_{ij} - \delta_{ij} \frac{2\kappa}{\sqrt{\pi}} \right].$$

Use rRESPA split:

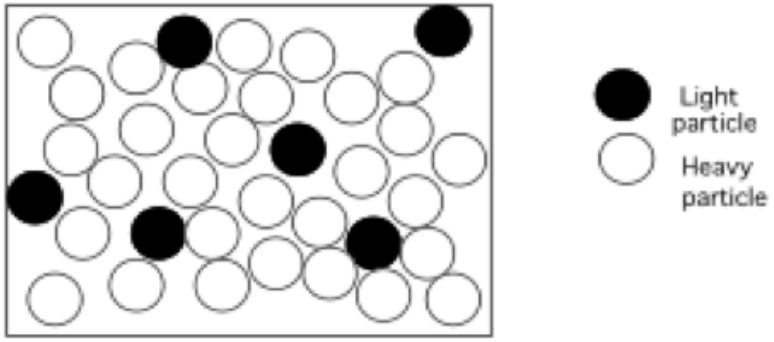
$$\mathbf{F}_{ij}^{(f)} = -\nabla_{\mathbf{r}_{ij}} V_{ij}^{\rm rs}$$

$$\mathbf{F}_{ij}^{(s)} = -\nabla_{\mathbf{r}_{ij}} V_{ij}^{\rm ks},$$

This is reasonable but not best split.

Disparate Mass System

Systems consisting of high frequency oscillators coupled to low frequency oscillators or slow baths,' and systems consisting of massive (slow moving) particles interacting with light (fast moving) particles.



System is subdivided into fast and slow degrees of freedom labeled x and y, respectively.

$$iL = iL_x + iL_y$$

$$\mathbf{iL_x} = \dot{\mathbf{x}}\frac{\partial}{\partial \mathbf{x}} + \mathbf{F_x}(\mathbf{x},\mathbf{y})\frac{\partial}{\partial \mathbf{p_x}} \qquad \mathbf{iL_y} = \dot{\mathbf{y}}\frac{\partial}{\partial \mathbf{y}} + \mathbf{F_y}(\mathbf{x},\mathbf{y})\frac{\partial}{\partial \mathbf{p_y}}$$

$$\mathbf{G_{xyx}}(\mathbf{\Delta t}) = \mathbf{e}^{\mathbf{iL_x}(\mathbf{\Delta t}/\mathbf{2})}\mathbf{e}^{\mathbf{iL_y}\mathbf{\Delta t}}\mathbf{e}^{\mathbf{iL_x}(\mathbf{\Delta t}/\mathbf{2})}$$

Systems with fast and slow degrees of freedom

Subdivision of Liouvillian $\quad \mathbf{iL = iL_f + iL_s}$

$$G_{sfs}(\Delta t) = e^{iL_s \frac{\Delta t}{2}} e^{iL_f \Delta t} e^{iL_s \frac{\Delta t}{2}}$$

$$e^{iL_f \Delta t} = \left(e^{\frac{\delta t}{2} F_f \frac{\partial}{\partial p_f}} e^{\delta t \dot{x}_f \frac{\partial}{\partial x_f}} e^{\frac{\delta t}{2} F_f (\frac{\partial}{\partial p_f}} \right)^n$$

$$e^{iL_s \Delta t/2} = e^{\frac{\Delta t}{4} F_s \frac{\partial}{\partial p_s}} e^{\frac{\Delta t}{2} \dot{x}_s \frac{\partial}{\partial x_s}} e^{\frac{\Delta t}{4} F_s \frac{\partial}{\partial p_s}} \qquad \text{where } \Delta t = n\delta t$$

Propagator

$$G_{sfs}(\Delta t) = \left(e^{\frac{\Delta t}{4} F_s \frac{\partial}{\partial p_s}} e^{\frac{\Delta t}{2} \dot{x}_s \frac{\partial}{\partial x_s}} e^{\frac{\Delta t}{4} F_s \frac{\partial}{\partial p_s}} \right)$$
$$\times \left(e^{\frac{\delta t}{2} F_f \frac{\partial}{\partial p_f}} e^{\delta t \dot{x}_f \frac{\partial}{\partial x_f}} e^{\frac{\delta t}{2} F_f (\frac{\partial}{\partial p_f}} \right)^n$$
$$\times \left(e^{\frac{\Delta t}{4} F_s \frac{\partial}{\partial p_s}} e^{\frac{\Delta t}{2} \dot{x}_s \frac{\partial}{\partial x_s}} e^{\frac{\Delta t}{4} F_s \frac{\partial}{\partial p_s}} \right)$$

Disparate Mass System

h=heavy or slow

$$v_h \leftarrow v_h + F_h \frac{\Delta t}{4m}$$
$$x_h \leftarrow x_h + v_h \frac{\Delta t}{2}$$
$$v_h \leftarrow v_h + F_h \frac{\Delta t}{4m}$$

do $i = 1, n$

l=light or fast or slow

$$v_l \leftarrow v_l + F_l \frac{\delta t}{2m}$$
$$x_l \leftarrow x_l + v_l \delta t$$
$$v_l \leftarrow v_l + F_l \frac{\delta t}{2m}$$

end do

h=heavy or slow

$$v_h \leftarrow v_h + F_h \frac{\Delta t}{4m}$$
$$x_h \leftarrow x_h + v_h \frac{\Delta t}{2}$$
$$v_h \leftarrow v_h + F_h \frac{\Delta t}{4m}$$

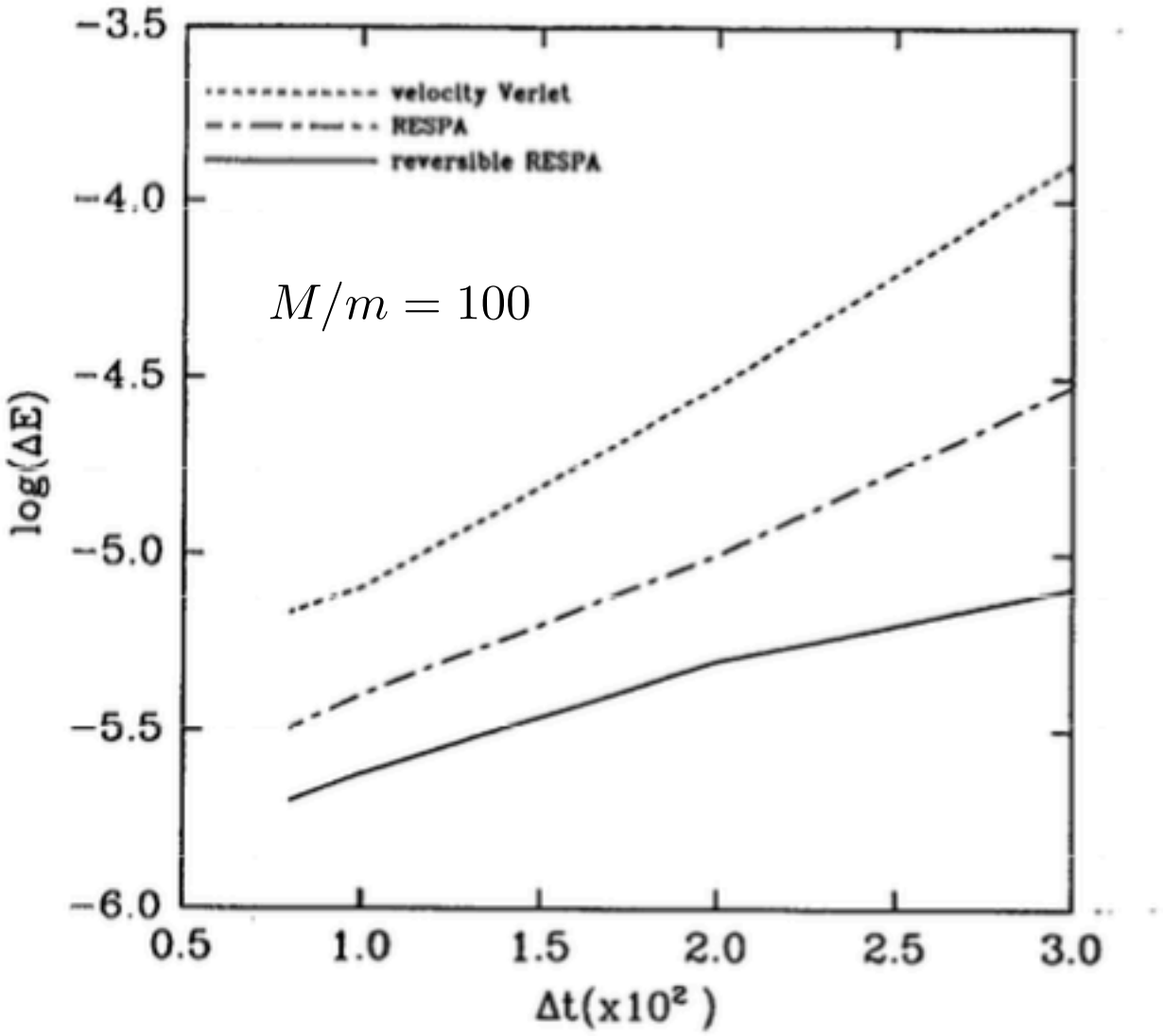
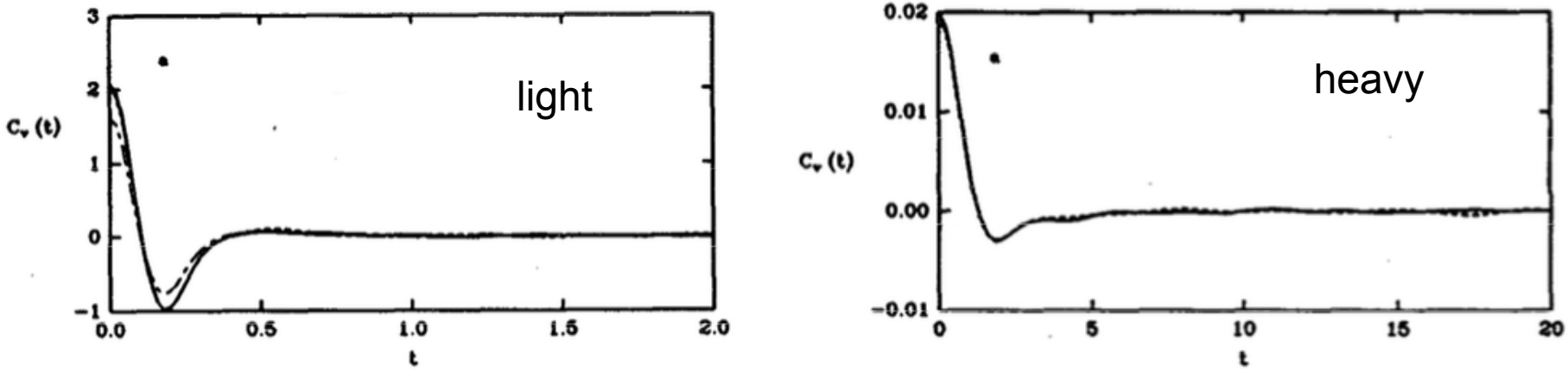


FIG. 3. Comparison of the energy conservation dependence on time step as given by Eq. (2.1) for straightforward velocity Verlet (double dashed line), nonreversible RESPA (dashed line), and the disparate mass breakup of Eq. (4.4) (bold line).

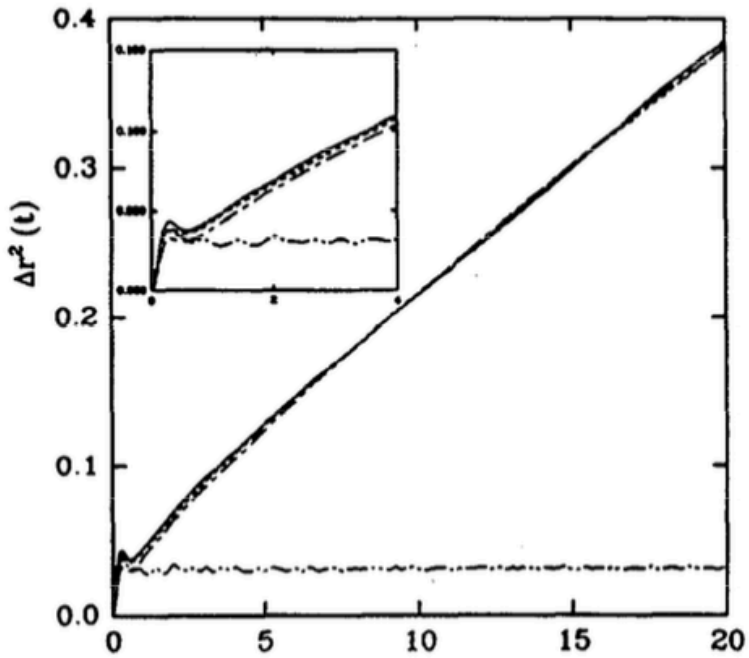
J. Chem. Phys. 97, 199-2001 (1992)

simulation of a mixture of Lennard-Jones spheres consisting of 824 heavy spheres of mass M = 100 and 40 light spheres of mass m = 1. the two time scales are $\Delta t_f = \sqrt{m_1 \sigma_1^2/\epsilon_1}$ and $\Delta t_s = \sqrt{m_2 \sigma_1^2/\epsilon_2}$

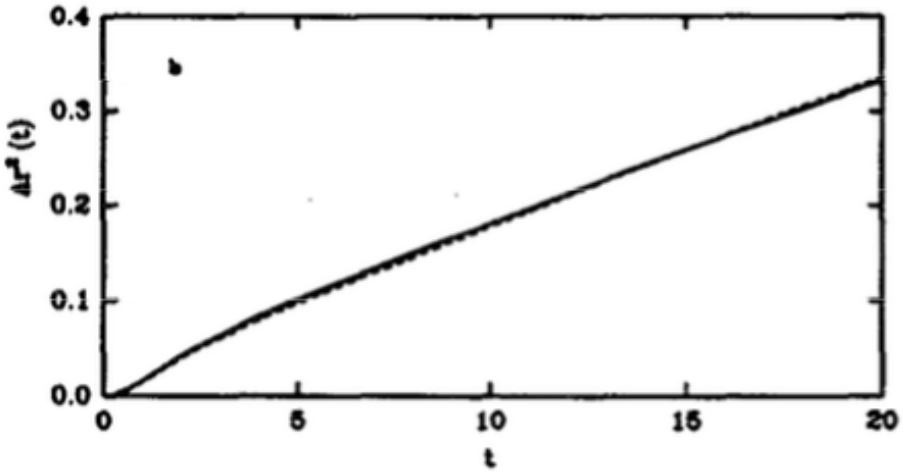


The unnormalized velocity correlation functions as a function of time for the light particles and for the heavy particles in the LennardJones mixture for RESPA (solid line), velocity Verlet (dashed line), and for TJMTS (dash-dot line).

Cannot tell difference between v-Verlet and RESPA!



The mean-square for the light for RESPA (solid line), velocity Verlet (dashed line), TJMTS (dash- dot line), and for motion of the light particles in a fixed heavy particle matrix

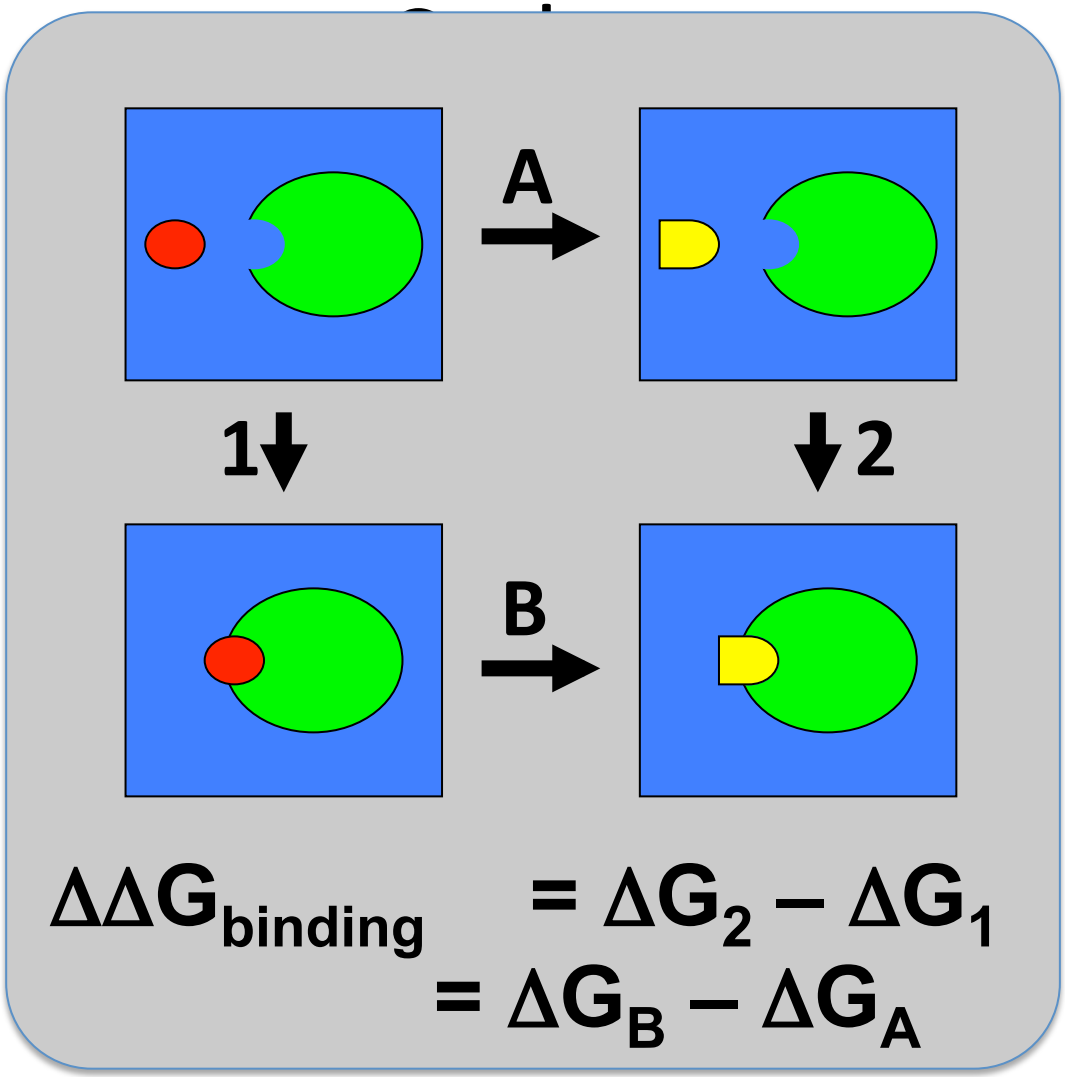


The mean square displacement for the heavy particles as computed by RESPA (solid line) and TJMTS (dashed line).

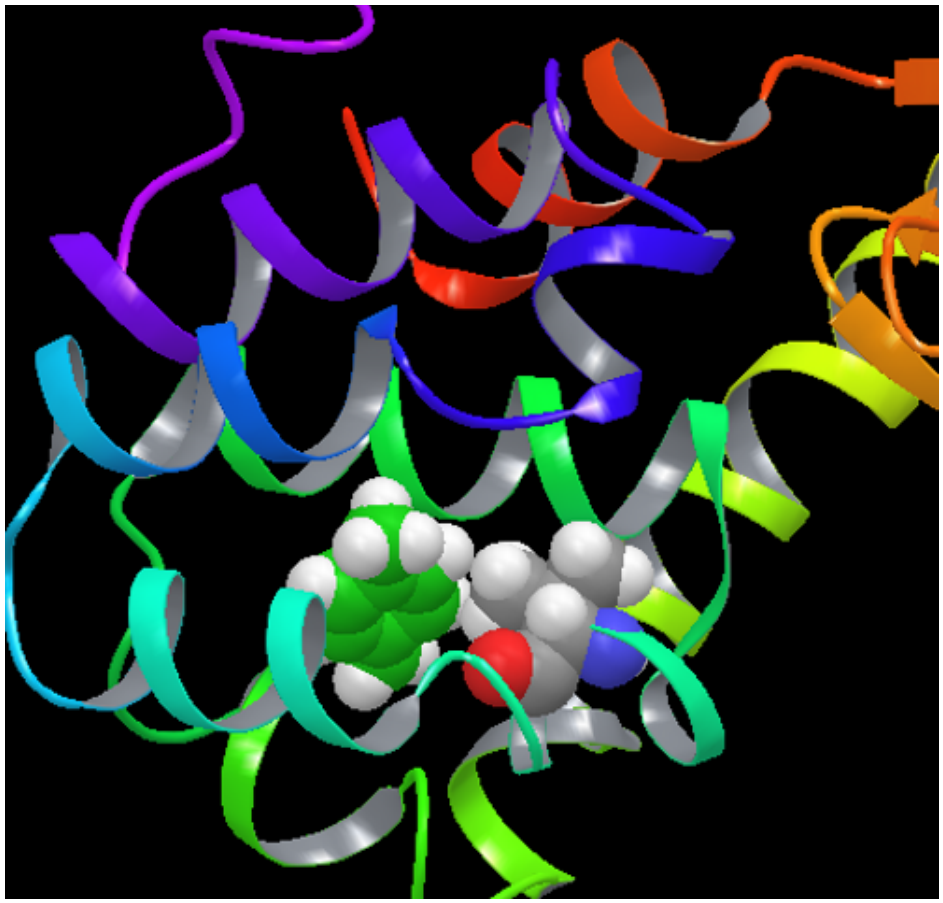
For light particles $D_{Verlet} = 2.00 \times 10^{-2}$, $D_{RESPA} = 2.00 \times 10^{-}$

For heavy particles $D = 1.72 \times 10^{-2}$.

Free Energies from Thermodynamic



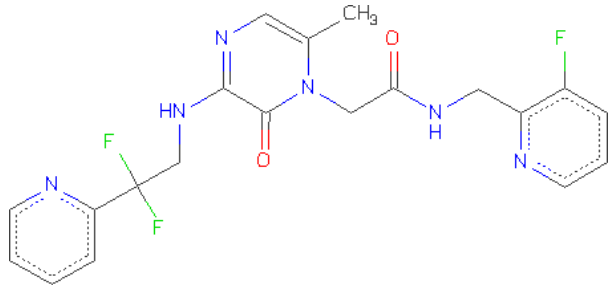
Example 1: structural reorganization in protein upon ligand binding (induced fit effect)



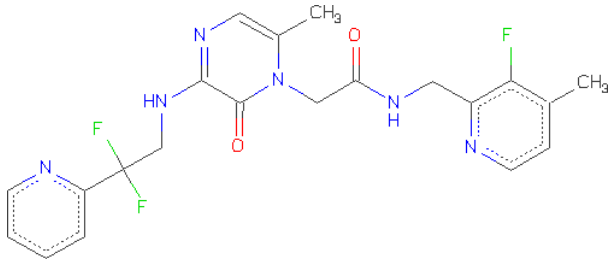
Nonpolar binding pocket of T4L/L99A
Key residue Valine 111

Conformation of Val_111
Apo: **trans**
Binding complexes of small ligands: **trans**
e.g.: Benzene, toluene
Binding complexes of larger ligands: **gauche**
e.g.: p-xylene, o-xylene

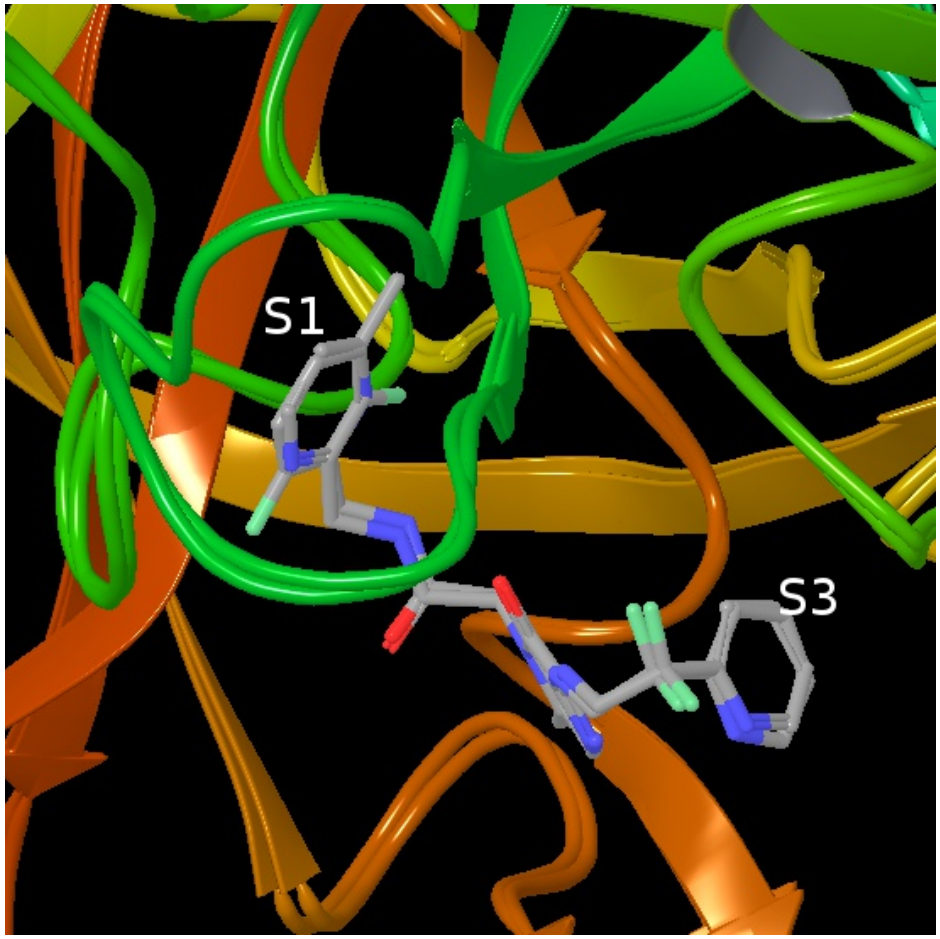
Example 2: structural reorganization in ligand upon alchemical transition



CDA



CDB



Thrombin

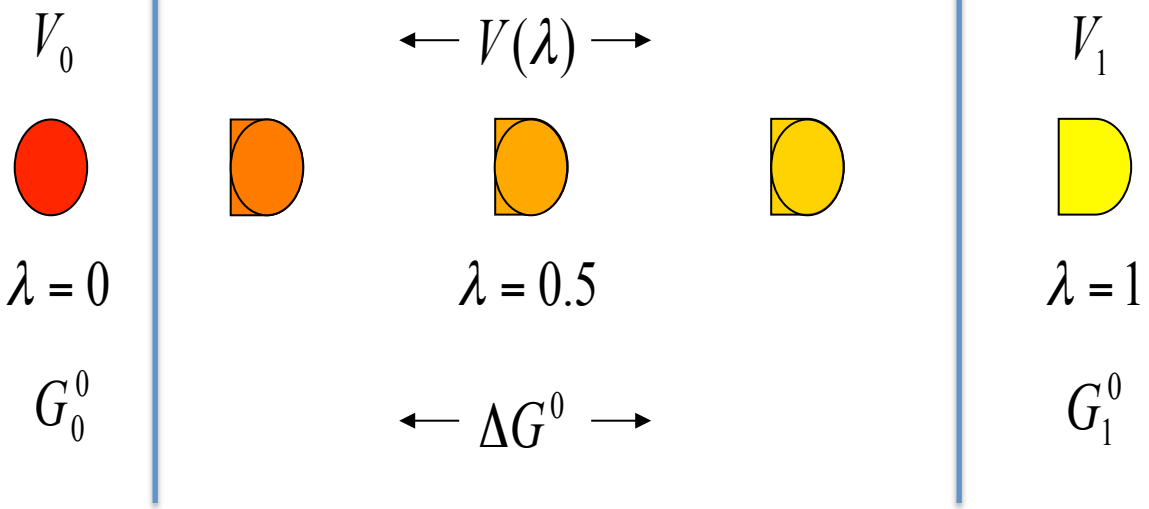
The pyridine ring flips with the addition of the methyl group.
CDA/complex: **F-out**
CDB/complex: **F-in**

Free Energy Perturbation: The Abstract λ-Coordinate

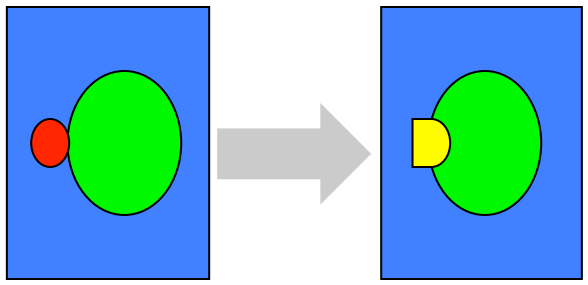
Real starting state

Alchemical intermediates

Real final state

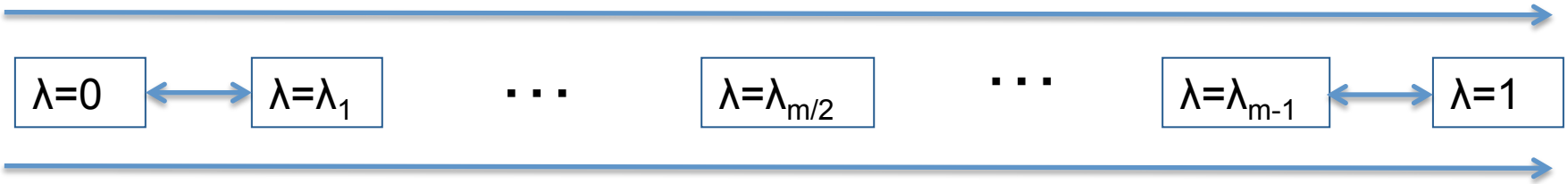


$$\Delta G^0 = G_1^0 - G_0^0 \approx -kT \ln\left\langle \exp\left(-\frac{\Delta E}{kT}\right)\right\rangle$$

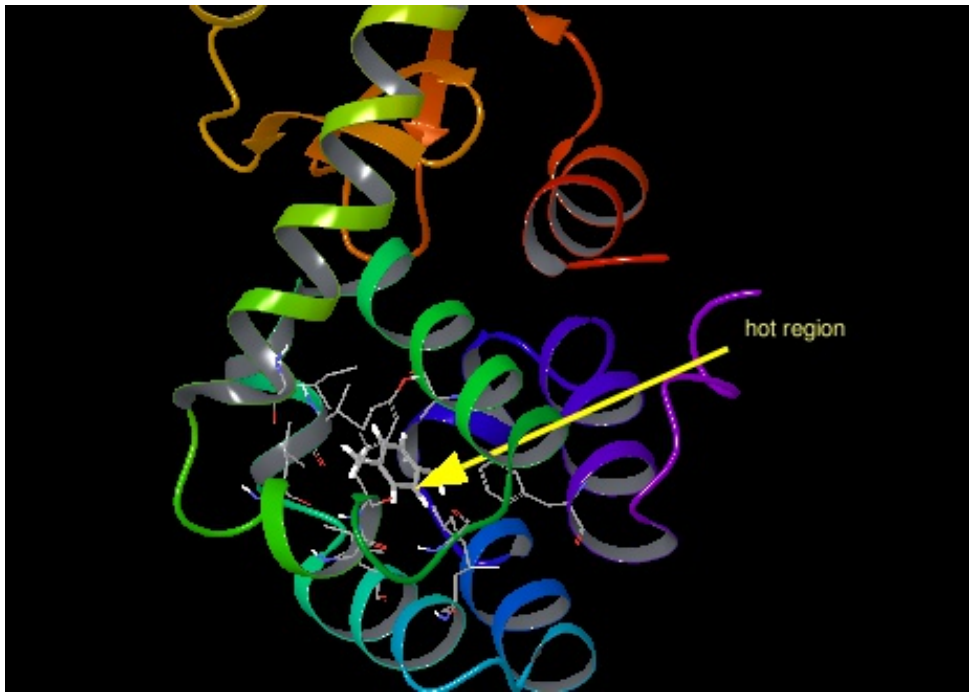


Free Energy Perturbation (FEP)

$$\textbf{For example: } \mathbf{E}(\lambda) = (1-\lambda)\mathbf{E_A} + \lambda\mathbf{E_B}$$



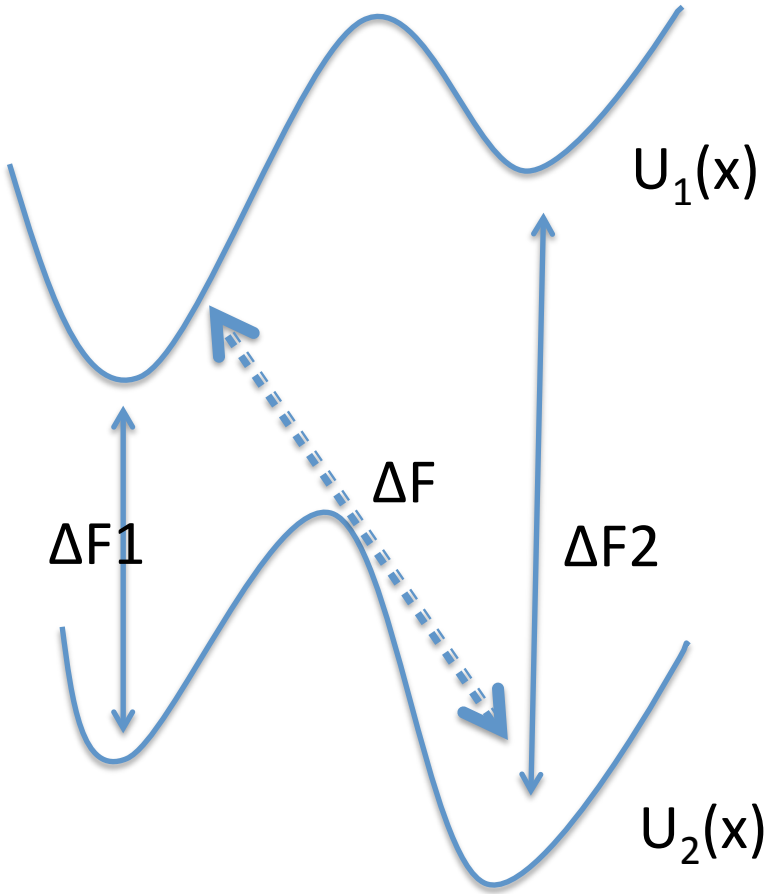
Thermodynamic axis, alchemical transition

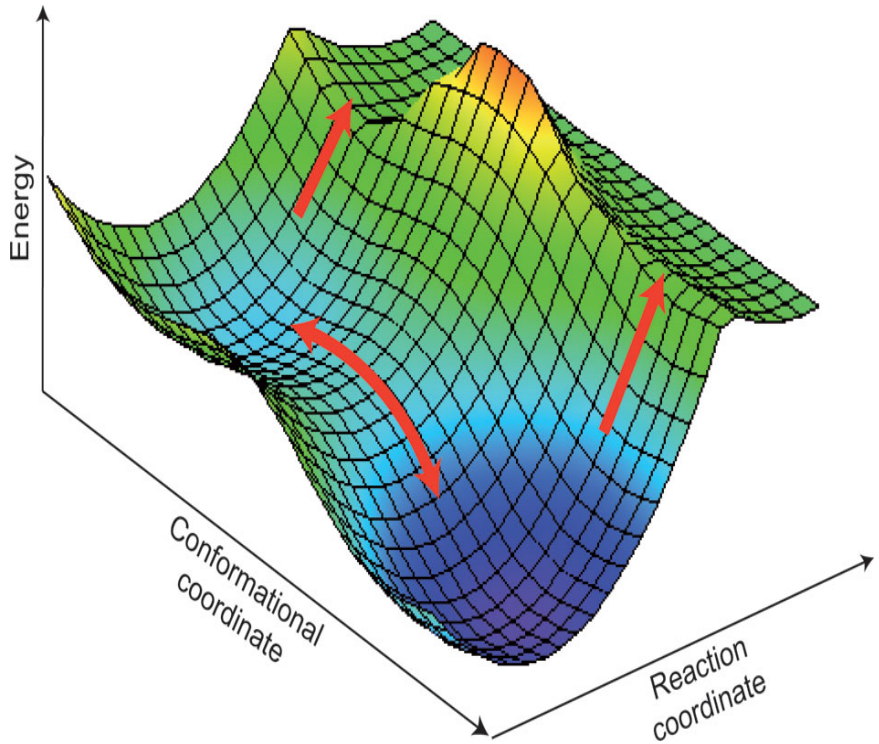


The ligand and the protein residues surrounding the binding pocket

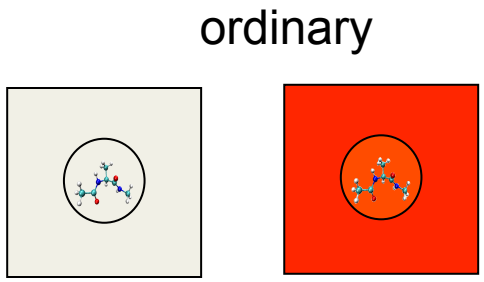
Quasi-nonergodicity problem in brute Force FEP

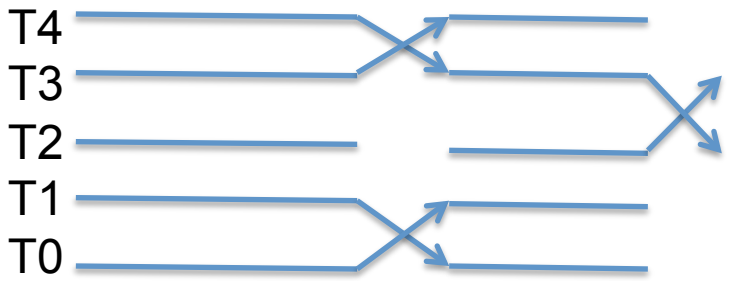
$$e^{-\beta(F_2-F_1)} = \left\langle e^{-\beta(U_2-U_1)} \right\rangle_1$$





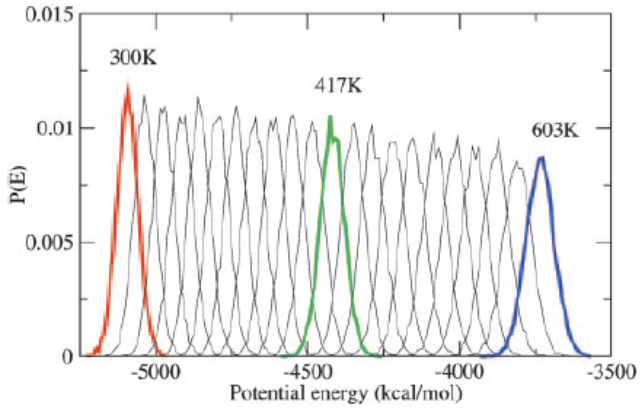
Temperature Replica exchange methods (TREM)





$$E_0(X) = E_p(X) + E_{pw}(X) + E_{ww}(X)$$

$$P(X_m;\beta_m)P(X_n;\beta_n)T(m \to n) = P(X_n;\beta_m)P(X_m\beta_n)T(n \to m)$$



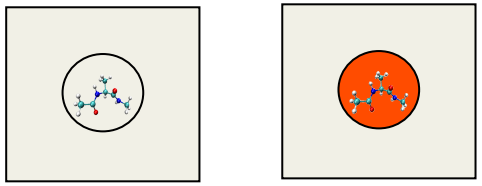
$$\frac{T(m \to n)}{T(n \to m)} = \frac{P(X_n;\beta_m)P(X_m;\beta_n)}{P(X_m;\beta_m)P(X_n;\beta_n)} = e^{-\Delta_{nm}}$$

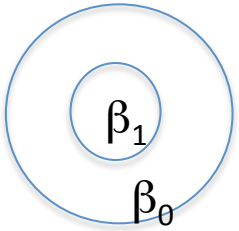
$$\Delta_{nm} = (\beta_m - \beta_n)(E(X_n) - E(X_m)) = \Delta\beta\Delta E$$

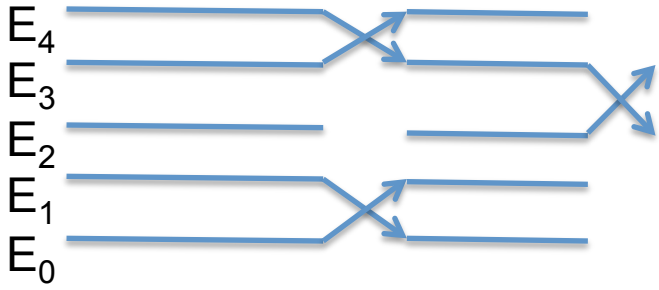
$$n \propto \sqrt{f}$$

Replica Exchange with Solute Tempering (REST)

REST



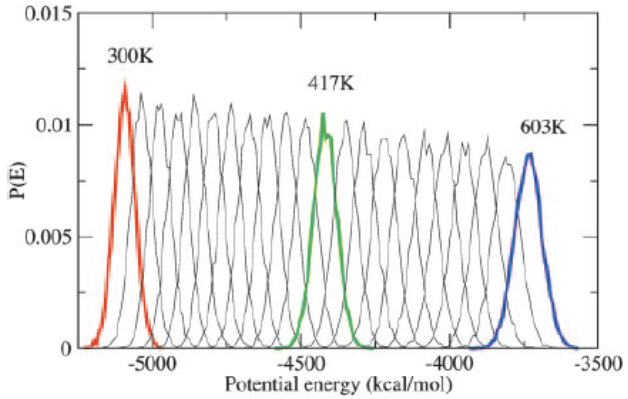




$$E_m^{\mathrm{REST2}}(X) = \frac{\beta_m}{\beta_0} E_{\mathrm{pp}}(X) + \sqrt{\frac{\beta_m}{\beta_0}} E_{\mathrm{pw}}(X) + E_{\mathrm{ww}}(X) \quad \overset{T_m \to \infty}{\longrightarrow} \mathbf{E_{ww}}$$

$$\mathbf{P(X_m;\beta_m)P(X_n;\beta_n)T(m \to n) = P(X_n;\beta_m)P(X_m\beta_n)T(n \to m)}$$

$$\frac{\mathbf{T(m \to n)}}{\mathbf{T(n \to m)}} = \frac{\mathbf{P(X_n;\beta_m)P(X_m;\beta_n)}}{\mathbf{P(X_m;\beta_m)P(X_n;\beta_n)}} = \mathbf{e^{-\Delta_{nm}}}$$



$$\Delta_{mn}(\mathrm{REST2}) = (\beta_m - \beta_n)\left[(E_{\mathrm{pp}}(X_n) - E_{\mathrm{pp}}(X_m)) + \frac{\sqrt{\beta_0}}{\sqrt{\beta_m} + \sqrt{\beta_n}}(E_{\mathrm{pw}}(X_n) - E_{\mathrm{pw}}(X_m))\right]$$

Note: The w-w interaction doesn't appear

Wang, Lingle; Friesner, R. A.; Berne, B. J. *JPCB, **115**, 943*, L.; *1–9438, 2011*
Liu P., Kim B., Friesner R.A., Berne B.J., *PNAS 2005, 13749-13754*

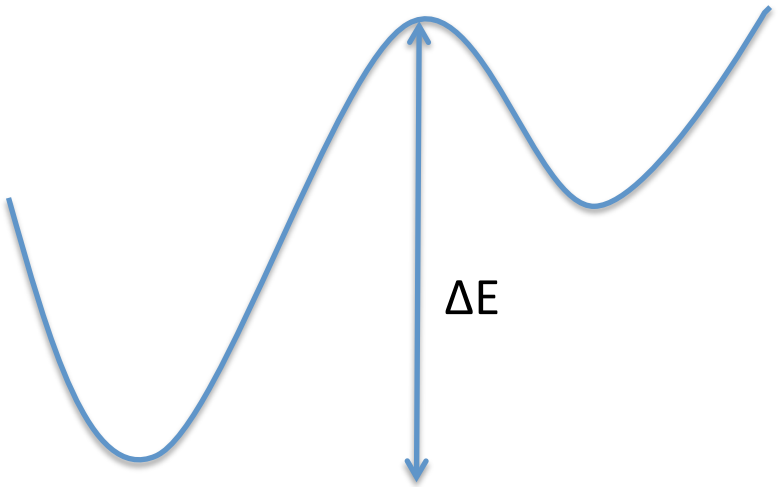
(a) In TREM high T replicas are unfolded and low temp replicas remain folded, thus no exchange between high and low temp replicas.

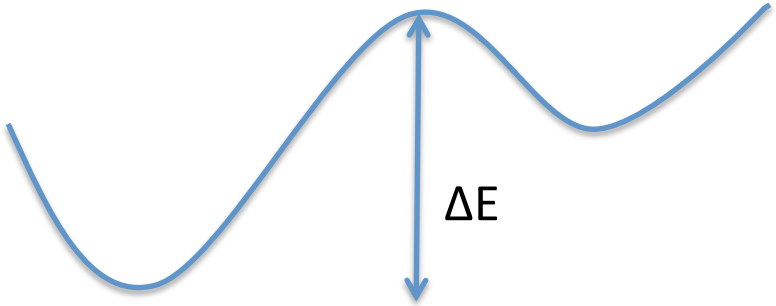
(b) In REST the high temperature replicas fold and unfold and there is very efficient replica exchange up and down the ladder.

$$\mathbf{E_m} \overset{\mathbf{T_m \to \infty}}{\longrightarrow} \mathbf{E_{ww}}$$

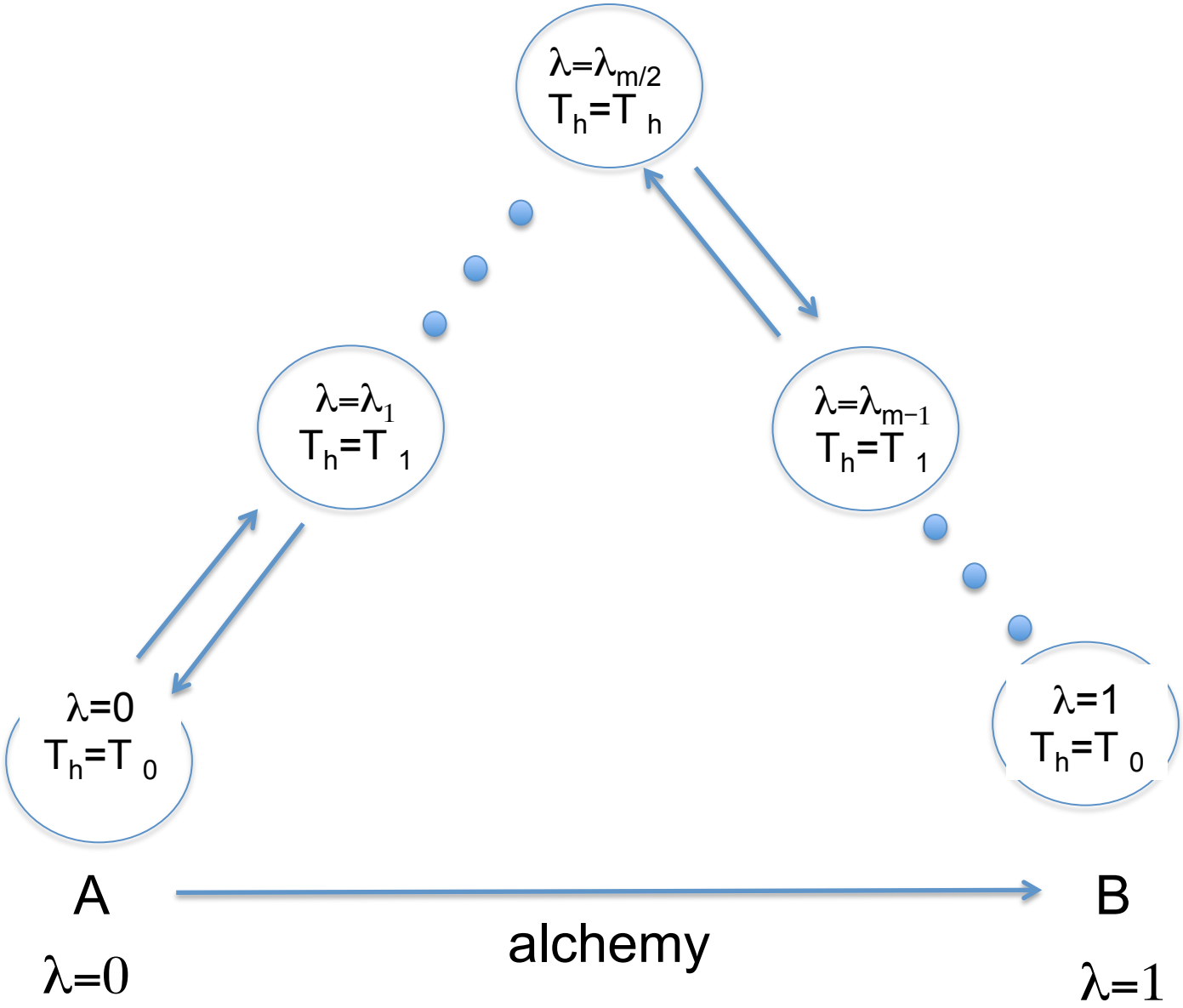
(c) In REST a much smaller number of replicas needed compared TREM.

(d) REST is excellent for ligand binding especially when combined with FEP.



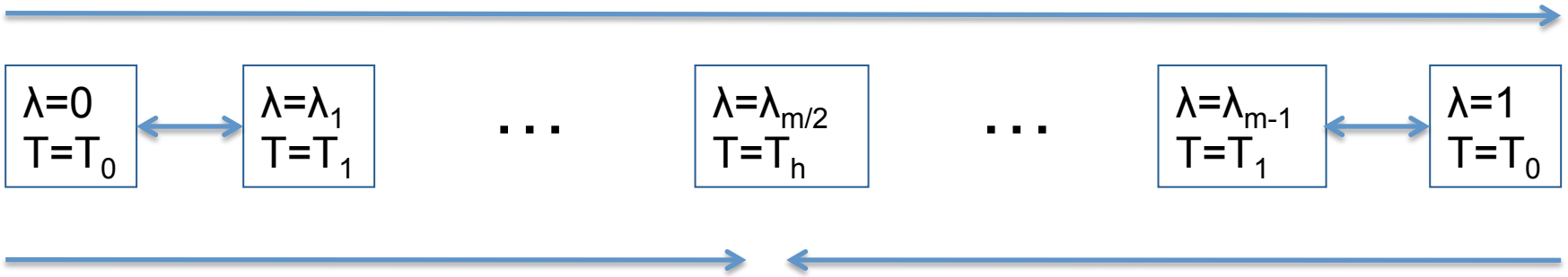


RESPA/REST

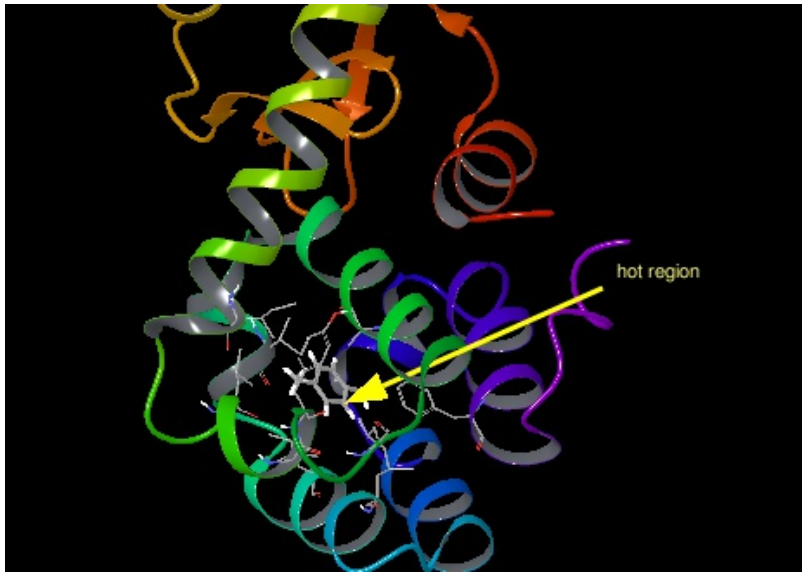


Combination of REST into FEP: FEP/REST

Thermodynamic axis, alchemical transition



Increasing the “temperature” of the hot region

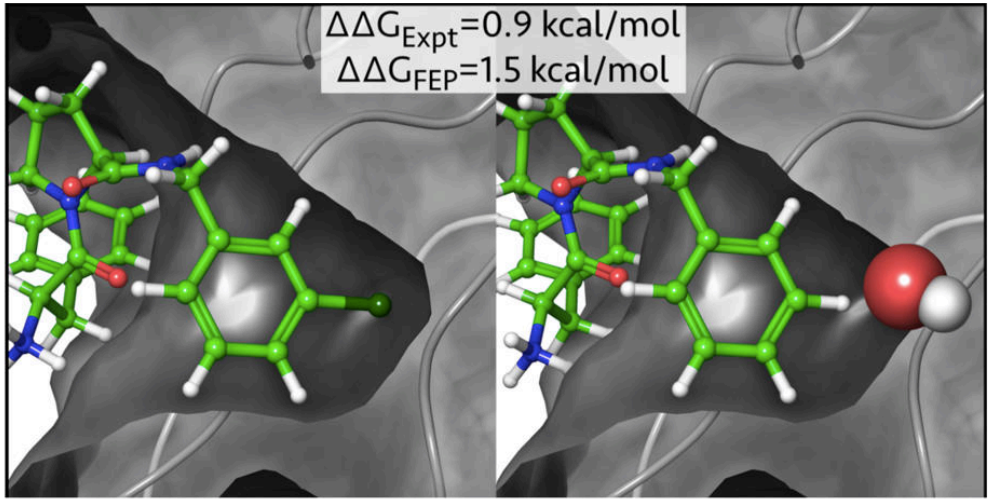


The ligand and the protein residues surrounding the binding pocket are in the hot region, the rest of the system stay cold.

Lingle Wang, Richard A. Friesner, and B. J. Berne, On achieving high accuracy and reliability in the calculation of relative protein-ligand binding affinities, Proc. Nat. Acad. Sci. USA, 109, 1937-1942 (2012)

Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field

Lingle Wang et al, J. Am. Chem. Soc. 2015, 137, 2695−2703



Here, we report an approach that achieves an unprecedented level of accuracy across a broad range of target classes and ligands, with retrospective results encompassing 200 ligands and a wide variety of chemical perturbations, many of which involve significant changes in ligand chemical structures. In addition, we have applied the method in prospective drug discovery projects and found a significant improvement in the quality of the compounds synthesized that have been predicted to be potent. Compounds predicted to be potent by this approach have a substantial reduction in false positives relative to compounds synthesized on the basis of other computational or medicinal chemistry approaches. Furthermore, the results are consistent with those obtained from our retrospective studies, demonstrating the robustness and broad range of applicability of this approach, which can be used to drive decisions in lead optimization.

Acknowledgements

- RESPA
  - Mark Tuckerman*
  - Glen Martyna
  - Ed Harder
  - Ruhong Zhou
  - Steve Stuart
  - Joe Morrone
- REST/FEP
  - Lingle Wang*
  - Piu Liu
  - Byungchan Kim
  - Richard Friesner
  - NSF, NIH, XCEDE

The End